



Begrippenkader

Bij het werkspoor datamaturiteit in de
stedelijke organisatie

Wergroep datamaturiteit - Kenniscentrum Vlaamse Steden



Voorwaarden voor het begrippenkader

Begrippen moeten laagdrempelig zijn, begrijpbaar voor iedereen

Weg van technologie en jargon

Aangepast aan de context van het lokaal bestuur (weg van bedrijfsjargon)

Keuze van de begrippen:

- Geleid door DAMA – DMBOK
- Afgestemd op bestaande Vlaamse en Nederlandse begrippenkaders:
 - [Vlaamse Terminologie voor Informatiemanagement](#)
 - [Begrippenkader GEMeentelijke Model Architectuur \(GEMMA\) - NL](#)

Ultiem willen we een referentie-lexicon samenstellen

- Een levend document
- Centraal beschikbaar

Algemene begrippen

Datageletterdheid

Datageletterdheid is **het vermogen van een individu** om de benodigde data te vinden, de kwaliteit ervan te beoordelen, data/informatie te lezen, te evalueren, te beargumenteren en ermee te werken. Het is een vaardigheid die iedereen in staat stelt de juiste vragen te stellen, kennis te vergaren, te beoordelen en daar doeltreffend met anderen over te kunnen communiceren.

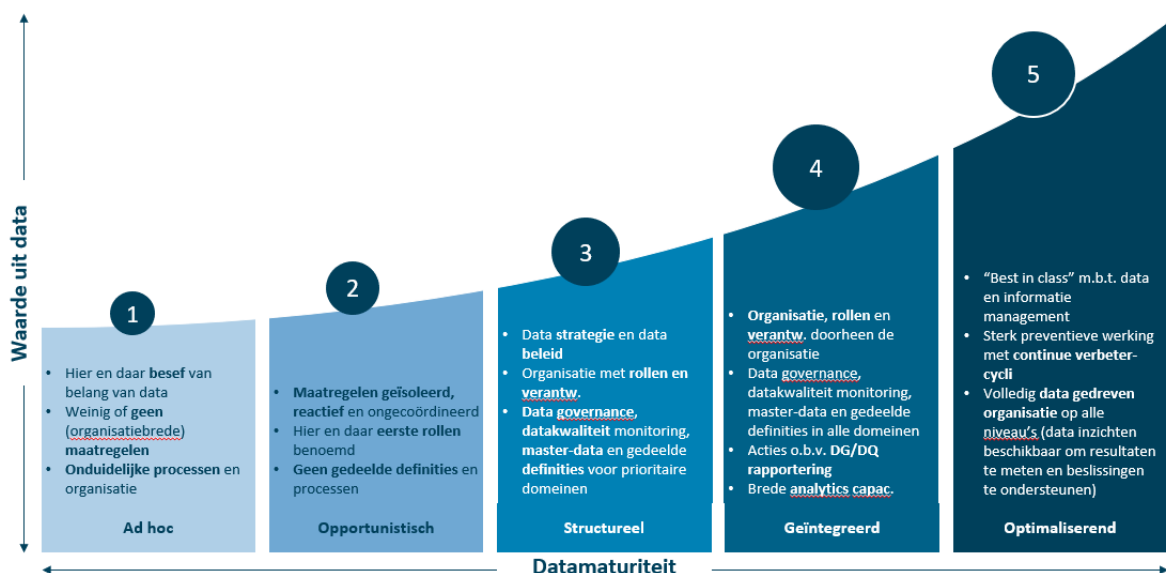
Datamaturiteit

Data-maturiteit verwijst naar de volwassenheid van een organisatie met betrekking tot het beheer, gebruik en strategische toepassing van data.

Focus: Het concentreert zich op de processen, systemen en beleidsmaatregelen die organisaties implementeren om gegevens effectief te verzamelen, op te slaan, te analyseren en te gebruiken.

Toepassing: Data-maturiteit is relevant op organisatieniveau en heeft betrekking op de mate waarin een organisatie in staat is om datagestuurde beslissingen te nemen, gegevenskwaliteit te waarborgen, en een gegevensgedreven cultuur te bevorderen.

Datamaturiteitsmodel



Figuur 1: Een voorbeeld van een maturiteitsmodel (Stad Gent)

Datageletterdheid vs Datamaturiteit

In het kort:

data-geletterdheid richt zich op individuele vaardigheden en kennis met betrekking tot gegevens, terwijl data-maturiteit zich richt op de volwassenheid en effectiviteit van datamanagement binnen een organisatie.

