



Alec  
Sharp

Ron  
Tolido

Mike  
Ferguson

prof.dr.  
Peter Boncz

Panos  
Alexopoulos

Jos  
van Dongen

Thomas  
Brinkman

Jan  
Henderyckx

**UTRECHT + LIVE VIDEO STREAM    27 MAART 2024**

**WORKSHOPS    28 MAART 2024**

# DATA WAREHOUSING & BUSINESS INTELLIGENCE SUMMIT 2024

**Key Changes in Data Architecture, Data Mesh Light, Gen-AI Impact on Data Management, Knowledge Graphs, Datamodelleren, Data Engineering & Analytics, Data Governance for Compliant AI**

- ▶ Data Architectuur - Key Changes
- ▶ Conceptual Data Modelling
- ▶ Hybrid Query Processing
- ▶ Impact van Generative AI op Data Management
- ▶ Data Mesh Light
- ▶ Data Governance for Compliant AI
- ▶ Mixed Source Data Engineering & Analytics
- ▶ Knowledge Graphs
- ▶ Democratisering van Data

Met internationale topsprekers en  
onafhankelijk analisten

Alec Sharp, Ron Tolido, Mike Ferguson,  
prof.dr. Peter Boncz, Panos Alexopoulos,  
Jos van Dongen, Thomas Brinkman en  
Jan Henderyckx

**INFORMATIE EN REGISTRATIE:**  
**WWW.DWBISUMMIT.COM**

## WORKSHOPS 28 MAART

- ▶ Data Products – From Design, to Build, to Publishing and Consumption | Mike Ferguson (halve dag, Engelstalig)
- ▶ Concept Modelling for Business Analysts | Alec Sharp (halve dag, Engelstalig)
- ▶ Knowledge Graphs – pragmatische aanpak en best practices | Panos Alexopoulos (halve dag, Engelstalig)



Follow us @AdeptEventsNL  
Event hash tag: #dwbisummit

**//Adept**Events



# DATA WAREHOUSING & BUSINESS INTELLIGENCE SUMMIT 2024

De DW&BI Summit is gericht op nieuwe ontwikkelingen, trends en technologieën op gebied van Datawarehousing, BI & Analytics en Data Science. Aan bod komen jaarlijks wisselende actuele en praktische thema's. De DW&BI Summit biedt voorts 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 en vertegenwoordigers van grote internationale organisaties. Profiteer van de kennis en ervaring van de eersterangs line-up van sprekers met *Alec Sharp, Ron Tolido, Mike Ferguson, prof.dr. Peter Boncz, Panos Alexopoulos, Jos van Dongen, Thomas Brinkman en Jan Henderyckx*.

Op de eerste congresdag kunt u naar onze accommodatie in Utrecht komen en presenteren alle sprekers live in de zaal. Bovendien is deze dag in het hybride format zodat u kunt kiezen tussen deelname in de zaal in Utrecht óf online via onze live video stream. De tweede congresdag bestaat uit drie workshops van een halve dag.

## Enkele onderwerpen aan bod komen

Op deze editie pakken we uit met liefst negen boeiende lezingen door topsprekers.

- Data Architectuur Ontwikkelingen - Key Changes in 2023/2024
- Conceptual Data Modelling en de Data-Process-Connection
- Hybrid Query Processing in MotherDuck
- Impact van Generative AI op Data Management, BI en Analytics
- Data Mesh Light
- Data Governance as Keystone for Compliant AI and Digital Trust

- Mixed Source Data Engineering & Analytics
- Knowledge Graphs - nieuwe mogelijkheden, kansen, uitdagingen en implementatie
- Democratisering van Data

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

## Parallel sessies en video opnames

Voor het optuigen van een optimaal en vol programma werken wij met parallel sessies. Of u nu in Utrecht deelneemt of online, u zult op 27 maart toch moeten kiezen. Sinds 2020 werken wij echter, om begrijpelijke redenen, met video opnames. Deelnemers aan het congres hebben nog enkele maanden na het congres toegang tot deze video opnames dus welke parallel sessie u ook kiest, de andere is altijd nog te bekijken.

## DW&BI SUMMIT is bestemd voor u!

Het congres is gericht op beslisseren 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 en Analytics én in de lange termijn ontwikkelingen. 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.

# PROGRAMMA



## WOENSDAG 27 MAART – LIVE + STREAMING

### Sessie 1

#### Data Architecture Evolution and the Impact on Analytics

Mike Ferguson, Managing Director, Intelligent Business Strategies Ltd

### Sessie 2A

#### Hybrid Query Processing in MotherDuck

Prof.dr. Peter Boncz, Database wetenschapper, Centrum Wiskunde & Informatica (CWI)

### Sessie 2B

#### Connecting Meaning, The promise and challenges of Knowledge

#### Graphs as providers of large-scale data semantics

Panos Alexopoulos, Founder, Panagiotis Alexopoulos

### Sessie 3A

#### Generative AI in Data Management and Analytics – A New Era of Assistance, Productivity and Automation

Mike Ferguson, Managing Director, Intelligent Business Strategies Ltd

### Sessie 3B

#### Democratisering van Data: Het Kwadrantenmodel in Actie

Thomas Brinkman, Data Architect, CJIB

### Sessie 4A

#### Data Governance as Keystone for Compliant AI and Digital Trust

Jan Henderyckx, Partner, BearingPoint

### Sessie 4B

#### Mixed Source Data Engineering & Analytics: a best of both worlds approach

Jos van Dongen, Director Erasmus Data Collaboratory, Erasmus Universiteit Rotterdam

