

EEN DATAMODEL VOOR GRAFZERKEN NEDERLAND

Wouter Bonhof



AUTEUR: Wouter Bonhof
DATUM: 15 maart 2025
CURSUS: Datamodelleren 2025
BEGELEIDER: Jan Meeuwsen

OMSLAG: Hans Verburg
BRON: www.istockphoto.com/nl/foto/grafsteen-vloer-gm847322684-142115553

Voorwoord

De cursus datamodelleren was anders dan ik van tevoren had verwacht. Ik had het idee dat we met talen als SQL en Python aan de slag zouden gaan om data te analyseren, grafieken te maken en rapporten op te stellen. Ik begreep dus niet goed wat de term ‘datamodelleren’ inhoudt. Door de cursus weet ik nu echter wel wat datamodelleren inhoudt en ook het enorme belang ervan. Niet enkel voor latere data-analyses is een goed ontworpen datamodel van belang, maar ook voor de klant/ opdrachtgever is een goed ontworpen datamodel cruciaal. Het documenteren van de beslissingen die zijn gemaakt tijdens het ontwerpen van het datamodel heeft ook enorm veel waarde. Zonder de documentatie kunnen toekomstige aanpassingen aan het model mogelijk niet goed worden geïmplementeerd en bestaat de kans dat een volledig nieuw model moet worden ontworpen.

Voor de aanvang van de cursus had ik al enkele keren gehoord en gelezen over het normaliseren van data en wist ik dat er verschillende normaalvormen zijn. Ik heb de verschillende normaalvormen echter nooit eigen kunnen maken en kon de eigenschappen van de normaalvormen niet uitleggen. Omdat we tijdens de cursus vele malen met de normaalvormen aan de slag zijn gegaan en hier vaak zelf hebben geknutseld zijn de normaalvormen mij nu hartstikke duidelijk. De 1e en 2e Meeuwsen normaalvormen waren mij nog onbekend, maar de meerwaarde van deze normaalvormen is onmiskenbaar. Zonder generalisatie zou er duplicatie kunnen ontstaan en het bijhouden van de mutatiegeschiedenis maakt herstelwerkzaamheden een stuk makkelijker.

De belangrijkste les die ik mee zal nemen van de cursus is het contact met de opdrachtgever. Als iets onbekend is kan je dit proberen af te leiden van de beschikbare informatie of gaan gokken, maar je kunt het beste echter navraag doen bij de opdrachtgever. Deze weet het beste wat er verwacht wordt en zal de meeste domeinkennis hebben. Mocht de opdrachtgever niet beschikbaar zijn, staat het je vrij om zelf iets te ontwerpen, maar noteer dan wel je ontwerpbeslissingen zodat je het model later goed kan uitleggen en bespreken.

Ik wil Jan enorm bedanken voor de vermakelijke en leerzame cursus. Het was een genot om wekelijks met de medecursisten in wisselende samenstellingen te overleggen over de opdrachten en deze vervolgens klassikaal te bespreken. Jan was altijd geduldig, geëngageerd en nieuwsgierig. Als we met een creatieve oplossing kwamen probeerde hij de gedachtegangen erachter volledig te begrijpen, om ons vervolgens in begrijpelijke woorden uit te leggen waarom zijn oplossing anders (en beter) was dan de onze. De informatie en manier van denken die ik tijdens deze cursus heb opgedaan zal ik nooit vergeten.

Inhoudsopgave

Voorwoord	3
Inleiding	5
Grafzerken Nederland	6
Probleemstelling	6
Use-cases	6
Datamodellen	8
Entiteiten	8
CDM	8
LDM	9
Datamodel in de 0 ^e normaalvorm:	9
Datamodel in de 1 ^e normaalvorm:	9
Datamodel in de 2 ^e normaalvorm:	10
Datamodel in de 3 ^e normaalvorm:	10
Datamodel in de 1 ^e Meeuwsen normaalvorm:	11
Datamodel in de 2 ^e Meeuwsen normaalvorm:	12
Metadata	14
Data principes	20
Conclusie	22

Inleiding

Bij de opzet van een nieuwe organisatie Grafzerken Nederland (GN) is het nodig dat er een volledig nieuwe database wordt opgesteld. Voordat deze database in elkaar wordt gezet wil de organisatie graag een rapport ontvangen, waarin een volledig uitgewerkt datamodel zien. In het datamodel moeten alle gegevens die GN in de database wil opnemen uitgewerkt zijn en moet duidelijk zijn dat de data zo efficiënt en dataveilig mogelijk worden opgeslagen. Het rapport moet ook de ontwerpbeslissingen van het datamodel bevatten, zodat latere aanpassingen in het datamodel hieraan getoetst kunnen worden.

Het onderstaande rapport is opgesteld door Wouter Bonhof, junior consultant bij House of Beta. Eerdere ervaring met het ontwerpen van databases heeft Wouter opgedaan tijdens zijn doctorale studie aan de Univeristy of Exeter (Verenigd Koninkrijk) en tijdens een eenjarige opdracht bij IVO Rechtspraak. Daarnaast heeft hij bij House of Beta meerdere cursussen over data voltooid, zoals Certified Data Practitioner en Datamodelleren.

Onderstaand rapport is opgedeeld in drie delen: Grafzerken Nederland, de datamodellen, en de data principes. De doelstellingen van organisatie Grafzerken Nederland worden in het gelijknamige gedeelte beschreven. Daarnaast wordt de context voor het huidige rapport geschetst en worden enkele voorbeelden gegeven van situaties waar GN het datamodel voor wil gebruiken. In de datamodellen worden de verscheidene entiteiten erkend en worden de relaties tussen deze entiteiten omschreven. Door middel van een volledig uitgeschreven Logisch Datamodel en enkel visualisaties wordt kenbaar gemaakt hoe de data worden opgeslagen en hoe deze data aan elkaar gekoppeld kan worden voor latere analyses. In de data principes wordt besproken op welke manieren het voorgestelde datamodel voldoet aan de algemeen erkende principes Governance, Kwaliteit en Beschikbaarheid.