Beide aspecten zijn echter belangrijk voor het realiseren van de volledige potentie van gegevens in zowel persoonlijke als zakelijke contexten.

Aspecten van Data Governance

Data Governance

Data governance heeft als primair doel om de **consistentie, betrouwbaarheid en ongehinderde beschikbaarheid van uw data te garanderen**, zodat de werkprocessen in uw organisatie optimaal ondersteund worden.

- Het is een set aan **afspraken en gekozen standaarden** die gezamenlijk vastleggen op welke manier uw organisatie met uw data omgaat.
- Data governance betreft een aantal verschillende werkzaamheden rondom het **gebruik en beveiligen** van data binnen een organisatie, alsmede de **controle over de processen** die dit mogelijk maken.
- Het heeft niet alleen met de technische en logistieke kant van het opslaan van data te maken. Het betreft ook een noodzakelijke combinatie voor het **naleven van wetten, regels en procedures**.
- Het draait om **de mensen** die met de data werken. Het is dus ook een set van **rollen en verantwoordelijkheden**.

Data governance creëert het overkoepelend kader dat het mogelijk maakt alle aspecten van datamanagement structureel aan te pakken en richt zich op de ontwikkeling van een datacultuur en datamature organisatie.

Vanuit dit opzicht draait data governance **meer om de mensen dan om de data**.

Data Management

- Doel van data management is om ervoor te zorgen dat de aanwezige data compleet, betrouwbaar en op tijd beschikbaar is voor toepassingen die van die data gebruik maken. Met andere woorden: data management gaat over het operationeel maken van de beschikbare data binnen de organisatie.
- Data management gaat over het **onderhouden, actualiseren, beheren en beveiligen** van data tijdens de hele [datalevenscyclus](#) met als doel het maximaliseren van de waarde van data als bedrijfsmiddel.

Datalevenscyclus

De datalevenscyclus verwijst naar de verschillende fasen die data doorgaat gedurende zijn hele bestaan, vanaf het moment van creatie of verwerving tot het moment van verwijdering of archivering. Deze levenscyclus omvat verschillende processen en activiteiten die betrekking hebben op het beheer van data binnen een organisatie. Hoewel specifieke fasen en benamingen kunnen variëren, omvatten de algemene stappen in de datalevenscyclus meestal het volgende:

- **Creatie of invoer:** Het punt waarop data wordt gegenereerd, gecreëerd of verzameld.
- **Opslag:** De fase waarin data wordt opgeslagen op een bepaalde locatie, zoals databases, servers of cloudopslag.
- **Gebruik en verwerking:** Data wordt geanalyseerd, bewerkt, en gebruikt voor verschillende doeleinden, zoals rapportage, besluitvorming of operationele processen.
- **Ontsluiting:** Het ontsluiten van data naar relevante partijen, applicaties of systemen.
- **Archivering:** Wanneer data niet langer actief wordt gebruikt maar nog steeds bewaard moet blijven vanwege wettelijke, regelgevende of bedrijfsvereisten, wordt het gearchiveerd.
- **Vernietiging of verwijdering:** Het verwijderen van data die niet langer nodig is, volgens vastgestelde procedures en richtlijnen.

Informatie-architectuur

De **informatie-architectuur** geeft een overzicht van de inhoudelijke relaties en samenhang tussen toepassingen en gegevensverzamelingen onderling.

De informatie-architectuur omvat de applicatie-architectuur en de data-architectuur.

Applicatie-architectuur

De **applicatie-architectuur** is een beschrijving van de structuur en interactie van de applicaties die belangrijke bedrijfsfuncties ondersteunen en gegevens beheren.

De applicatie-architectuur vormt samen met de [data-architectuur](#) de [informatie-architectuur](#).

Data-architectuur

De **data-architectuur** geeft een schematisch overzicht van de aanwezige en benodigde data-componenten in een organisatie.

Het is een beschrijving van de belangrijkste soorten en bronnen van data in de organisatie.

Eén van de aspecten binnen het opstellen van een data-architectuur is de [datamodellering](#).

De data-architectuur vormt samen met de [applicatie-architectuur](#) de [informatie-architectuur](#).

Business Informatie Model

Een Business Informatie Model (BIM) is een **conceptuele representatie of raamwerk** dat bepaalt hoe informatie georganiseerd, gestructureerd en gerelateerd is binnen een organisatie.

Het dient als een blauwdruk voor het beheer en benutten van informatie-assets ter ondersteuning van de doelstellingen en besluitvormingsprocessen van de organisatie.