### Sessie 5

#### Data Mesh Light – getting there, step by step, avoiding the Mess

Ron Tolido, CTO and Executive VP, Capgemini

### Sessie 6

#### Concept Modelling and The Data-Process Connection

Alec Sharp, Founder, Clariteq Systems Consulting

## DONDERDAG 28 MAART – WORKSHOPS

### 9:00 – 12:30

#### Data Products – From Design, to Build, to Publishing and Consumption

Mike Ferguson, Managing Director, Intelligent Business Strategies Ltd

### 9:00 – 12:30

#### Concept Modelling for Business Analysts

Alec Sharp, Founder, Clariteq Systems Consulting

### 13:30 – 17:00

#### Knowledge Graphs – pragmatische aanpak en best practices

Panos Alexopoulos, Founder, Panagiotis Alexopoulos

### Globale dagindeling voor 27 maart:

09:00 – 09:15 Opening

09:15 – 10:15 Sessie 1

10:15 – 10:30 Koffiepauze

10:30 – 11:30 Sessie 2A en 2B

11:30 – 12:30 Sessie 3A en 3B

12:30 – 13:30 Lunchpauze

13:30 – 14:30 Sessie 4A en 4B

14:30 – 15:30 Sessie 5

15:30 – 15:45 Koffiepauze

15:45 – 16:45 Sessie 6

16:45 – 16:50 Afsluiting

16:50 Borrel



## 1. Data Architecture Evolution and the Impact on Analytics

**Mike Ferguson, Managing Director, Intelligent Business Strategies Ltd**

In the last 12-18 months we have seen many different architectures emerge from many different vendors who claim to be offering 'the modern data architecture solution' for the data-driven enterprise. These range from streaming data platforms to data lakes, to cloud data warehouses supporting structured, semi-structured and unstructured data, cloud data warehouses supporting external tables and federated query processing, lakehouses, data fabric, and federated query platforms offering virtual views of data and virtual data products on data in data lakes and lakehouses. In addition, all of these vendor architectures are claiming to support the building of data products in a data mesh. It's not surprising therefore, that customers are confused as to which option to choose.

However, in 2023, key changes have emerged including much broader support for open table formats such as Apache Iceberg, Apache Hudi and Delta Lake in many other vendor data platforms. In addition, we have seen significant new milestones in extending the ISO SQL Standard to support new kinds of analytics in general purpose SQL. Also, AI has also advanced to work across any type of data.

The key question is what does this all mean for data management? What is the impact of this on analytical data platforms and what does it mean for customers? This session looks at this evolution and helps customers realise the potential of what's now possible and how they can exploit it for competitive advantage.

- The demand for data and AI
- The need for a data foundation to underpin data and AI initiatives
- The emergence of data mesh and data products
- The challenge of a distributed data estate
- Data fabric and how can they help build data products
- Data architecture options for building data products
- The impact of open table formats and query language extensions on architecture modernisation
- Is the convergence of analytical workloads possible?

## 2A. Hybrid Query Processing in MotherDuck

**Prof.dr. Peter Boncz, Database wetenschapper, Centrum Wiskunde & Informatica (CWI)**

MotherDuck is a new service that connects DuckDB to the cloud. It introduces the concept of "*hybrid query processing*": the ability to execute queries partly on the client and partly in the cloud. The talk covers the motivation for MotherDuck and some of its use cases; as well as the main characteristics of its system

architecture, which heavily uses the extension mechanisms of DuckDB. To provide context, the talk will therefore also provide a brief overview of the DuckDB architecture.

- DuckDB
- History: MonetDB, VectorWise, Snowflake
- MotherDuck: DuckDB in the cloud
- Hybrid Query Processing
- Applications: Data Teams & Low-latency Web Analytics.

## 2B. Connecting Meaning, The promise and challenges of Knowledge Graphs as providers of large-scale data semantics

**Panos Alexopoulos, Founder, Panagiotis Alexopoulos**

Ever since Google announced that "their knowledge graph allowed searching for things, not strings", the term "knowledge graph" has been widely adopted, to denote any graph-like network of interrelated typed entities and concepts that can be used to integrate, share and exploit data and knowledge.

This idea of interconnected data under common semantics is actually much older and the term is a rebranding of several other concepts and research areas (semantic networks, knowledge bases, ontologies, semantic web, linked data etc). Google popularized this idea and made it more visible to the public and the industry, the result being several prominent companies, developing and using their own knowledge graphs for data integration, data analytics, semantic search, question answering and other cognitive applications.

As the use of knowledge graphs continues to expand across various domains, the need for ensuring the accuracy, reliability, and consensus of semantic information becomes paramount. The intricacies involved in constructing and utilizing knowledge graphs present a spectrum of challenges, from data quality assurance to ensuring scalability and adaptability to evolving contexts.

In this talk, we will delve deeper into the significance of knowledge graphs as facilitators of large-scale data semantics. The discussion will encompass the core concepts, challenges, and strategic considerations that architects and decision-makers encounter while initiating and implementing knowledge graph projects.

The session will cover:

- Understanding Knowledge Graphs: Exploring the fundamental concepts and significance of knowledge graphs in integrating, organizing, and harnessing data across diverse domains
- Challenges in Building Knowledge Graphs: Identifying and dissecting primary hurdles such as data quality assurance, schema alignment, scalability, and ongoing maintenance



- Strategic Dilemmas: Examining critical decision points and dilemmas faced by architects and executives when designing and executing knowledge graph initiatives
- Crafting an Effective Strategy: Outlining guidelines to formulate a robust knowledge graph strategy tailored to specific organizational goals, considering scalability, interoperability, and domain relevance.

### 3A. Generative AI in Data Management and Analytics – A New Era of Assistance, Productivity and Automation

**Mike Ferguson, Managing Director, Intelligent Business Strategies Ltd**

The emergence of generative AI has been described as a major breakthrough in technology. It has reduced the time to create new content and triggered a new wave of innovation that is impacting almost every type of software. New tools, applications and functionality are already emerging that are dramatically improving productivity, simplifying user experiences and paving the way for new ways of working. In this keynote session, Mike Ferguson, Europe's leading IT industry analyst on Data Management and Analytics, looks at the impact generative AI is having on Data Management, BI and Data Science and what it can do to help shorten time to value.

- What is generative AI?
- What are the business benefits of generative AI?
- How is generative AI being used in data management?
- How is generative AI being used in data science and BI
- What does this mean for business going forward?
- What should you do to get started?