Grafzerken Nederland

Grafzerken Nederland is een organisatie opgezet door de Nederlandse overheid in samenwerking met commerciële archeologische bedrijven en Nederlandse universiteiten met een archeologische afdeling. Het doel van GN is het documenteren van alle grafzerken die in Nederlandse kerken te vinden zijn. De organisatie wil deze informatie kunnen analyseren voor onderzoeken, maar het moet ook toegankelijk zijn voor geïnteresseerde partijen uit de publieke sector (bijvoorbeeld burgers, onderzoekers en reisgidsen). De initiële focus van de organisatie is op de grafzerken die zich binnen kerken bevinden, maar op termijn moet het ook mogelijk worden om grafzerken die zich buiten de kerk bevinden toe te voegen aan de database.

Probleemstelling

Historische **kerken** in Nederland zijn gevuld met grafstenen (**grafzerken**). Deze begravingen bevatten voornamelijk de restanten van fortuinlijke en invloedrijke **personen**. Doordat de grafzerken soms de vloer van een kerk volledig bedekken, zijn de meeste grafzerken niet beschermd en wordt er dagelijks over gelopen. Hierdoor gaat de informatie die op de grafzerken te vinden is, zoals de namen van begravenen en hun geboorte- en sterfdata, verloren. Er is daarom besloten een database op te zetten waarin de grafzerken en de daarop gegraveerde informatie gedocumenteerd worden. Deze informatie gaat over de begraven personen en **relaties** tussen personen, informatie over het fysieke aspect van de grafzerk, zoals de steensoort en positie (liggend of staand).

Als onderdeel van verscheiden **projecten** is voor een groot deel van de Nederlandse grafzerken al relevante informatie verzameld door **onderzoekers**. Deze projecten zijn uitgevoerd voor **opdrachtgevers**, zoals de **beheerders** van de kerken (kerkgenootschappen en de Rijksdienst voor Cultureel Erfgoed (RCE)) maar soms ook de bedrijven die door de beheerders zijn ingehuurd voor bijvoorbeeld renovatiewerkzaamheden. Naar aanleiding van deze projecten, maar ook als onderdeel van onderzoeksprojecten door studenten en onderzoekers zijn vele wetenschappelijke **artikelen** gepubliceerd.

De database met alle informatie moet beschikbaar worden voor het **publiek**. Gebruikers uit het publiek kunnen informatie over de grafzerken inzien via de nog te ontwikkelen website en app. Het publiek heeft ook de mogelijkheid om grafzerken te documenteren als deze nog niet in de database zijn opgenomen. Door middel van deze burgerwetenschap hoopt GN op een hoger tempo alle grafzerken gedocumenteerd te hebben. De grafzerken informatie die door burgers wordt ingevoerd moet vervolgens nog wel door een interne onderzoeker van Grafzerken Nederland of een externe geaffilieerde onderzoeker gevalideerd worden.

Use-cases

- Bij het aanleggen van een kelder wil de kerk de grafzerken die in de bovenliggende vloer aanwezig zijn laten documenteren als onderdeel van een project.
- Een onderzoeker registreert de gegevens van een grafzerk en onderscheidt grafzerken door de coördinaten (noorderbreedte en oosterlengte in decimale graden) van de linkerbovenhoek en de rechteronderhoek van de grafzerk te bepalen m.b.v. GPS-apparatuur (zoals een Total Station of een telefoon).
- Een tourgids wil laten alle grafzerken van de Grote Kerk te Breda kunnen inzien om te zien of hier ook leden van de Van Gogh familie begraven liggen.

- Onderzoekers willen de gemiddelde leeftijd van de begravenen van alle kerken van Amsterdam berekenen.
- Een onderzoeker wil graag inzien welke grafzerken zijn opgenomen in de analyse van een recente publicatie.
- In de app kan een kerkbezoeker informatie zien over de grafzerken waar de kerkbezoeker overheen loopt.
- Via de app wil een kerkbezoeker een nieuw grafzerk invoeren, omdat de kerkbezoeker op de plattegrond ziet dat deze nog niet in de database is opgenomen.
- Na invoer van een nieuwe grafzerk door een standaard gebruiker moet een onderzoeker de ingevoerde gegevens valideren.
- Nadat een project is afgerond kunnen er publicaties komen, die gekoppeld moeten worden aan de relevante projecten, grafzerken en onderzoekers (tevens auteurs van de publicatie).

Datamodellen

Entiteiten

In de Probleemstelling en de Use-cases worden verscheidene entiteiten genoemd. Deze zijn als volgt gedefinieerd:

Tabel 1 De herkende entiteiten en de bijbehoren definities.

Entiteit	Definitie
Bedrijf	Beheerders van kerken en/ of opdrachtgevers voor projecten.
Gebruiker	Leden van het algemene publiek die interesse hebben in grafzerken.
Onderzoeker	Personen die werkzaam zijn bij Grafzerken Nederland of externe geaffilieerde onderzoekers.
Kerk	Religieuze gebouwen waarin/ waarbuiten de grafzerken zich bevinden.
Grafzerk	Steen die gebruikt wordt om de begraafplaats van één of meerdere overledenen te markeren. Bevatten veelal inscripties met informatie over de overledene.
Begravene	Personen die begraven zijn en een grafzerk hebben.
Relatie	De verhouding tussen twee begravene uit één of meerdere begravingen.
Project	Tijdgebonden activiteit tijdens welke nieuwe informatie over grafzerken verzameld is.
Publicatie	Wetenschappelijk en openbaar beschikbaar artikel waarin grafzerken onderdeel zijn van de data.

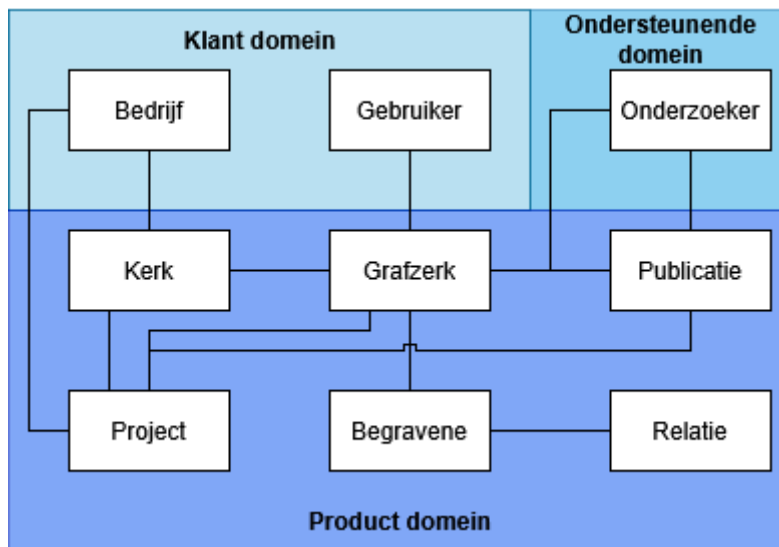
CDM

De hierboven genoemde entiteiten kunnen verdeeld tussen het Klant domein, het Product domein en het Ondersteunende domein. Het Klant domein omvat de entiteiten die van buiten de organisatie Grafzerken Nederland afkomstig zijn. In het Product domein zijn de entiteiten opgenomen waarvan de gegevens door GN verzameld worden en vervolgens ook publiekelijk beschikbaar worden gesteld. In het Ondersteunende domein zijn de interne onderzoekers van GN en de externe geaffilieerde onderzoekers opgenomen (onder de gedeelde naam 'Onderzoeker'), omdat deze intern bij GN betrokken zijn, maar hun informatie niet publiekelijk beschikbaar wordt gesteld.