Het BIM richt zich op **het conceptuele niveau** en biedt een **overzicht** van de belangrijkste entiteiten, attributen, relaties en processen binnen een organisatie.

Een BIM heeft tot doel de organisatie te helpen **begrijpen hoe informatie wordt georganiseerd, gestructureerd en gebruikt** om de bedrijfsdoelstellingen te ondersteunen.

Het BIM wordt vaak gebruikt als een communicatiemiddel tussen zakelijke belanghebbenden en IT-teams, waardoor een gemeenschappelijk begrip ontstaat van de bedrijfsgegevens en processen.

Datamodellering en -ontwerp

Het proces van het ontdekken, analyseren, visualiseren en communiceren van datavereisten in een precieze vorm die het [datamodel](#) wordt genoemd.

Dat proces beschrijft de verschillende data-entiteiten, hun attributen en hun onderlinge relatie, binnen één proces of applicatie.

Dit proces is een belangrijk onderdeel binnen het opstellen van een [data-architectuur](#).

Datamodel

Met een datamodel, of gegevensmodel, wordt beschreven hoe de gegevens in een systeem gestructureerd zijn.

Men onderscheidt het conceptueel datamodel, het logisch datamodel en het fysiek datamodel.

- Het **conceptuele datamodel** beschrijft de structuur van en de relaties tussen de conceptuele gegevensobjecten, genaamd entiteiten.

- Het **logische datamodel** beschrijft de structuur van en de referenties tussen de logische gegevensobjecten, genaamd tabellen. Het conceptuele model is verbonden aan het logische model doordat entiteiten worden omgezet in tabellen (of preciezer: tabeldefinities), en doordat relaties worden omgezet in foreign key constraints.

Het **fysieke datamodel** beschrijft de manier waarop gegevens in een individuele datacomponent zijn opgeslagen.

Dataopslag en -beheer

De opslag van data met bijhorend het ontwerp, de technische implementatie en de bijhorende ondersteuning om de waarde ervan te maximaliseren.

Databeheer biedt ondersteuning gedurende de gehele [datalevenscyclus](#), van planning tot en met het verwijderen van data.

We kunnen daarbij het [datawarehouse](#) zien als één manier om aan data opslag te doen.

Data-integratie en interoperabiliteit

Data-integratie is het proces van het combineren van gegevens uit verschillende bronnen (binnen een organisatie) om een completer en meer samenhangend beeld van de data te verkrijgen.

Het doel is om verschillende datasets te verbinden, zodat ze effectiever kunnen worden gebruikt voor analyse, rapportage en besluitvorming.

Interoperabiliteit verwijst naar de mogelijkheid van verschillende systemen, applicaties of componenten om effectief en naadloos met elkaar te communiceren, samen te werken en informatie uit te wisselen. Het doel van interoperabiliteit is het vergemakkelijken van de gegevensuitwisseling en samenwerking tussen verschillende technologische systemen.

Daarbij is een onderscheid te maken tussen

- technische interoperabiliteit
- semantische interoperabiliteit
- organisationele interoperabiliteit
- juridische interoperabiliteit

Informatiebeheer

Informatiebeheer verwijst naar het proces van het organiseren, duurzaam beheren, bewaren en beschermen van informatie binnen een organisatie.

Het is het creëren, beheren en bewaren van alle informatie die voortkomt uit werkprocessen, taken en functies waarvoor de lokale overheid bestuurlijk verantwoordelijk is.

Het omvat alle activiteiten die gericht zijn op het effectief verzamelen, duurzaam beheren, bewaren, verwerken en ontsluiten van informatie om de doelstellingen van een organisatie te ondersteunen.

Belangrijke deelaspecten van informatiebeheer zijn:

- [Informatieveiligheid](#)
- [Documentbeheer](#)
- [Contentmanagement](#)

Informatieveiligheid

Zorgt ervoor dat de privacy en vertrouwelijkheid van de data behouden blijft, dat er geen inbreuk wordt gemaakt op de data en dat de data op de juiste manier worden benaderd.