### 3B. Democratisering van Data: Het Kwadrantenmodel in Actie

**Thomas Brinkman, Data Architect, CJIB**

Traditioneel zijn datawarehouses primair ontworpen voor het oplossen van analysevraagstukken. Met de opkomst van data-democratisering groeit de behoefte om data breder binnen organisaties in te zetten. Dataconsumenten willen de beschikbare gegevens vrijer benutten, en historische data in datawarehouses wordt steeds waardevoller als bron voor het trainen van AI-modellen. In dit evoluerende landschap wordt het integreren van privacy by design in de architectuur essentieel. Het moet niet langer worden gezien als een hindernis, maar eerder als een katalysator voor deze vooruitgang. Het kwadrantenmodel van Damhof biedt hierbij een leidraad. Door deze benadering toe te passen, ontstaat niet alleen de mogelijkheid om te voldoen aan de groeiende eisen van dataconsumptie en AI-ontwikkelingen, maar leggen we ook een solide basis waarop innovatie wordt gestimuleerd.



- Datawarehouses en de rol binnen datasience
- Privacy by Design als katalysator
- Kwadrantenmodel in combinatie met datavirtualisatie
- Kostenreductie van experimenten.

### 4A. Data Governance as Keystone for Compliant AI and Digital Trust

**Jan Henderyckx, Partner, BearingPoint**

Data governance is the process of managing the availability, usability, integrity, and security of data in an organization. It is essential for ensuring that data is used ethically, responsibly, and in compliance with regulations and standards. Data governance also enables the development and deployment of AI systems that are aligned with the values, goals, and expectations of the stakeholders and the society. In this keynote, we will discuss how data governance can serve as a keystone for building ethical AI and digital trust. We will explore the challenges and opportunities of data governance in the context of AI, and present some best practices and frameworks for implementing data governance in AI projects. We will also share some examples and case studies of how data governance can help achieve ethical AI and digital trust outcomes. The keynote will conclude with some recommendations and future directions for data governance in the AI era.

By the end of this session, you will be able to:

- Define data governance and its importance for data and AI systems
- Identify the challenges and opportunities of data governance in the context of AI
- How to apply best practices and frameworks for data governance, such as data lifecycle management, data stewardship, data ethics principles, and data audit and assessment
- Explain how data governance can support ethical AI and digital trust outcomes, such as fairness, privacy, explainability, and reliability
- Recognize the roles and responsibilities of various actors and stakeholders in the AI ecosystem for data governance.

### 4B. Mixed Source Data Engineering & Analytics: a best of both worlds approach

**Jos van Dongen, Director Erasmus Data Collaboratory, Erasmus Universiteit Rotterdam**

Erasmus University Rotterdam (EUR) is one of the largest academic institutions of the country whose mission is 'creating a positive societal impact', and where the United Nations Sustainable Development Goals serve as a compass for research and education alike. With the variety and diversity of topics within EUR, an open, flexible, affordable, and easy to use data & analytics solution is key to support data & AI projects. At the same time there are



many internal and external factors that need to be considered: the adoption of and migration to cloud solutions, the push for open science and open source, an ever faster changing technology landscape, and finally the breathtaking speed with which AI solutions are coming to market. Making future proof choices in this environment is a daunting task as one could imagine. Nevertheless, choices have been made and consist of a mix of open source and proprietary solutions, both on-premise and in the cloud, and guided by modern software engineering principles. This session will highlight the following:

- The influence of modern software engineering principles like CI/CD on data engineering, data management, and analytics
- How to remain independent and prevent lock in from any vendor or cloud provider
- The tradeoff between building, buying, and renting hard and software
- How to standardize on tools and technology and remain flexible at the same time.

## 5. Data Mesh Light – getting there, step by step, avoiding the Mess

**Ron Tolido, CTO and Executive VP, Capgemini**

The Data Mesh approach has been well on its way as an alternative data management approach that does justice to the federative nature of most organizations and the need to provide ownership of data as close as possible to the business domains – where data is actually created and used. However, the transformational impact of Data Mesh is potentially big, and many organizations have found it difficult to implement the approach in all of its dimensions at once. Why not take a lighter approach, reaping benefits one by one, rather than going for an unprepared, deep dive into the Data Mesh rabbit hole?

- Recap: the key elements of the Data Mesh approach
- Best and worst practices from real life
- Crafting a step-by-step approach

- Architectural and technological considerations
- Adding semantics to the Data Mesh
- Using generative AI to augment a Data Mesh.

## 6. Concept Modelling and The Data-Process Connection

**Alec Sharp, Founder, Clariteq Systems Consulting**

Whether you call it a conceptual data model, a domain map, a business object model, or even a “thing model,” a concept model is invaluable to process and architecture initiatives. Why? Because processes, capabilities, and solutions act on “things” – Settle Claim, Register Unit, Resolve Service Issue, and so on. Those things are usually “entities” or “objects” in the concept model, and clarity on “what is one of these things?” contributes immensely to clarity on what the corresponding processes are.

After introducing methods to get people, even C-level executives, engaged in concept modelling, we’ll introduce and get practice with guidelines to ensure proper naming and definition of entities/concepts/business objects. We’ll also see that success depends on recognising that a concept model is a description of a business, not a description of a database. Another key – don’t call it a data model!

Drawing on almost forty years of successful modelling, on projects of every size and type, this session introduces proven techniques backed up with current, real-life examples. Topics include:

- Concept modelling essentials – things, facts about things, and the policies and rules governing things
- “Guerrilla modelling” – how to get started on concept modelling without anyone realising it
- Naming conventions and graphic guidelines – ensuring correctness, consistency, and readability
- Concept models as a starting point for process discovery
- Practical examples of concept modelling supporting process work, architecture work, and commercial software selection.



# INTERNATIONALE TOPSPREKERS



**ALEC SHARP**, senior consultant bij Clariteq Systems Consulting, heeft diepgaande expertise in een zeldzame combinatie van werkgebieden – gespreksleiding, strategie ontwikkeling, specificatie van applicatie-requirements, datamodellering en, uiteraard,

business proces analyse en herontwerp, van verhoogde verbetering door zakelijke verandering. Zijn 35 jaar ervaring in hands-on advies, praktische benaderingen en wereldwijde reputatie in modelgedreven methodieken hebben hem gemaakt tot een veelgevraagd adviseur in uiteenlopende locaties als Ierland, Illinois en India.

Hij is tevens een populaire spreker op conferenties gerelateerd aan Business Process Management, Business Analyse en Data Management, waar hij inhoud en inzicht mixt met lichtvoetigheid en humor. Tot zijn vele hooggewardeerde presentaties behoren "The Lost Art of Conceptual Modelling," "The Human Side of Data Modelling," "Getting Traction for 'Process' – What the Experts Forget," "Mind the Gap! – Integrating Process, Data, and Requirements Modelling" en "Adventures in Reverse-Engineering – What You've Got and Why You Don't Like It."