Tabel 2 De verdeling van de entiteiten tussen de Klant, Product en Ondersteunende domeinen.

Klant domein	Product domein	Ondersteunende domein
Gebruiker	Grafzerk	Onderzoeker
Bedrijf	Kerk	
	Begravene	
	Project	
	Publicatie	
	Relatie	

Op basis van de entiteiten is onderstaand conceptueel datamodel (CDM) gemaakt. Hierin worden de relaties tussen de verscheidene entiteiten getoond. Wederom worden de domeinen waartoe de entiteiten behoren getoond.



Figuur 1 Het conceptuele datamodel.

LDM

In het logisch datamodel (LDM) zijn de entiteiten van het CDM verder uitgewerkt en worden alle bijbehorende attributen genoemd. Het LDM maakt daarbij ook duidelijk op welke manier de verschillende entiteiten aan elkaar gerelateerd zijn. Het ontwerpen van het LDM is hieronder stapsgewijs beschreven en de ontwerpbeslissingen worden per stap toegelicht.

Datamodel in de 0^e normaalvorm:

Grafzerk, Steensoort, Slijtage, Positie, Coördinaat linkerbovenhoek, Coördinaat rechteronderhoek, **Begravene naw**, **Relatie met andere begravene**, **Kerknaam**, **Kerkadres**, **Periode in gebruik**, **Denominatie**, **Beheerder naw**, **Projectnaam**, **Startdatum project**, **Einddatum project**, **Opdrachtgever naw**, **Validatie**, **Datum**, **Onderzoeker naw**, **Affiliatie**, **Gebruiker naw**, **Publicatie**, **Titel**, **Tijdschrift**, **Datum publicatie**, **DOI**, **Auteur naw**

Datamodel in de 1^e normaalvorm:

GRAFZERK: Grafzerk_ID, Steensoort, Slijtage, Positie, Coördinaat_lb_N, Coördinaat_lb_O, Coördinaat_ro_N, Coördinaat_ro_O, **Kerknaam**, **Kerkadres**, **Periode in gebruik**, **Denominatie**, **Beheerdersnaam**, **Straat**, **Huisnummer**, **Toevoeging**, **Postcode**, **Stad**, **Telefoonnummer**, **Emailadres**, **URL**, **Validatie**, **Datum**, **Onderzoeker**, **Voornaam**, **Tussenvoegsel**, **Achternaam**, **Telefoonnummer**, **Emailadres**, **Affiliatie**, **Gebruiker**, **Voornaam**, **Tussenvoegsel**, **Achternaam**, **Telefoonnummer**, **Emailadres**

BEGRAVENE: Begravene_ID, **Voornaam**, **Doopnaam**, **Tussenvoegsel**, **Achternaam**, **Geboortedag**, **Geboortemaand**, **Geboortjaar**, **Sterfdag**, **Sterfmaand**, **Sterfjaar**, **Validatie status**, **Grafzerk_ID** (FK van GRAFZERK)

RELATIE: Begravene_ID (FK van BEGRAVENE), Begravene_ID (FK van BEGRAVENE), **Relatiesoort**

PROJECT: Project_ID, Projectnaam, Startdatum project, Einddatum project, Opdrachtgevernaam, Straat, Huisnummer, Postcode, Stad, Telefoonnummer, Emailadres, URL

GRAFZERK-PROJECT: Grafzerk_ID (FK van GRAFZERK), Publicatie_ID (FK van PUBLICATIE)

PUBLICATIE: Publicatie_ID, Titel, Tijdschrift, Publicatiedatum, DOI, Grafzerk_ID (FK van GRAFZERK)

AUTEUR: Auteur_ID, Voornaam, Tussenvoegsel, Achternaam, Telefoonnummer, Emailadres, Publicatie_ID (FK van PUBLICATIE)

AUTEUR-PUBLICATIE: Auteur_ID (FK van AUTEUR), Publicatie_ID (FK van PUBLICATIE)

Ontwerpbeslissingen:

- Entiteiten en attributen staan altijd in enkelvoud genoteerd.
- Een grafzerk kan voor meerdere personen gemaakt zijn en er is daarom een aparte entiteit gemaakt voor begravenen.
- De gegevens van een grafzerk kunnen door slechts één gebruiker of één onderzoeker worden ingevoerd.
- Voor begravenen zijn de geboortedatum en sterfdatum opgesplitst naar dag, maand en jaar, omdat het mogelijk is dat één of meerdere van deze attributen op de grafzerk ontbreken. Als de geboorte/sterfdatum als één attribuut wordt ingevoerd zal er door het systeem automatisch toch een getal worden ingevoerd voor het ontbrekende onderdeel, wat niet wenselijk is.
- Een begravenen kan met meerdere personen een relatie hebben en er is daarom een aparte entiteit gemaakt voor de relaties tussen begravenen.
- Voor de onderzoekers en gebruikers zijn de adresgegevens niet relevant en daarom niet in het datamodel opgenomen.
- Een publicatie kan meerdere auteurs hebben en dit zou daarom een repeterende groep worden. Vandaar dat er een aparte entiteit voor auteur is gemaakt.
- Omdat een auteur bij meerdere publicaties betrokken kan zijn, is er een aparte Auteur-Publicatie entiteit gemaakt.

Datamodel in de 2^e normaalvorm:

Het datamodel in de 1^e normaalvorm voldoet reeds aan de vereisten voor de 2^e normaalvorm en er zijn daarom geen wijzigingen aangebracht.