Belangrijke aspecten van informatieveiligheid zijn:

- **Vertrouwelijkheid:** Garanderen dat informatie alleen toegankelijk is voor geautoriseerde personen.
- **Integriteit:** Zorgen dat informatie accuraat en ongewijzigd blijft. Dit omvat het voorkomen van ongeautoriseerde wijzigingen in gegevens.
- **Beschikbaarheid:** Waarborgen dat informatie beschikbaar is wanneer geautoriseerde gebruikers deze nodig hebben. Dit houdt in dat systemen en gegevens toegankelijk blijven en niet onnodig worden belemmerd.
- **Authenticiteit:** Verzekeren dat de oorsprong van informatie betrouwbaar en verifieerbaar is. Dit is belangrijk om te voorkomen dat informatie wordt gemanipuleerd door ongeautoriseerde bronnen.
- **Fysieke beveiliging:** Beschermen van de fysieke apparaten en infrastructuur die informatie opslaan en verwerken, zoals servers, datacenters en opslagapparaten.
- **Toegangscontrole:** Beperken van de toegang tot informatie op basis van de rollen en verantwoordelijkheden van individuen binnen een organisatie. Dit omvat het gebruik van wachtwoorden, authenticatieprocedures en autorisatiemechanismen.
- **Bewustwording en training:** Het opleiden van medewerkers en gebruikers over de risico's van informatieveiligheid en hen voorzien van de kennis en vaardigheden om veilig met informatie om te gaan.

Documentbeheer

Documentbeheer verwijst naar het proces van het organiseren, opslaan, beheren en volgen van digitale, gedigitaliseerde én analoge documenten binnen een organisatie.

Het doel van documentmanagement is om efficiëntie te bevorderen, informatie gemakkelijk toegankelijk te maken, versiebeheer te waarborgen, de samenwerking te verbeteren en de naleving van regelgeving te garanderen.

Contentmanagement

Contentmanagement verwijst naar het **proces van het creëren, organiseren, publiceren en beheren van digitale content**. Deze content kan verschillende vormen aannemen, waaronder tekst, afbeeldingen, video's, audio en andere multimediale elementen.

Het doel van contentmanagement is om ervoor te zorgen dat de juiste informatie op het juiste moment beschikbaar is voor de juiste doelgroep.

Contentmanagement kan worden ondersteund door specifieke softwaretoepassingen, bekend als **contentmanagementsystemen (CMS)**. Deze systemen bieden tools en functies om het gehele contentmanagementproces te vergemakkelijken, van het maken van content tot het publiceren ervan. CMS'en worden vaak gebruikt voor websites, maar ze kunnen ook worden toegepast in bredere zakelijke contexten waar digitale content een centrale rol speelt.

Referentiegegevens

Referentiegegevens zijn gegevens die worden gebruikt om andere gegevens te identificeren, categoriseren, beschrijven of relateren. Deze gegevens dienen als een standaard of referentiepunt voor het organiseren en begrijpen van andere gegevens in een systeem of context. Referentiedata zijn meestal stabiel en veranderen niet vaak. Ze bieden een gemeenschappelijk kader voor het interpreteren van andere gegevens binnen een organisatie of industrie.

Voorbeelden van referentiegegevens zijn een postcode, klantcode of het rijksregisternummer.

Stamgegevens (Master Data)

Stamgegevens, ook wel bekend als master data, zijn de kerngegevens die essentieel zijn voor de dagelijkse activiteiten van een organisatie. Ze vertegenwoordigen de **fundamentele bedrijfsinformatie** die nodig is om transacties uit te voeren en beslissingen te nemen.

Stamgegevens worden vaak **centraal beheerd** en gebruikt in verschillende bedrijfsprocessen en systemen binnen een organisatie. Het nauwkeurig beheren van stamgegevens is cruciaal voor het waarborgen van consistente informatie en het voorkomen van fouten in operationele processen.

Stamgegevens **vormen een basislaag** voor transactiegegevens en rapportage.

Data opslaan en structureren

Er bestaan verschillende manieren om data op te slaan en te verwerken. Elke manier heeft een ander doel en impact naar manier van ontsluiten en raadplegen van de data.

- **Datawarehouse:**
een (centrale) opslagplaats van actuele en historische gegevens uit één of meer verschillende bronnen met als doel het aanleveren van informatie, inzichten, analyses en rapporten. De gegevens worden gestructureerd bijgehouden, zoals in een echt magazijn in onderverdelingen (die we dan datamarts kunnen noemen).
- **Datalake:**
Een opslagplaats met als doel het capteren en opslaan van verschillende “stromen” gestructureerde en ongestructureerde ruwe data. In een datalake wordt de data zelf in zijn ruwe vorm en ongeorganiseerd bijgehouden. Zoals de naam het zegt kan je het zien als een meer dat gevoed wordt door verschillende rivieren. De toepassingsmogelijkheden zijn onder andere machine learning.
- **Data lakehouse:**
Combineert de beste eigenschappen van een Datawarehouse en een datalake. Je brengt zo structuur in het datalake.
- **Databank:**
Is een gestructureerde verzameling van gegevens die (inhoudelijk) bij elkaar horen en met logische relaties tussen de verschillende componenten ervan.
Een database kan gedefinieerd worden als een gestructureerde verzameling van gegevens.
- **Data Mashup:**
Een data mashup is een integratie van twee of meer gegevenssets in een enkele grafische interface. Een dashboard is een uitstekend voorbeeld van een data mashup. De meeste dashboards halen gegevens uit meerdere, ongelijksoortige bronnen om eindgebruikers in één oogopslag een beeld te geven van hun bedrijfsprestaties.