Alec Sharp is ontvanger van de prestigieuze DAMA Professional Achievement Award, een internationale prijs die eens per jaar wordtgereikt aan de professional die het meest heeft bijgedragen aan het Data Management vakgebied.

Alec schreef letterlijk hét boek over business process modeling – hij is auteur van de tweede editie van "Workflow Modeling: Tools for Process Improvement and Application Development". Het boek is populair bij professionals op het gebied van procesverbetering, business analisten en consultants, is aanhoudend bestseller op het gebied van business process modelleren en wordt veel gebruikt als MBA-leerboek.

Alec's educatieve workshops worden virtueel en in-person uitgevoerd bij vele bekende organisaties. Deze omvatten *Business-Oriented Data Modelling*, *Business-Oriented Data Modelling – Masterclass*, *Working With Business Processes*, *Advanced Business Process Techniques*, alsmede *Model-Driven Business Analysis Techniques*. Zijn lessen zijn praktisch, energiek en leuk, en krijgen stevast uitstekende beoordelingen.



**RON TOLIDO** is Executive Vice President, CTO en Master Architect van Capgemini's Insights & Data global business line. Hij publiceerde verschillende boeken en talrijke artikelen en berichten over innovatie, AI, data-architectuur en door technologie aangedreven

transformatie. Hij is de hoofdauteur van de jaarlijkse trendreeks 'TechnoVision' van Capgemini en de hoofdredacteur van het tijdschrift 'Data-powered Innovation Review'. Ron spreekt vaak en is gastdocent aan de TIAS Business School. Hij is tien jaar lang lid geweest van de bestuursraad van The Open Group.



**MIKE FERGUSON** is oprichter van Intelligent Business Strategies Ltd. en als analist en consultant gespecialiseerd in business intelligence, big data, data management en enterprise business integration. Hij kan bogen op meer dan 40 jaar ervaring in IT, ondermeer

op gebied van BI en Corporate Performance Management, Data Management en Big Data Analytics (Hadoop, MapReduce, Hive, Graph DBMSs). Mike opereert afwisselend op bestuursniveau, IT management niveau en ook gespecialiseerde technische IT niveau's voor de terreinen BI, corporate performance management strategie, technologie- en toolselectie, enterprise architectuur, MDM and data-integratie. Hij is een veelgevraagd spreker op internationale conferenties en heeft veelvuldig artikelen gepubliceerd in de vakbladen en via weblogs.

Eerder was Mike partner en mede-oprichter van Codd and Date Europe Limited, Chief Architect bij NCR voor het Teradata DBMS en Europees directeur van Database Associates. Hij verzorgt regelmatig seminars en workshops met als onderwerpen Data Warehouse Modernisation, Big Data Architecture & Technology, Centralised Data Governance of a Distributed Data Landscape, Practical Guidelines for Implementing a Data Mesh, Embedded Analytics, Intelligent Apps & AI Automation, Migrating Data Warehouse to the Cloud, Modern Data Architecture and Data Virtualisation.

Mike heeft gesproken op eerdere edities van ons jaarlijkse congres Datawarehousing & Business Intelligence Summit.



**JAN HENDERYCKX** is Partner bij BearingPoint waar hij verantwoordelijk is voor het Belgisch kantoor en het Data Governance centre of excellence leidt. Hij is sinds 1986 gefocust op data. Hij heeft op vele internationale congressen en usergroup-bijeenkomsten over de gehele wereld presentaties gegeven en workshops gehouden, of als moderator opgetreden. Als consultant bij talloze bedrijven, variërend van Fortune top 500 tot kleine ondernemingen, is Jan getuige geweest van de interne gang van zaken bij een scala aan ondernemingen. Deze ervaringen gecombineerd met zijn deskundigheid op het gebied van data strategy het data centric maken van organisaties, komen goed van pas als u worstelt met het behalen van waarde uit uw data assets.

Jan heeft gesproken op eerdere edities van ons jaarlijkse congres Datawarehousing & Business Intelligence Summit.



**PETER BONCZ** is al vier decennia actief in de databasegemeenschap, wat hem tot een veteraan maakt. Hij leidt de onderzoeksgroep Database Architecturen van onderzoeksinstiutuut Centrum Wiskunde & Informatica (CWI) in Amsterdam en is tot nu toe betrokken geweest bij zes startups. Hij is onlangs benoemd tot ACM Fellow voor zijn bijdragen aan moderne database architecturen en is tevens hoogleraar aan de Vrije Universiteit in Amsterdam, gespecialiseerd in analytische databases. Hij is ook de oprichter en voorzitter van de grafiekdatabase organisatie Linked Data Benchmark Council (LDBC), hoewel hij dit jaar met verlof is van deze laatste functie, tijdens zijn sabbatical verblijf bij MotherDuck.

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 architecturbewust 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.

Peter Boncz heeft een sterke staat van dienst in het overbruggen van de kloof tussen de academische wereld en commerciële toepassingen. Hij 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 netwerk databases, en databases die gebruik kunnen maken van heterogene processors en moderne opslagmedia.



**JOS VAN DONGEN** is directeur van het Erasmus Data Collaboratory – House of AI, waar data & AI samen komen. Het EDC is onderdeel van het Erasmus Centre for Data Analytics, waarin Jos zitting heeft als lid van het management team. Voordat hij in juli 2023 bij het ECDA kwam, werkte hij als analytics adviseur en architect bij SAS Institute. Jos is consultant, docent en data analytics expert sinds 1991. Sinds 2006 schrijft en presenteert hij over nieuwe ontwikkelingen op datagebied en spreekt regelmatig op nationale en internationale conferenties.