Datamodel in de 3^e normaalvorm:

GRAFZERK: Grafzerk_ID, Kerk_ID (FK van KERK), Steensoort_ID (FK van STEENSOORT), Slijtage_ID (FK van SLIJTAGE), Positie_ID (FK van POSITIE), Coördinaat_lb_N, Coördinaat_lb_O, Coördinaat_ro_N, Coördinaat_ro_O, Validatie_ID (FK van VALIDATIE), Gebruiker_ID (FK van GEBRUIKER)

STEENSOORT: Steensoort_ID, Steensoort

SLIJTAGE: Slijtage_ID, Slijtage

POSITIE: Positie_ID, Positie

VALIDATIE: Validatie_ID, Validatie, Datum, Onderzoeker_ID (FK van ONDERZOEKER)

BEGRAVENE: Begravene_ID, Voornaam, Doopnaam, Tussenvoegsel, Achternaam, Geboortedag, Geboortemaand, Geboortjaar, Sterfdag, Sterfmaand, Sterfjaar, Validatie, Grafzerk_ID (FK van GRAFZERK)

RELATIE: Begravene_ID (FK van BEGRAVENE), Begravene_ID (FK van BEGRAVENE), Relatiesoort
RELATIESOORT: Relatiesoort_ID, Relatiesoort

KERK: Kerk_ID, Naam, Straat, Huisnummer, Postcode, Stad, Beheerder_ID (FK van BEHEERDER), Begin-datum-actief, Eind-datum-actief, Denominatie, URL

BEHEERDER: Beheerder_ID, Naam, Straat, Huisnummer, Postcode, Stad, Telefoonnummer, Emailadres, URL, KvK nummer

ONDERZOEKER: Onderzoeker_ID, Voornaam, Doopnaam, Tussenvoegsel, Achternaam, Geboortedatum, Sterfdatum, Telefoonnummer, Email, Affiliatie

GEBRUIKER: Gebruiker_ID, Voornaam, Doopnaam, Tussenvoegsel, Achternaam, Geboortedatum, Sterfdatum, Telefoonnummer, Email,

PUBLICATIE: Publicatie_ID, Titel, Tijdschrift, Publicatiedatum, DOI

AUTEUR: Publicatie_ID (FK van PUBLICATIE), Onderzoeker_ID (FK van PARTIJ)

PUBLICATIE-GRAFZERK: Publicatie_ID (FK van Publicatie), Grafzerk_ID (FK van Grafzerk)

PROJECT: Project_ID, Titel, Begindatum, Einddatum, Kerk_ID (FK van KERK), Opdrachtgever (FK van OPDRACHTGEVER)

OPDRACHTGEVER: Opdrachtgever_ID, Opdrachtgevernaam, Straat, Huisnummer, Postcode, Stad, Telefoonnummer, Emailadres, URL, KvK nummer

PROJECT-PUBLICATIE: Publicatie_ID (FK van PUBLICATIE), Project_ID (FK van PROJECT)

Ontwerpbeslissingen:

- Om invoerfouten te voorkomen zijn er losse entiteiten gemaakt voor de Validatie, Steensoort, Slijtage, en Positie.
- Hetzelfde geldt voor Relatiesoort.
- Om bij te houden door wie de informatie over de grafzerken gevalideerd is, is hier een aparte entiteit voor gemaakt.

Datamodel in de 1^e Meeuwsen normaalvorm:

GRAFZERK: Grafzerk_ID, Kerk_ID, Steensoort_ID (FK van STEENSOORT), Slijtage_ID (FK van SLIJTAGE), Positie_ID (FK van POSITIE), Coördinaat_lb_N, Coördinaat_lb_O, Coördinaat_ro_N, Coördinaat_ro_O, Project_ID (FK van PROJECT), Invoerder (Partij_ID FK van Partij), Validatie_ID (FK van VALIDATIE)

STEENSOORT: Steensoort_ID, Steensoort

SLIJTAGE: Slijtage_ID, Slijtage

POSITIE: Positie_ID, Positie

VALIDATIE: Validatie_ID, Validatie, Datum, Partij_ID (FK van PARTIJ)

KERK: Kerk_ID, Naam, Straat, Huisnummer, Postcode, Stad, Beheerder (Partij_ID FK van PARTI), Begin-datum-actief, Eind-datum-actief, Denominatie, URL

BEGRAVENE: Grafzerk_ID (FK van GRAFZERK), Partij_ID (FK van PARTIJ), Validatie_ID (FK van VALIDATIE)

RELATIE: Persoon 1 (Partij_ID FK van PARTIJ), Persoon 2 (Partij_ID FK van PARTIJ), Relatiesoort_ID (FK van RELATIESOORT)

RELATIESOORT: Relatiesoort_ID, Relatiesoort

BEDRIJF: Bedrijf_ID, Bedrijfsnaam, Straat, Huisnummer, Postcode, Stad, Telefoonnummer, Email, URL, KvK nummer

PERSOON: Persoon_ID, Voornaam, Doopnaam, Tussenvoegsel, Achternaam, Geboortedag, Geboortemaand, Geboortjaar, Sterfdag, Sterfmaand, Sterfjaar, Telefoonnummer, Email, Affiliatie

PARTIJ: Partij_ID, Persoon_ID (FK van PERSOON), Bedrijf_ID (FK van BEDRIJF), Partijsoort_ID (FK van PARTIJSOORT)

PARTIJSOORT: Partijsoort_ID, Partij-soort, Mutatie_ID (FK van MUTATIE)

PUBLICATIE: Publicatie_ID, Titel, Tijdschrift, Publicatiedatum, DOI

AUTEUR: Publicatie_ID (FK van PUBLICATIE), Onderzoeker (Partij_ID FK van PARTIJ)

PUBLICATIE-GRAFZERK: Publicatie_ID (FK van PUBLICATIE), Grafzerk_ID (FK van GRAFZERK)

PROJECT: Project_ID, Titel, Begindatum, Einddatum, Kerk_ID (FK van KERK), Opdrachtgever (Partij_ID FK van PARTIJ)

PROJECT-PUBLICATIE: Publicatie_ID (FK van PUBLICATIE), Project_ID (FK van PROJECT)

Ontwerpbeslissingen:

- De beheerder van een kerk en een opdrachtgever voor een project zijn beide bedrijven en zijn daarom tot één entiteit Bedrijf gemaakt.
- De gebruikers, onderzoekers, auteurs en begravenen zijn alle personen en voor hen kan (grotendeels) dezelfde informatie worden opgevraagd. Deze zijn daarom onder de entiteit Persoon geplaatst.
- Bijkomend voordeel van de generalisatie van personen is dat er nu ook een relatie tussen levende personen (gebruikers/onderzoekers) en begravenen kan worden gemaakt.

