

# DATAMODELLERING ER DIAGRAM

## Inleiding

In dit whitepaper wordt de datamodelleervorm ER diagram beschreven. Deze modelleervorm staat in verhouding tot een aantal andere modelleervormen. Wil je een beeld krijgen van welke modelleervormen er zijn bekijk dan het whitepaper wat een introductie geeft tot datamodelleervormen en deze serie van whitepapers zie Inleiding datamodellering.

Dit whitepaper is een onderdeel van meerdere whitepapers over modellering in de fysieke laag. ER modellering is specifiek voor relationele databases. Daarnaast zijn modelleerwijzen rond XSD en JSON relevant. Informatie over deze modelleervormen is te vinden in deze serie van van whitepapers.

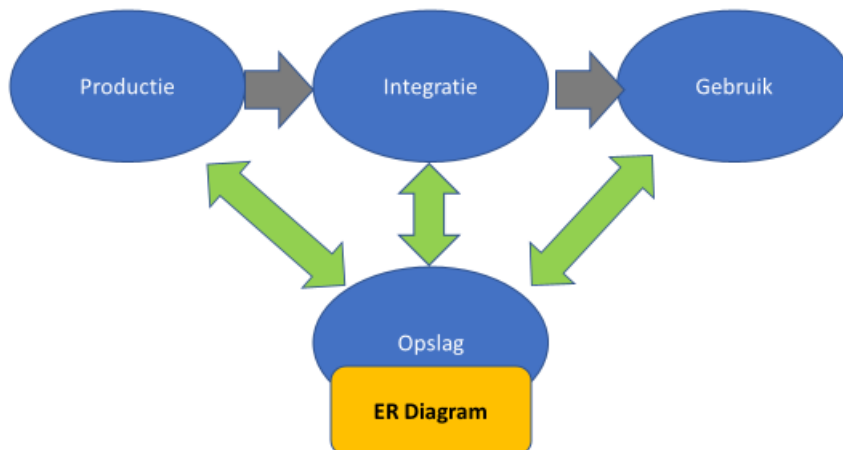
## Doel

ER staat voor Entity-Relationship en ER diagrammen geven een gedetailleerde notatiewijze voor het modelleren van deze twee concepten. Daarnaast kun je met ER diagrammen de kolommen in een tabel modelleren en eventueel een aantal belangrijke constraints in relationele databases zoals primary- en foreign key constraints

ER diagrammen hebben tot doel om tot in detail de fysieke structuur van een relationele database te modelleren. Veelal wordt daarbij al rekening gehouden met de eisen die vanuit een specifiek databaseplatform gesteld worden aan de inrichting van een database. Hierdoor is het mogelijk om deze ER modellen rechtstreeks om te zetten naar een fysieke inrichting in een database. Dit staat bekend als de Data Definition Language (DDL).

# Raamwerken

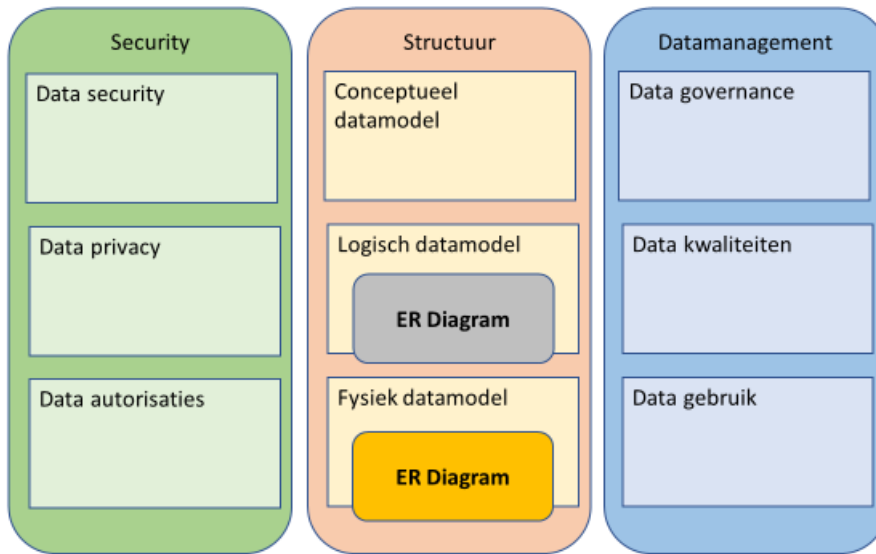
## DATALEVENSLOOP



Vanwege de focus op relationele databases van de ER diagrammen wordt deze notatiewijze alleen op de data opslag afgebeeld in de data levensloop. ER diagrammen zijn heel specifiek gericht op het beschrijven van de opslagstructuren en -beperkingen binnen een relationele database.