**PANOS ALEXOPOULOS** werkt sinds 2006 op het snijvlak van data, semantiek en software en draagt bij aan de bouw van intelligente systemen die waarde leveren aan het bedrijfsleven en de maatschappij. Panos werkt momenteel als Hoofd Ontologie bij Textkernel BV in Amsterdam, waar hij een team van data professionals leidt bij het ontwikkelen en opleveren van een grote cross-lingual Knowledge Graph voor het HR en Recruitment werkveld. Daarnaast ontwikkelt en verzorgt hij praktische trainingsworkshops over datasemantiek, natuurlijke taalverwerking en kunstmatige intelligentie. Panos heeft verschillende artikelen gepubliceerd op internationale conferenties, in tijdschriften en boeken, en hij is een regelmatige spreker op zowel academische als bedrijfspodiums. Daarbij streeft hij ernaar om de kloof tussen de academische wereld en het bedrijfsleven te overbruggen zodat ze van elkaar kunnen profiteren. Hij is auteur van het O'Reilly boek "Semantic Modeling for Data – Avoiding Pitfalls and Dilemmas", een praktische handleiding voor data professionals die willen leren hoe semantische datamodellering wordt toegepast.

0101001010100101  
010010101010101010  
010000110110110101



**THOMAS BRINKMAN** is momenteel werkzaam in het Innovatielab van het CJIB, waar hij zich verdiept in de rol van architectuur in combinatie met maatschappelijke vraagstukken. Als deels zelfstandig ondernemer deelt hij zijn kennis en ervaring graag met anderen. Thomas is vanaf jonge leeftijd gefascineerd door computers. In 2001 begon hij zijn carrière als database-ontwikkelaar, voornamelijk werkend aan ERP-applicaties zoals PeopleSoft. Al snel zag hij de noodzaak van datawarehousing voor geïntegreerde data-analyses. Enkele jaren later greep hij zijn kans in de vroege dagen van servervirtualisatie, gedreven door zijn liefde voor nieuwe technologieën. Desondanks keerde Thomas terug naar het domein van Business Intelligence, dit keer bij Norsk-Tipping. Zijn technische achtergrond, gecombineerd met de ervaring in BI en zijn betrokkenheid bij maatschappelijke vraagstukken, resulteerden in de ontwikkeling van steeds meer architecturale oplossingen.

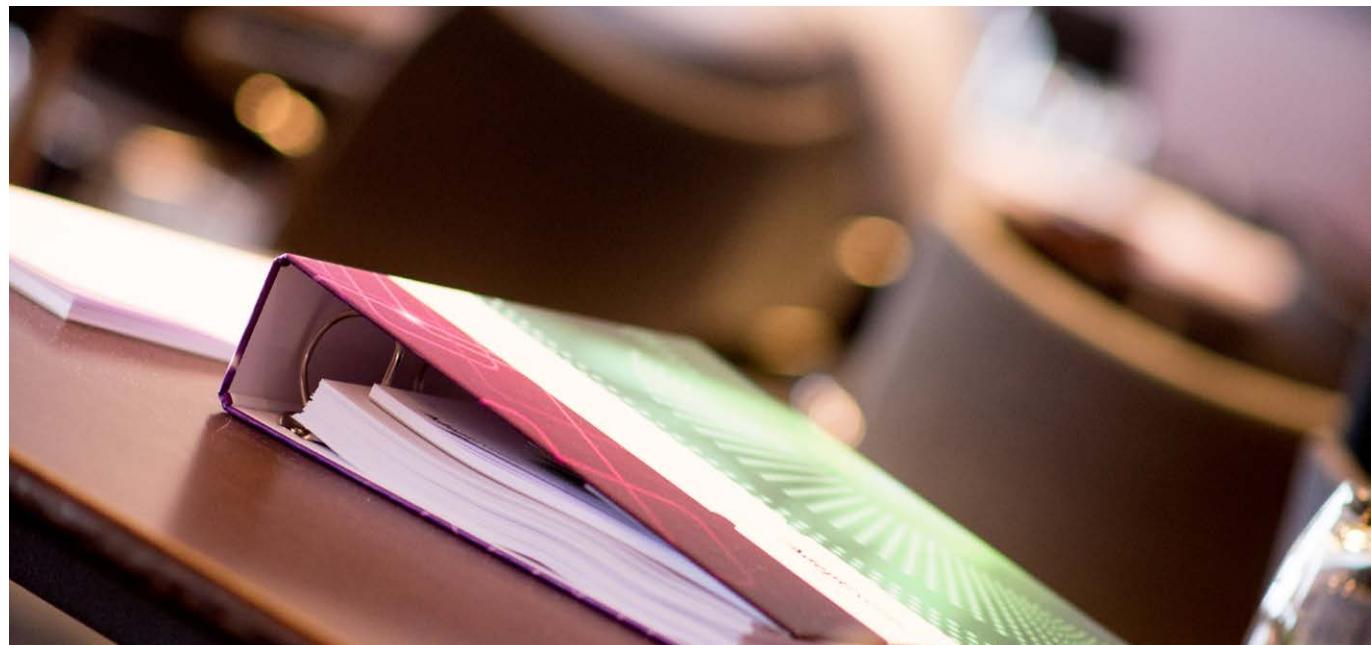
## CONGRES-APP

### DW&BI SUMMIT Download de DW&BI Summit Congres-App

(onderdeel van de BI-Platform App)

The screenshot shows the mobile application interface for the DW&BI Summit. At the top, it displays the event name "DW & BI Summit" and the date "27 maart 2024". Below this is a navigation bar with icons for Home, Schedule, Speakers, and Partners. The main content area is a scrollable list of sessions:

- 09:00 - 09:15 Opening** (Plenair, Zaal 1) by Werner Schmitz. Includes a "See Live Stream" button.
- 09:15 - 10:15 Data Architecture Evolution and the Impact on Analytics** (Zaal 1) by Mike Ferguson. Includes a detailed description of the session's focus on changes in open table formats like Apache Iceberg, Apache Flink, and Delta Lake, and how AI has advanced in this field.
- 10:30 - 11:30 Hybrid Query Processing in MotherDuck** (Zaal 1) by Peter Boers. Includes a description of MotherDuck as a service connecting DuckDB to the cloud, introducing hybrid query processing.
- 10:30 - 11:30 Connecting Meaning: The promise and challenges of Knowledge Graphs as providers of large-scale data semantics** (Zaal 2) by Renes Alvesco. Includes a description of knowledge graphs as facilitators of large-scale data semantics and the challenges for architects and decision-makers.