Datamodel in de 2^e Meeuwsen normaalvorm:

GRAFZERK: Grafzerk_ID, Kerk_ID, Steensoort_ID (FK van STEENSOORT), Slijtage_ID (FK van SLIJTAGE), Positie_ID (FK van POSITIE), Coördinaat_lb_N, Coördinaat_lb_O, Coördinaat_ro_N, Coördinaat_ro_O, Project_ID (FK van PROJECT), Validatie_ID (FK van VALIDATIE), Mutatie_ID (FK van MUTATIE)

STEENSOORT: Steensoort_ID, Steensoort, Mutatie_ID (FK van MUTATIE)

SLIJTAGE: Slijtage_ID, Slijtage, Mutatie_ID (FK van MUTATIE)

POSITIE: Positie_ID, Positie, Mutatie_ID (FK van MUTATIE)

VALIDATIE: Validatie_ID, Validatie, Datum, Partij_ID (FK van PARTIJ)

KERK: Kerk_ID, Naam, Straat, Huisnummer, Postcode, Stad, Beheerder (Partij_ID, FK van PARTIJ), Begin-datum-actief, Eind-datum-actief, Denominatie, URL, Mutatie_ID (FK van MUTATIE)

BEGRAVENE: Grafzerk_ID (FK van GRAFZERK), Partij_ID (FK van PARTIJ), Validatie_ID (FK van VALIDATIE), Mutatie_ID (FK van MUTATIE)

RELATIE: Persoon 1 (Partij_ID FK van PARTIJ), Persoon 2 (Partij_ID FK van PARTIJ), Relatiesoort_ID (FK van RELATIESOORT)

RELATIESOORT: Relatiesoort_ID, Relatiesoort, Mutatie_ID (FK van MUTATIE)

BEDRIJF: Bedrijf_ID, Bedrijfsnaam, Straat, Huisnummer, Postcode, Stad, Telefoonnummer, Email, URL, KvK nummer, Mutatie_ID (FK van MUTATIE)

PERSOON: Persoon_ID, Voornaam, Doopnaam, Tussenvoegsel, Achternaam, Geboortedag, Geboortemaand, Geboortejaar, Sterfdag, Sterfmaand, Sterfjaar, Telefoonnummer, Email, Affiliatie

PARTIJ: Partij_ID, Persoon_ID (FK van PERSOON), Bedrijf_ID (FK van BEDRIJF), Partijsoort_ID (FK van PARTIJSOORT), Mutatie_ID (FK van MUTATIE)

PARTIJSOORT: Partijsoort_ID, Partij-soort, Mutatie_ID (FK van MUTATIE)

PUBLICATIE: Publicatie_ID, Titel, Tijdschrift, Publicatiedatum, DOI, Mutatie_ID (FK van MUTATIE)

AUTEUR: Publicatie_ID (FK van PUBLICATIE), Onderzoeker_ID (Partij_ID, FK van PARTIJ), Mutatie_ID (FK van MUTATIE)

GRAFZERK-PUBLICATIE: Publicatie_ID (FK van PUBLICATIE), Grafzerk_ID (FK van GRAFZERK), Mutatie_ID (FK van MUTATIE)

PROJECT: Project_ID, Titel, Begindatum, Einddatum, Kerk_ID (FK van KERK), Opdrachtgever (Partij_ID, FK van PARTIJ), Mutatie_ID (FK van MUTATIE)

PROJECT-PUBLICATIE: Publicatie_ID (FK van PUBLICATIE), Project_ID (FK van PROJECT), Mutatie_ID (FK van MUTATIE)

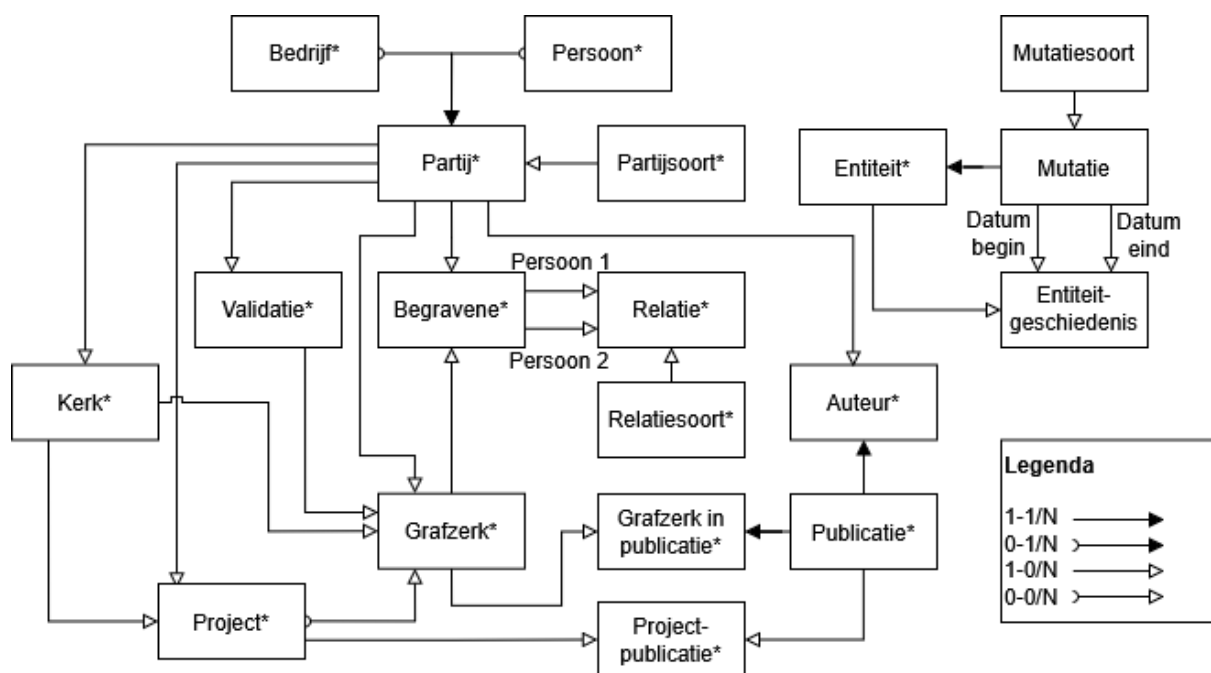
MUTATIE: Mutatie_ID, Mutatiesoort_ID (FK van MUTATIESOORT), Datum

MUTATIESOORT: Mutatiesoort_ID, Mutatiesoort

ENTITEITGESCHIEDENIS: Entiteit_ID, Mutatie_ID (FK van MUTATIE - begin mutatie), Mutatie_ID (FK van MUTATIE - eind mutatie), overige attributen

Ontwerpbeslissingen:

- Er is gekozen om voor alle entiteiten de mutatiegeschiedenis bij te houden.