Wordt een ER diagram gebruik voor logische modellering dan zou deze modellering ook gebruikt kunnen worden voor modellering van de andere onderdelen van de data levensloop. Echter hiervoor zijn naar mijn mening andere, meer geschikte, notatiewijzen beschikbaar.

## DATARAAMWERK



In het data raamwerk zien we het ER diagram alleen terug in de datastructuur kolom. In het beschrijven van fysieke data structuren is het ER diagram een zeer goed toepasbare notatiewijze die goed aansluit bij de behoefte van de stakeholders.

Het ER diagram is met name geschikt voor de fysieke data modellering en wordt daar ook veelvuldig toegepast. Dit vanwege de mogelijkheden om de fysieke structuur van een relationele database tot in detail te beschrijven in een ER diagram. Inclusief de specifieke aspecten van een bepaald database platform, zoals specifieke datatypes voor dat platform of specifieke constraints zoals autonummering implementaties.

ER diagrammen kunnen gebruikt worden voor het opstellen van logische modellen. Dit zijn feitelijk vereenvoudigde modellen die niet databaseplatform specifiek zijn. Een aantal tools ondersteund deze opzet. Echter voor logische datamodellering is het UML klassediagram meer geschikt.

## Stakeholders

Fysieke modellering met ER richt zich op stakeholders die zich voornamelijk in het ICT werkveld bevinden. Vanwege de technische details is deze modelleerwijze minder geschikt voor interactie met stakeholders binnen de business.

Fysieke datamodellering dient dan ook zorg te dragen dat de implementatie in een relationele database tot in detail beschreven is en desgewenst een geautomatiseerde transformatie van model naar implementatie mogelijk is.

Onderstaande opsomming geeft de belangrijkste stakeholders:

- **Technisch (Applicatie) Beheerders en DBA**, een groep stakeholders die diepgaande kennis heeft van (standaard) informatiesystemen en het onderliggende fysieke datamodel inclusief de door de (software en databaseplatform) leverancier gebruikte datastructuren en opslagmechanismen
- **Software- of Informatiearchitect**, zal veelal kennis hebben van het fysieke datamodel in de huidige- en de gewenste situatie van een geïmplementeerde applicatieverandering. Ze hebben daarnaast veelal diepgaande kennis van data patronen, data structuren en data modellen.
- **Ontwikkelaars**, software ontwikkelaars met name van back end systemen dienen kennis te hebben van de opslag van hun structuren in relationele databases als dit het opslagmedium is.

## Concepten

Binnen ER diagrammen worden voornamelijk fysieke datamodellen weergegeven in een verrijkte graaf. ER diagrammen zijn opgebouwd uit de volgende concepten:

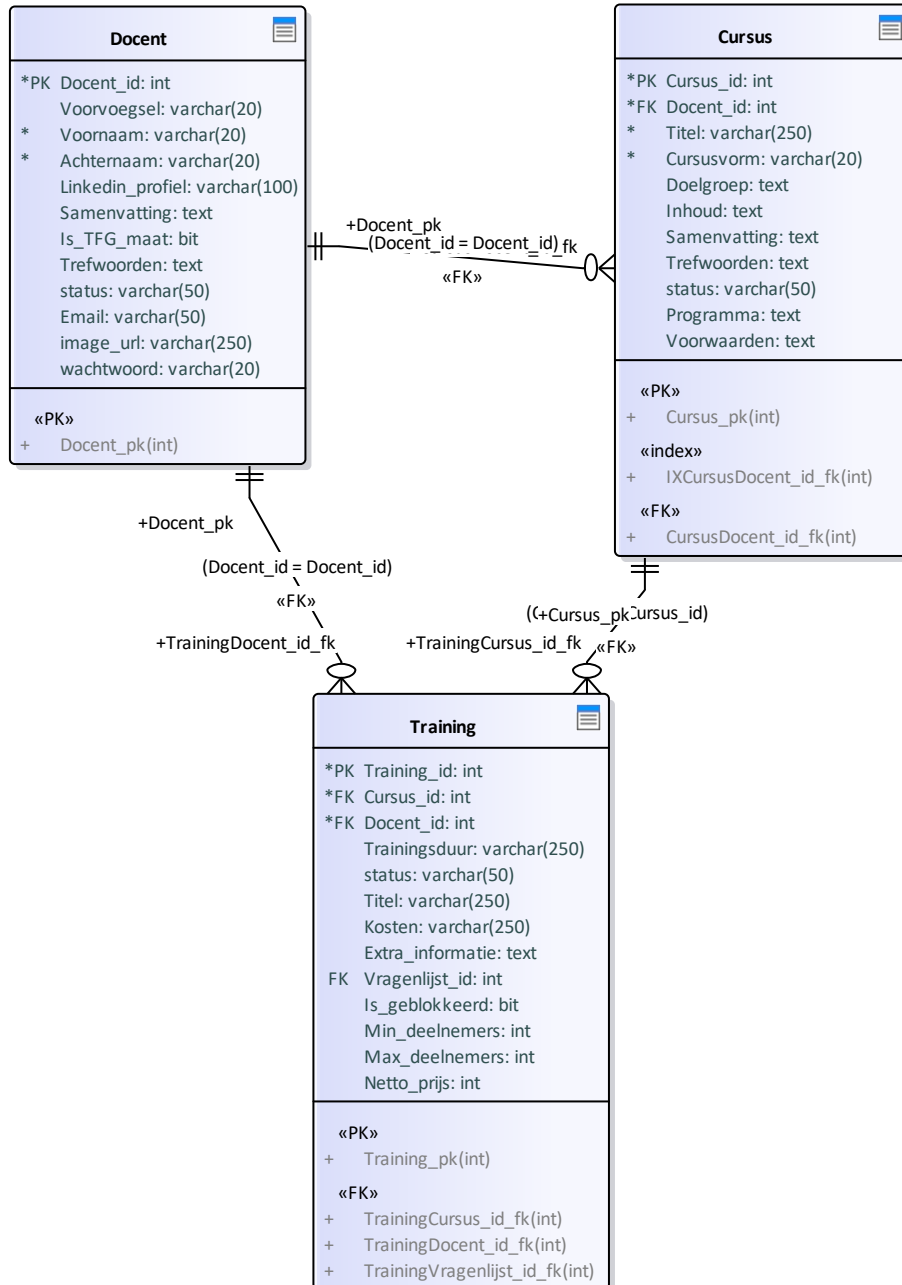
- **Tabel**: is het hoofdconcept. Een tabel kenmerkt zich in dat het is opgebouwd uit kolommen en dat het een relatie kan hebben met andere tabellen. Tabellen zijn het hoofdconcept voor het opslaan bij elkaar behorende gegevens in een relationele database. Het is de fysieke implementatie van het logische concept klasse.
- **Kolom**, zijn de onderdelen van een tabel die ervoor zorgdragen dat de eigenschappen uit een klasse (in het logische model) binnen een tabel van elkaar onderscheiden kunnen worden. Kolommen worden bij de implementatie gebruikt bij het ophalen en wegschrijven van gegevens uit een database. Bijvoorbeeld binnen Select en Insert/Update/Delete commando's. Voor een kolom kunnen in een ER diagram extra gegevens worden bijgehouden. Bijvoorbeeld het datatype van de inhoud maar ook of de kolom null (leeg) waarden mag bevatten.
- **Relatie**, is een verbinding tussen twee tabellen en geeft verbanden tussen fysieke entiteiten weer. Het draagt er zorg voor dat bij de implementatie in een relationele database deze afdwingt dat de relaties aan de eisen voldoen van deze database. Hiermee wordt ervoor zorggedragen dat verwijzingen van de ene tabel naar de andere altijd correct zijn.
- **Primary keys of primaire sleutel**, dragen zorg voor het uniek identificeren van rijen in een tabel. Op basis hiervan zijn tabellen op correcte wijze aan elkaar te verbinden. Is een belangrijk aspect van tabellen in een relationele database.
- **Foreign keys of verwijzende sleutels**, bij bepaalde relaties in een relationele database wordt gebruik gemaakt van een ouder – kind relatie. Verwijzende sleutels dragen zorg dat in de kind tabel op correcte wijze een verwijzing is opgenomen naar de sleutel in de ouder-tabel. De database bewaakt te allen tijde de correctheid van deze verwijzingen.

Op basis van deze vijf basis concepten zijn zeer krachtige modellen op te stellen die zorgdragen voor dataopslag van data die aan bepaalde kwaliteitseisen op het vlak van correctheid en referentiele integriteit voldoen.

## Notatie

Het ER diagram is een verrijkte graaf waarbij met name de relaties een aantal verrijkingen kent die de taal de zeggingskracht geeft. In onderstaande afbeelding een voorbeeld van een ER diagram

dm Fysiek voorbeeld project



In de afbeelding is te zien dat een tabel bestaat uit een naam en dat een tabel kolommen bevat. Voor de kolommen geldt dat het data type getoond worden en de cardinaliteit [0..1] als het optioneel is of verplicht [1..1].