# WORKSHOPS 28 MAART

9:00 – 12:30

## DATA PRODUCTS – FROM DESIGN, TO BUILD, TO PUBLISHING AND CONSUMPTION

Mike Ferguson, Managing Director, Intelligent Business Strategies Ltd

This half-day workshop looks at the development of data products in detail. It also looks at the strengths and weaknesses of data mesh implementation options for data product development. Which architecture is best to implement this? How do you co-ordinate multiple domain-oriented teams and use common data infrastructure software like Data Fabric to create high-quality, compliant, reusable, data products in a Data Mesh. Is there a methodology for creating data products? Also, how can you use a data marketplace to share and govern the sharing of data products?

Most companies today are storing data and running applications in a hybrid multi-cloud environment. Analytical systems tend to be centralised and siloed like data warehouses and data marts for BI, cloud storage data lakes for data science and stand-alone streaming analytical systems for real-time analysis. These centralised systems rely on data engineers and data scientists working within each silo to ingest data from many different sources and engineer it for use in a specific analytical system or machine learning models. There are many issues with this centralised, siloed approach including multiple tools to prepare and integrate data, reinvention of data integration

pipelines in each silo and centralised data engineering with poor understanding of source data unable to keep pace with business demands for new data.

To address these issues, a new approach called Data Mesh emerged in late 2019 attempting to accelerate creation of data for use in multiple analytical workloads. Data Mesh is a decentralised business domain-oriented approach to data ownership and data engineering to create a mesh of reusable data products that can be created once and shared across multiple analytical systems and workloads.

This half-day workshop looks at the development of data products in detail and also, how can you use a data marketplace to share and govern the sharing of data products across the enterprise to shorten time to value.

### Learning Objectives:

- Strengths and weaknesses of centralised data architectures used in analytics
- The problems caused in existing analytical systems by a hybrid, multi-cloud data landscape
- The emergence of data mesh and data products
- What exactly a data product is and the types of data products that you can create
- The benefits that data products offer and what are the implementation options?
- How to organise to create data products in a decentralised environment so you avoid chaos?
- How business glossaries can help ensure data products are formally defined, understood by business users and semantically linked
- The critical importance of a data catalog in understanding what data is available
- What software is required to build, operate and govern a data mesh of data products for use in a data lake, a data lakehouse or data warehouse?
- What is data fabric software, how does it integrate with data catalogs and connect to data in your data estate
- An Implementation methodology to produce ready-made, trusted, reusable data products
- Collaborative domain-oriented development of modular and distributed DataOps pipelines to create data products
- How a data catalog and automation software can be used to generate DataOps pipelines



# WORKSHOPS 28 MAART

- Managing data quality, privacy, access security, versioning, and the lifecycle of data products
- Publishing semantically linked data products in a data marketplace for others to consume and use
- Governing the sharing and use of data products in a data marketplace
- Consuming data products in an MDM system
- Consuming and assembling data products in multiple analytical systems like data warehouses, lakehouses and graph databases to shorten time to value.

## Who is it for?

This seminar is intended for business data analysts, data architects, chief data officers, master data management professionals, data scientists, IT ETL developers, and data governance professionals. It assumes you understand basic data management principles and data architecture plus a reasonable understanding of data cleansing, data integration, data catalogs, data lakes and data governance.

## Detailed course outline

Most companies today are storing data and running applications in a hybrid multi-cloud environment. Analytical systems tend to be centralised and siloed like data warehouses and data marts for BI, cloud storage data lakes or Hadoop for data science and stand-alone streaming analytical systems for real-time analysis. These centralised systems rely on data engineers and data scientists working within each silo to ingest data from many different sources, clean and integrate it for use in a specific analytical system or machine learning models. There are many issues with this centralised, siloed approach including multiple tools to prepare and integrate data, reinvention of data integration pipelines in each silo and centralised data engineering with poor understanding of source data unable to keep pace with business demands for new data. Also, master data is not well managed. To address these issues, a new approach emerged in late 2019 attempting to accelerate creation of data for use in multiple analytical workloads. That approach is Data Mesh. Data Mesh is a decentralised business domain-oriented approach to data ownership and data engineering to create a mesh of reusable data products that can be created once and shared across multiple analytical systems and workloads. A Data Mesh can be implemented in a number of ways. These include using one or more cloud storage accounts on cloud storage, on an organised data lake, on a Lakehouse, on a data cloud, using Kafka or using data virtualisation. Data products can then be consumed in other pipelines for use in streaming analytics, Data Warehouses or

Lakehouse Gold Tables, for use in business intelligence, feature stores for use data science, graph databases for use in graph analysis and other analytical workloads.

This half-day workshop looks at the development of data products in detail. It also looks at the strengths and weaknesses of data mesh implementation options for data product development. Which architecture is best to implement this? How do you coordinate multiple domain-oriented teams and use common data infrastructure software like Data Fabric to create high-quality, compliant, reusable, data products in a Data Mesh. Is there a methodology for creating data products? Also, how can you use a data marketplace to share and govern the sharing of data products? The objective is to shorten time to value while also ensuring that data is correctly governed and engineered in a decentralised environment. It also looks at the organisational implications of Data Mesh and how to create sharable data products for use as master data, in a data warehouse, in data science, in graph analysis and in real-time streaming analytics to drive business value? Technologies discussed includes data catalogs, data fabric for collaborative development of data integration pipelines to create data products, DataOps to speed up the process, data orchestration automation, data observability and data marketplaces.

- What are data products?
- What makes creating data products different from other approaches to creating data for use analytical workloads?
- A best practice methodology for creating data products
- How to design semantically linked data products to enable rapid consumption and use of data to produce new insights
- Quick start mechanisms to speed up data product design
- Defining common business data names for data products in a business glossary
- Data modelling techniques for data products
- Discovering data needed to build data products using a data catalog
- Developing DataOps pipelines to engineer the data needed using data fabric
- Publishing data products – the role of the data marketplace
- Governing access to and use of data products across the enterprise
- Consuming and assembling data products for use in multiple analytical workloads
- Technologies and skills needed.