Figuur 2 ERD van het LDM.

Metadata

Voor ieder attribuut uit het datamodel in de 2^e Meeuwsen normaalvorm is een beslissing genomen om wat voor data type het gaat, de maximale lengte van de invoerwaarde, of dit attribuut leeg mag worden gelaten bij het invoeren van een nieuwe rij, en of het attribuut een Primary- of Foreign Key is en in het laatste geval, naar welk attribuut van een externe tabel verwezen wordt. Daarnaast is er voor iedere attribuut een korte omschrijving geschreven. Deze informatie is in onderstaande tabel overzichtelijk gemaakt.

Tabel 3 De metadata voor het voorgestelde datamodel voor Grafzerken Nederland.

Naam	Type	NULL	Key	Sleutel verwijzing	Omschrijving
tblGRAFZERKEN					
Grafzerk_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Kerk_ID	INT (11)	NO	FK	tblKerk.Kerk_ID	Link naar gerelateerd kerk.
Steensoort_ID	INT (11)	YES	FK	tblSteensoort.Steensoort_ID	Link naar gerelateerd steensoort.
Slijtage_ID	INT (11)	YES	FK	tblSlijtage.Slijtage_ID	Link naar gerelateerd slijtage.
Positie_ID	INT (11)	YES	FK	tblPositie.Positie_ID	Link naar gerelateerd positie.
Coördinaat_lb_N	DECIMAL (3,6)	NO			Noorderbreedte coördinaat van de linkerbovenhoek van de grafzerk in decimale graden.
Coördinaat_lb_O	DECIMAL (3,6)	NO			Oosterlengte coördinaat van de linkerbovenhoek van de grafzerk in decimale graden.
Coördinaat_ro_N	DECIMAL (3,6)	NO			Noorderbreedte coördinaat van de rechteronderhoek van de grafzerk in decimale graden.
Coördinaat_ro_O	DECIMAL (3,6)	NO			Oosterlengte coördinaat van de rechteronderhoek van de grafzerk in decimale graden.

Project_ID	INT (11)	YES	FK	tblProject.Project_ID	Link naar gerelateerd project.
Gebruiker	INT (11)	NO	FK	tblPartij.Partij_ID	Link naar de gerelateerde gebruiker die de informatie over de grafzerk heeft ingevoerd.
Validatie	INT (11)	NO	FK	tblValidatie.Validatie_ID	Link naar gerelateerd validatie.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblBEGRAVENE					
Begravene_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Partij_ID	INT (11)	NO	FK	tblPartij.Partij_ID	Link naar de gerelateerde persoon.
Validatie_ID	INT (11)	NO	FK	tblValidatie.Validatie_ID	Link naar gerelateerd validatie.
Mutatie_ID	INT (11)	YES	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblSTEENSOORT					
Steensoort_ID	INT (11)	NO	PK		Unieke ID van Steensoort. Wordt automatisch gegenereerd.
Steensoort	VARCHAR (34)	NO			De steensoort waaruit de grafzerk gehouden is.
tblSLIJTAGE					
Slijtage_ID	INT (11)	NO	PK		Unieke ID van Slijtage. Wordt automatisch gegenereerd.
Slijtage	VARCHAR (34)	NO			De slijtage fase van de grafzerk.
tblPOSITIE					
Positie_ID	INT (11)	NO	PK		Unieke ID van Positie. Wordt automatisch gegenereerd.
Positie	VARCHAR (34)	NO			De positie van de grafzerk (liggend of staand).
tblVALIDATIE					
Validatie_ID	INT (11)	NO	PK		Unieke ID van Validatie. Wordt automatisch gegenereerd.
Validatie	VARCHAR (34)	NO			De validatie status van de grafzerk of begravene informatie.
Datum	DATA	NO			Datum waarop de informatie van de grafzerk is gevalideerd.
Gevalideerd-door	INT (11)	NO	FK	tblPartij.Part_ID	Link naar de gerelateerde partij die de validatie heeft uitgevoerd.
tblKERK					
Kerk_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Naam	VARCHAR (34)	NO			Naam van de kerk.
Straat	VARCHAR (34)	NO			Straat waarop de kerk zich bevindt.
Huisnummer	INT (11)	NO			Huisnummer waarop de kerk zich bevindt.
Postcode	VARCHAR (34)	NO			Postcode waarop de kerk zich bevindt.
Stad	VARCHAR (34)	NO			Stad waarin de kerk zich bevindt.

Begin-datum-actief	DATE	YES			Begindatum van begravingen in/ rondom de kerk.
Eind-datum-actief	DATE	YES			Einddatum van begravingen in/ rondom de kerk.
Denominatie	VARCHAR (34)	NO			Religieuze stroom van de kerk.
Beheerder	INT (11)	NO	FK	tblPartij.Partij_ID	Link naar de gerelateerde partij die de beheerder is van de kerk.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblRELATIE					
Persoon_1	INT (11)	NO	FK	tblPartij.Partij_ID	Link naar de eerste gerelateerde persoon in de relatie.
Persoon_2	INT (11)	NO	FK	tblPartij.Partij_ID	Link naar de tweede gerelateerde persoon in de relatie.
Relatiesoort_ID	INT (11)	NO	FK	tblRelatie.Relatiesoort_ID	Link naar de relevante gerelateerde relatiesoort.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblREALTIESOORT					
Relatiesoort_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Relatiesoort	VARCHAR (34)				Omschrijving van de relatiesoort.
tblBEDRIJF					
Bedrijf_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Bedrijfsnaam	VARCHAR (34)	NO			Naam van het bedrijf.
Straat	VARCHAR (34)	NO			Straat waarop het bedrijf geregistreerd is.
Huisnummer	INT (11)	NO			Huisnummer waarop het bedrijf geregistreerd is.
Postcode	VARCHAR (34)	NO			Postcode waarop het bedrijf geregistreerd is.
Stad	VARCHAR (34)	NO			Stad waarin het bedrijf geregistreerd is.
Telefoonnummer	INT (13)	NO			Telefoonnummer waarop het bedrijf te bereiken is.
Emailadres	VARCHAR (34)	NO			Emailadres waarop het bedrijf te bereiken is.
URL	VARCHAR (34)	YES			URL naar de website van het bedrijf.
KvK nummer	INT (8)	YES			Het KvK nummer van het bedrijf.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblPERSOON					
Persoon_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Voornaam	VARCHAR (34)	YES			Voornaam van de persoon.
Doopnaam	VARCHAR (34)	YES			Doopnaam van de persoon.