Voor de relaties zijn er meerdere extra kenmerken te zien. Bijvoorbeeld welke rol de tabel vervuld in de relatie (veelal op basis van primary keys en foreign keys). Daarnaast wordt voor de relatie aangegeven

wat de cardinaliteit is. Hiermee kan een database zorgdragen voor het correct houden van de relaties tussen tabellen binnen de database.

Als laatste zijn de indexen, primary- en foreign keys te noemen. Dit zijn beperkingen in tabellen die gemodelleerd kunnen worden in een ER diagram. In bovenstaande afbeelding zijn deze constraints opgenomen onder de kolommen.

## Kenmerken

ER diagram heeft de volgende kenmerken:

- Een internationale en breed geaccepteerde standaard
- Rijke notatie met name voor fysieke datamodellering
- Interactie met leveranciers wordt eenvoudiger vanwege standaardisatie
- Het genereren van diverse andere output zoals DDL scripts voor databases
- Notatie is rijk en modellen kunnen complex worden
- Grote mate van detaillering
- Mogelijkheid om bestaande relationele databases geautomatiseerd om te zetten naar een ER model
- Het modelleren van relaties met niet relationele entiteiten is niet mogelijk in modellering
- Moeilijk te begrijpen voor Niet ICTers

## Gebruikstoepassingen

Het ER diagram wordt vooral toegepast binnen de fysieke data modellering en in een aantal gevallen binnen de logische modellering. Binnen dit werkveld is het een zeer krachtige en veel gebruikte modelleerwijze. Het heeft de mogelijkheid om voldoende database specifieke details toe te voegen (met name bij de relaties, de kolommen en de keys).

Een extra gebruikstoepassing van het ER diagram is de inzet van deze notatie binnen database modellering en -ontwikkeling. Hierbij wordt het ER diagram omgezet tot een daadwerkelijke database implementatie of omgekeerd een database implementatie wordt omgezet tot een ER diagram.

Houdt er rekening mee dat ondanks de eenvoud van de notatiewijze het opstellen van een ER diagram een complex traject kan zijn, zeker bij een complex domein of binnen een organisatiecontext waar rond de informatievoorziening en logische modellering weinig volwassenheid is.

## Gerelateerde notatiewijzen

Deze ER notatie heeft met veel notatiewijzen, met name in de fysieke en logische modellering een relatie. In onderstaande opsomming een overzicht:

- **ArchiMate Datamodellering** of begrippenlijst en -boom, hiermee worden de verbanden gelegd vanuit het conceptuele datamodel naar het logische datamodel
- **UML klassediagram** waarmee een koppeling gelegd wordt naar het logische model in UML klassediagram vanuit een fysieke opslagstructuur in een relationele database
- **XSD schema** waarmee een verband wordt gelegd van het fysieke datamodel (veelal via een logisch model in UML) naar een fysieke inrichting voor gestandaardiseerde berichtuitwisseling.

- **CRUD matrix**, voor het modelleren van welke persoon, rol of database user of -group toegang krijgt tot de gegevens in een tabel.

## Tooling

Voor de ER diagrammen zijn een aantal algemene en specifieke tools aanwezig:

- Visio met ERD extensie
- Sparx Enterprise Architect
- Powerdesigner
- ERWin
- BizDesign
- ADO-IT
- ARIS

## Evaluatie

ER diagram is een veelgebruikte notatiewijze met name voor het opstellen van fysieke datamodellen voor implementatie in relationele databases. Het legt daarmee een verbinding tussen de logische modellen en de fysieke implementatie in een relationeel database platform. Het is daarmee een onmisbare schakel in de data modelleerketen.

Het ER diagram wordt met name toegepast binnen software- en databaseontwikkeling. De basisnotatie biedt een ruime hoeveelheid mogelijkheden om complexe modellen op te stellen. Dit is enerzijds de kracht van het ER diagram en anderzijds een zwakte omdat de modellen veelal te complex zijn voor stakeholders met minder modelleerervaring.

Voor ER diagrammen is een veelheid aan tooling aanwezig, in dit artikel slechts een beperkte opsomming.

## Over de auteur



Bert Dingemans is trainer op het vlak van data architectuur, data management en Big Data. Hij heeft een passie voor modelleren, modelleertools en het effectief inzetten van geautomatiseerde hulpmiddelen om modellen effectief in te zetten in de praktijk. Bert is te bereiken via [bert@interactory.nl](mailto:bert@interactory.nl)