# WORKSHOPS 28 MAART

9:00 – 12:30

## CONCEPT MODELLING FOR BUSINESS ANALYSTS

Alec Sharp, Founder, Clariteq Systems Consulting

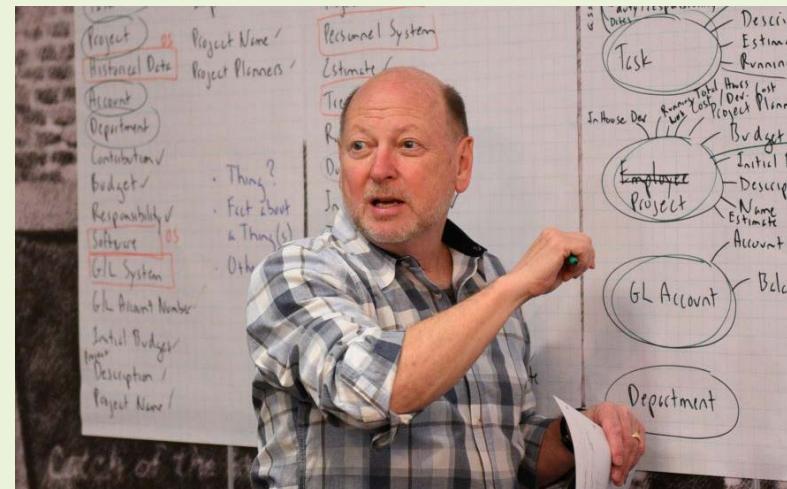
**Concept Modelling (or Conceptual Data Modelling) has seen an amazing resurgence of popularity in recent years, and Alec Sharp illustrates the many reasons for this along with practical techniques and guidelines to ensure useful models and business engagement.**

Whether you call it a conceptual data model, a domain model, a business object model, or even a “thing model,” the concept model is seeing a worldwide resurgence of interest. Why? Because a concept model is a fundamental technique for improving communication among stakeholders in any sort of initiative. Sadly, that communication often gets lost – in the clouds, in the weeds, or in chasing the latest bright and shiny object. Having experienced this, Business Analysts everywhere are realizing Concept Modelling is a powerful addition to their BA toolkit. This session will even show how a concept model can be used to easily identify use cases, user stories, services, and other functional requirements.

Realizing the value of concept modelling is also, surprisingly, taking hold in the data community. “Surprisingly” because many data practitioners had seen concept modelling as an “old school” technique. Not anymore! In the past few years, data professionals who have seen their big data, data science/AI, data lake, data mesh, data fabric, data lakehouse, etc. efforts fail to deliver expected benefits realise it is because they are not based on a shared view of the enterprise and the things it cares about. That’s where concept modelling helps. Data management/governance teams are (or should be!) taking advantage of the current support for Concept Modelling. After all, we can’t manage what hasn’t been modelled!

The Agile community is especially seeing the need for concept modelling. Because Agile is now the default approach, even on enterprise-scale initiatives, Agile teams need more than some user stories on Post-its in their backlog. Concept modelling is being embraced as an essential foundation on which to envision and develop solutions. In all these cases, the key is to see a concept model as a description of a business, not a technical description of a database schema.

This workshop introduces concept modelling from a non-technical perspective, provides tips and guidelines for the analyst, and explores entity-relationship modelling at conceptual and logical levels using techniques that maximise client engagement and understanding. We’ll also look at techniques for facilitating concept modelling sessions (virtually and in-person), applying concept modelling within other disciplines (e.g., process change or business analysis,) and moving into more complex modelling situations.



Drawing on over forty years of successful consulting and modelling, on projects of every size and type, this session provides proven techniques backed up with current, real-life examples.

### Topics include:

- The essence of concept modelling and essential guidelines for avoiding common pitfalls
- Methods for engaging our business clients in conceptual modelling without them realizing it
- Applying an easy, language-oriented approach to initiating development of a concept model
- Why bottom-up techniques often work best
- “Use your words!” – how definitions and assertions improve concept models
- How to quickly develop useful entity definitions while avoiding conflict
- Why a data model needs a sense of direction
- The four most common patterns in data modelling, and the four most common errors in specifying entities
- Making the transition from conceptual to logical using the world’s simplest guide to normalisation
- Understand “the four Ds of data modelling” – definition, dependency, demonstration, and detail
- Tips for conducting a concept model/data model review presentation

# WORKSHOPS 28 MAART

- Critical distinctions among conceptual, logical, and physical models
- Using concept models to discover use cases, business events, and other requirements
- Interesting techniques to discover and meet additional requirements
- How concept models help in package implementations, process change, and Agile development

## Learning Objectives:

- Understand the essential components of a concept model
  - things (entities) facts about things (relationships and attributes) and rules

- Use entity-relationship modelling to depict facts and rules about business entities at different levels of detail and perspectives, specifically conceptual (overview) and logical (detailed) models
- Apply a variety of techniques that support the active participation and engagement of business professionals and subject matter experts
- Develop conceptual and logical models quickly using repeatable and Agile methods
- Draw an Entity-Relationship Diagram (ERD) for maximum readability
- Read a concept model/data model, and communicate with specialists using the appropriate terminology.

KIJK VOOR UITGEBREIDE INFORMATIE OP [WWW.ADEPEVENTS.NL/CMB](http://WWW.ADEPEVENTS.NL/CMB)

13:30 – 17:00

## KNOWLEDGE GRAPHS – PRAGMATICHE AANPAK EN BEST PRACTICES

Panos Alexopoulos, Founder, Panagiotis Alexopoulos

This seminar explores the strategic implementation of Knowledge Graph initiatives within organizations, offering a comprehensive framework that blends cutting-edge techniques with real-world case studies. It equips participants with the crucial understanding needed to make informed decisions, optimize initiatives, and unlock the transformative potential of Knowledge Graphs in today's data-driven landscape.

In today's data-driven landscape, the concept of a knowledge graph has emerged as a pivotal framework for managing and utilizing interconnected data and information. Stemming from Google's proclamation that shifted the focus from searching for strings to understanding entities and relationships, the term encapsulates a network of interconnected entities and concepts,

facilitating data integration, sharing, and utilization within organizations.

Amid the widespread adoption of knowledge graphs across diverse domains, ensuring the accuracy, reliability, and consensus of semantic information becomes an imperative. The construction and utilization of these graphs present multifaceted challenges, ranging from ensuring data quality to scaling and adapting to evolving contexts.