Tussenvoegsel	VARCHAR (34)	YES			Tussenvoegsel in de naam van de persoon.
Achternaam	VARCHAR (34)	YES			Achternaam van de persoon.
Geboortedag	INT (2)	YES			Geboortedag van de persoon.
Geboortemaand	INT (2)	YES			Geboortemaand van de persoon.
Geboortejaar	INT (4)	YES			Geboortejaar van de persoon.
Sterfdag	INT (2)	YES			Sterfdag van de persoon.
Sterfmaand	INT (2)	YES			Sterfmaand van de persoon.
Sterfjaar	INT (4)	YES			Sterfjaar van de persoon.
Telefoonnummer	INT (13)	YES			Telefoonnummer waarop de persoon te bereiken is. Dit kan een buitenlands nummer zijn.
Emailadres	VARCHAR (34)	YES			Emailadres waarop de persoon te bereiken is.
Affiliatie	VARCHAR (34)	YES			(Onderzoeks)instituut waaraan de persoon geaffilieerd is.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblPARTIJ					
Partij_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Bedrijf_ID	INT (11)	YES	FK	tblBedrijf.Bedrijf_ID	Link naar het gerelateerde bedrijf.
Persoon_ID	INT (11)	YES	FK	tblPersoon.Persoon_ID	Link naar de gerelateerde persoon.
Partijsoort_ID	INT (11)	NO	FK	tblPartijsoort.Partijsoort_ID	Link naar de gerelateerde partijsoort.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblPARTIJSOORT					
Partijsoort_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Partijsoort	VARCHAR (34)				Omschrijving van de partijsoort.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblPUBLICATIE					
Publicatie_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Titel	VARCHAR (34)				Titel van de publicatie.
Tijdschrift	VARCHAR (34)				Tijdschrift waarin de publicatie verschenen is.
Publicatiedatum	DATE				Datum waarop de publicatie beschikbaar kwam.
DOI	VARCHAR (34)				De unieke Digital Object Identifier waarmee de publicatie online gevonden kan worden.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de

					mutatiegeschiedenis van de entiteit.
tblAUTEUR					
Auteur	INT (11)	NO	FK	tblPartij.Partij_ID	Link naar de partij van de gerelateerde auteur.
Publicatie_ID	INT (11)	NO	FK	tblPublicatie.Publicatie_ID	Link naar de gerelateerde publicatie waar de partij auteur bij is.
tblGRAFZERK-PUBLICATIE					
Grafzerk_ID	INT (11)	NO	PK	tblGrafzerk.Grafzerk_ID	Link naar de gerelateerde grafzerk.
Publicatie_ID	INT (11)	NO	PK	tblPublicatie.Publicatie_ID	Link naar de publicatie waarin de gerelateerde grafzerk onderdeel is van de data.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblPROJECT					
Project_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Titel	VARCHAR (34)	NO			Titel van het project.
Begindatum	DATE	NO			Datum waarop het project begonnen is.
Einddatum	DATE	YES			Datum waarop het project afgerond is.
Kerk_ID	INT (11)	NO	FK	tblKerk.Kerk_ID	Link naar de gerelateerde kerk.
Opdrachtgever	INT (11)	NO	FK	tblPartij.Partij_ID	Link naar de gerelateerde opdrachtgever.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblPROJECT-PUBLICATIE					
Project_ID	INT (11)	NO	PK	tblProject.Project_ID	Link naar gerelateerd project.
Publicatie_ID	INT (11)	NO	PK	tblPublicatie.Publicatie_ID	Link naar de gerelateerde publicatie.
Mutatie_ID	INT (11)	NO	FK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van de mutatiegeschiedenis van de entiteit.
tblMUTATIE					
Mutatie_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Mutatiesoort_ID	INT (11)	NO	FK	tblMutatiesoort.Mutatiesoort_ID	Link naar de gerelateerde mutatiesoort.
Datum	DATE	NO			Datum waarop de mutatie uitgevoerd is.
tblMUTATIESOORT					
Mutatiesoort_ID	INT (11)	NO	PK		Unieke ID van Grafzerken. Wordt automatisch gegenereerd.
Mutatiesoort	VARCHAR (34)	NO			Omschrijving van de mutatiesoort.
tblENTITEITGESCHIEDENIS					

Entiteit_ID	INT (11)	NO	PK	tblEntiteit.Entiteit_ID	Link het unieke ID van de relevante entiteit.
Mutatie_ID	INT (11)	NO	PK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van begindatum van de staat van de entiteit.
Mutatie_ID	INT (11)	YES	PK	tblMutatie.Mutatie_ID	Link naar de mutatietabel voor het bijhouden van einddatum van de staat van de entiteit.
Overige attributen					

Data principes

Voor het datamodel en de daaruit volgende database van Grafzerken Nederland zijn de Data principes in het achterhoofd gehouden. De Data principes zijn gecategoriseerde aspecten van data die algemeen erkend worden als cruciale richtlijnen over dataopslag en dataprovisie. De data principes zijn opgedeeld in de drie hoofdcategorieën: Governance, Kwaliteit, en Beschikbaarheid. Het voorgestelde datamodel voor GN voldoet bijna volledig aan alle data principes.

Tabel 4 Een overzicht van de data principes en hoe deze bij het voorgestelde datamodel worden toegepast.

Categorie	Subcategorie	Situatie in het voorgestelde datamodel
Governance	Data hebben een eigenaar	De data zijn in beheer bij Grafzerken Nederland.
	Data hebben een definitie	Voor alle entiteiten en attributen is een omschrijving gegeven van de inhoud en het datatype.
	Data voldoen aan de regels	Er worden persoonsgegevens opgeslagen van personen die gevoelig kunnen zijn (zoals namen en geboortedata). Er zal onderzocht moeten worden of deze data noodzakelijk is en hoe er met deze data moet worden omgegaan.
Kwaliteit	Data zijn accuraat	Alle ingevoerde data worden gecontroleerd door hiervoor aangewezen personen (de onderzoekers). Het is mogelijk dat incorrecte data worden ingevoerd door gebruikers, maar deze worden vervolgens door de onderzoekers gevalideerd. De nauwkeurigheid van persoonlijke gegevens van de gebruikers kan niet gegarandeerd worden.
	Data zijn volledig	De data kunnen niet altijd volledig worden ingevuld, doordat grafzerken versleten kunnen zijn en de persoonsgegevens dan niet afgeleid kunnen worden. Omdat dit daarnaast een project is dat alle al bestaande grafzerken wil documenteren, maar nog lang niet voor alle kerken/grafzerken documentatie is zal er veel data ontbreken die nog door onderzoekers/ gebruikers moet worden toegevoegd.
	Data zijn consistent	Doordat er zo veel mogelijk aparte entiteiten zijn aangemaakt voor attributen die als meerkeuze menu kunnen dienen (zoals Steensoort) wordt voorkomen dat de opties op verschillende manieren gespeld worden. Daarnaast worden alle ingevoerde data over de grafzerken door bevoegde personen gevalideerd en wordt voor alle entiteiten een mutatiegeschiedenis bijgehouden.

Beschikbaarheid	Data zijn tijdig	Er is normaliter geen tijdsdruk voor het invoeren van de data, aangezien Grafzerken Nederland een centrale, publieke database is. Data die door gebruikers worden ingevoerd zullen handmatig moeten worden gecontroleerd en er kan daardoor niet gegarandeerd worden dat deze data snel gevalideerd en daardoor beschikbaar kan worden gemaakt.
	Data zijn toegankelijk	Omdat de data voor het publiek beschikbaar zal worden gemaakt via een app en een website zal de meeste data toegankelijk zijn. Enkele entiteiten, zoals de mutatiegeschiedenis en persoonsgegevens van gebruikers en onderzoekers zullen enkel beschikbaar zijn voor geautoriseerde personen.
	Data zijn beschermd	In het systeem zal ontworpen worden dat gebruiker accounts direct aangemaakt kunnen worden. Personen die een Onderzoekers-account nodig hebben zullen eerst een persoonsverificatie moeten doorlopen. Zo worden cruciale aspecten van het datamodel beschermd. Dankzij het bijhouden van de mutatiegeschiedenis kunnen data hersteld worden.

De data van GN worden secuur beheerd en voldoen aan bijna alle gestelde eisen van de Data principes. Voornamelijk de persoonsgegevens van gebruikers zijn een obstakel waar meer onderzoek naar moet worden gedaan. De nauwkeurigheid van deze informatie kan namelijk moeilijk gecontroleerd worden en ook aan de opslag van deze gegevens zullen regels verbonden zijn die gecontroleerd moeten worden. Voor de toegankelijkheid en bescherming van de data moeten degelijke systemen worden opgezet, om te voorkomen dat onbevoegden toegang krijgen tot gevoelige data. Deze systemen vallen echter buiten de scope van dit rapport en zullen onderdeel moeten zijn van een volgende fase in het opzetten van de database.

Op alle andere vlakken voldoet het voorgestelde datamodel naar behoren aan de Data principes. De data hebben genuanceerde definities gekregen en er is een validatie proces ontworpen om de nauwkeurigheid van de data te waarborgen. Door middel van het bijhouden van een mutatiegeschiedenis kan ook gegarandeerd worden dat bij ongewenste aanpassingen een vorige versie van de data hersteld kan worden. Terwijl de ‘buitenschil’ van de database nog verder ontwikkeld wordt (zoals de opslag van gebruikers gegevens) kan er dus al wel begonnen worden aan de productie van het fysieke datamodel voor de aspecten van het Product domein. Zodra dat ontworpen is kan de database op secure wijze gevuld worden met gegevens over grafzerken voordat de database openbaar toegankelijk wordt gemaakt.

Conclusie

In opdracht van Grafzerken Nederland is in dit rapport een mogelijk datamodel beschreven, dat gebruikt kan worden als basis voor het opstellen van een solide en efficiënte database. Alle gegevens die voor Grafzerken Nederland van belang zijn zijn in het model opgenomen. Iedere entiteit en de daaraan gerelateerde attributen komen in het LDM slechts éénmalig voor. Fouten die ontstaan doordat data meerdere malen moeten worden ingevoerd worden hiermee voorkomen. In het LDM zijn ook enkele entiteiten opgenomen die enkel bestaan uit twee Foreign keys, waarmee veel-tot-veel relaties toch mogelijk worden gemaakt. Tevens is er via de Partij-entiteit generalisatie toegepast, waardoor bedrijven en personen meerdere rollen kunnen hebben. Doordat gebruikers uit het publiek, onderzoekers van GN en begraven personen gegeneraliseerd zijn tot de entiteit Persoon is het zelfs mogelijk geworden om in het systeem relaties tussen begraven en nog levende personen te documenteren.

Het datamodel voldoet grotendeels aan de data principes, die ingaan op de Governance, Kwaliteit en Beschikbaarheid. In het beschreven datamodel voldoen de data aan de regels en worden deze strikt beheerd. Omdat een onderdeel van Grafzerken Nederland burgerwetenschap is, zal er ook veel data worden opgenomen die door het algemene publiek wordt aangedragen. Er kan niet gegarandeerd worden dat deze data compleet en correct is. Er is daarom besloten dat aangedragen data gevalideerd moet worden door geautoriseerde personen. Na controle door één van deze personen wordt de status van de data aangepast en kunnen de data niet meer door het publiek worden aangepast. Op deze manier wordt ervoor gezorgd dat de data beschikbaar doch beschermd zijn.

Het datamodel kan direct in gebruik worden genomen voor het produceren van een fysiek datamodel. Het logisch datamodel en de tabel met de metadata kunnen hiervoor als basis dienen. We hebben er alle vertrouwen in dat de data daarmee op de meest optimale manier opgeslagen en verwerkt worden.