Implementing a successful Knowledge Graph initiative within an organization demands strategic decisions before and during its execution. Often overlooked are critical considerations such as managing trade-offs between knowledge quality and other factors, prioritizing knowledge evolution, and allocating resources effectively. Neglecting these facets can lead to friction and suboptimal outcomes.

This half-day seminar delves into the technical, business, and organizational dimensions essential for data practitioners and executives embarking on a Knowledge Graph initiative. Offering insights gleaned from real-world case studies, the seminar provides a comprehensive framework that combines cutting-edge techniques with pragmatic advice. It equips participants to navigate the complexities of executing a knowledge graph project successfully.

Moreover, the session addresses pivotal strategic dilemmas encountered during the design and execution phases of knowledge graph projects, and outlines potential approaches to tackle these challenges, empowering attendees with actionable strategies to optimize their initiatives.



# WORKSHOPS 28 MAART

## Learning Objectives

- Understand the key factors determining the feasibility and viability of implementing a knowledge graph in an organization.
- Identify and articulate the fundamental questions crucial for preparing and launching a successful knowledge graph initiative.
- Learn techniques to determine and prioritize the content requirements of a knowledge graph.
- Grasp best practices in schema design for knowledge graphs, addressing real-world challenges of uncertainty and vagueness.
- Explore strategies and guidelines for populating a knowledge graph, evaluating available knowledge extraction systems.
- Gain insights into assessing and prioritizing quality dimensions within a knowledge graph.
- Explore practical applications of knowledge graphs, such as entity disambiguation and semantic search, optimizing performance through design principles.
- Gain insights into methodologies for ongoing maintenance and evolution of knowledge graphs, ensuring their sustained relevance and adaptability across time.

## Who is it for?

- Data practitioners: Data scientists, data engineers, data analysts, and database administrators seeking to deepen their understanding of knowledge graphs, their implementation, and the technical intricacies involved.
- Technology Leaders: Architects, CTOs, and IT professionals exploring or leading initiatives involving data integration, semantic technologies, and knowledge management systems.
- Business Executives and Managers: Leaders and decision-makers responsible for overseeing data strategies, innovation, and organizational transformation, aiming to comprehend the strategic implications and business value derived from knowledge graph initiatives.

## Course Outline

The seminar will walk participants through 8 key stages of introducing, developing, delivering and evolving Knowledge Graphs in an organization. These are:

### Stage 1 – “Knowing where you are getting into”

- Clarification of the knowledge graph concept
- Key factors influencing the ease or difficulty of building a knowledge graph
- Evaluating feasibility and viability of implementing a knowledge graph in a specific organization and for a particular business problem

### Stage 2 – “Setting up the stage”

- Exploring 5 key questions essential before initiating knowledge graph development
- Defining what, why, how, who, and the stakeholders involved in the project
- Outlining actions required to seek and discover answers to these questions

### Stage 3 – “Deciding what to build”:

- Delving into knowledge graph specification
- Use of competency questions for gap analysis between organizational knowledge capabilities and needs
- Scoping and prioritizing knowledge graph content

### Stage 4 – “Giving it a shape”

- Schema design using Ontology Representation and Engineering
- Identification of conceptual modeling best practices, dilemmas, and pitfalls
- Addressing uncertainty and vagueness

### Stage 5 – “Giving it substance”

- Exploring the challenging task of knowledge graph population
- Description of population tasks and associated difficulties
- Designing optimal population pipelines

### Stage 6 – “Ensuring it’s good”:

- Assessing knowledge graph quality, defining dimensions, and metrics
- Insights into quality trade-offs and prioritization of dimensions
- Measuring quality and effective prioritization of focus areas

### Stage 7 – “Making it useful”:

- Typical knowledge graph applications
- Guidelines and best practices for optimizing knowledge graph usefulness and value

### Stage 8 – “Making it last”:

- Addressing the challenge of knowledge graph maintenance and evolution
- Detecting, measuring, and monitoring concept drift
- Best practices for enabling continuous improvement and preventing knowledge graph obsolescence over time.

# INFORMATIE DATA WAREHOUSING & BUSINESS INTELLIGENCE SUMMIT 2024



## DATUM EN TIJD

Het congres DW&BI SUMMIT vindt plaats op 27 en 28 maart. Op 27 maart begint het programma om 09.00 uur en duurt tot 16.45 uur. Registratie is mogelijk vanaf 08.00 uur. Op 28 maart beginnen de workshops op verschillende tijden, zie daarvoor de website.

## PLAATS

Van der Valk Hotel Utrecht  
Winhontlaan 4-6  
3526 KV Utrecht  
Telefoon 030 8000 800  
E-mail: [utrecht@valk.nl](mailto:utrecht@valk.nl)  
Website hotel: [www.vandervalkhotelutrecht.nl](http://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](http://www.dwbisummit.com). Geeft u de voorkeur aan schriftelijk aanmelden? Stuur de PDF van uw aanmelding of inkoopopdracht naar [seminars@adeptevents.nl](mailto: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 conferentie loont.

Deelnameprijs	On-premise	Online
Best rate (geldig tot 24 januari 2024)	€ 586,50	€ 416,50
Vroegboekkorting (25 januari t/m 21 februari)	€ 621	€ 441
Regulier tarief (vanaf 22 februari)	€ 690	€ 490

De kosten voor een workshop van een halve dag bedragen slechts € 370,- in combinatie met congresdeelname. Ook voor deze workshops geldt een vroegboekkorting (geen best rate) van 10% tot 22 februari 2024. Aanmelding daarvoor kan ook afzonderlijk, via de website van **Adept Events**.

Deelnemers aan het congres hebben bovendien nog enkele maanden toegang tot de video opnames, dus als u een sessie moet missen is er geen man overboord. Alle prijzen exclusief BTW. Documentatie, lunch, en koffie zijn inbegrepen. Leden van KNVI afdeling BI&A 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 voor de tweede en derde deelnemer een korting van 10%. 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
-  [www.dwbisummit.com](http://www.dwbisummit.com)
-  [seminars@adeptevents.nl](mailto:seminars@adeptevents.nl)
-  [@AdeptEventsNL | #dwbisummit](https://twitter.com/AdeptEventsNL)
-  <https://www.linkedin.com/company/adept-events>
-  Bezoek ook onze Business Intelligence en Datawarehousing website [www.biplatform.nl](http://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 door de onderstaande mediapartners.