IOT Data

IoT (Internet of Things)-data verwijst naar de informatie die wordt gegenereerd en verzameld door apparaten die deel uitmaken van het Internet of Things-ecosysteem.

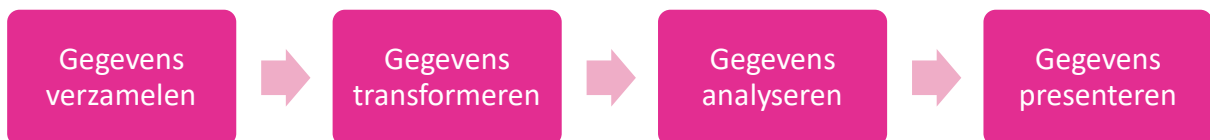
Het Internet of Things verwijst naar het netwerk van onderling verbonden apparaten die zijn uitgerust met sensoren, software en andere technologieën waarmee ze gegevens kunnen verzamelen en uitwisselen via het internet.

Deze apparaten kunnen variëren van alledaagse objecten zoals huishoudelijke apparaten en draagbare apparaten tot industriële machines en slimme stadsinfrastructuur.

Business Intelligence

Het overkoepelende begrip **business intelligence** kun je omschrijven als het omzetten van gegevens naar informatie. Deze informatie helpt organisaties om operationele en strategische beslissingen te nemen en te onderbouwen.

Je kunt het business-intelligenceproces heel eenvoudig voorstellen:



1. Gegevens **verzamelen** uit verschillende interne en externe systemen (ERP, CRM, web, enz.) en eventueel in een datawarehouse plaatsen.
2. De verzamelde gegevens omvormen (**transformeren**) zodat de gegevens uit de verschillende systemen hetzelfde formaat hebben.
3. De verzamelde gegevens **analyseren** en omvormen tot informatie die bruikbaar is voor de organisatie.
4. De gevonden informatie **presenteren** in een **dashboard**.

Datakwaliteitsmanagement

Datakwaliteitsmanagement is het proces van plannen, implementeren en beheren van activiteiten en maatregelen om de **kwaliteit van data** binnen een organisatie te waarborgen zodat data geschikt wordt om een vooropgesteld doel te realiseren.

Het doel is om ervoor te zorgen dat de data die een organisatie gebruikt **accuraat, consistent, volledig, actueel en relevant** is.

Goede datakwaliteit is essentieel voor **betrouwbare** besluitvorming, rapportage en andere bedrijfsprocessen.

Technische begrippen

ETL / ELT

ETL staat voor Extract => Transform => Load

ELT staat voor Extract => Load => Transform

Extract: inlezen van brondata

Transform: aanpassen van de gegevens aan hoe de gegevens gebruikt moeten gaan worden

Load: gegevens ergens naartoe laden (bijvoorbeeld Excel werkblad)

Belangrijkste verschil is het moment waarop de gegevens worden getransformeerd. Dat kan voordat ze naar de doelomgeving worden geladen (ETL) of nadat ze werden opgeladen (ELT)

API

API staat voor Application Programming Interface.

Het is een set definities en protocollen waarmee verschillende softwaretoepassingen met elkaar kunnen communiceren. Een API maakt het mogelijk voor ontwikkelaars om functionaliteit van een bepaalde softwarecomponent te gebruiken zonder dat ze toegang hebben tot de interne broncode.

Kortom, een API fungeert als een interface waarmee verschillende softwarecomponenten met elkaar kunnen praten en informatie kunnen uitwisselen.

Het maakt de ontwikkeling van software efficiënter en bevordert de integratie tussen verschillende systemen.

CRM

CRM staat voor Customer Relationship Management

Het is een softwareoplossing die gericht is op het beheren en optimaliseren van de interacties en relaties tussen een organisatie en haar klanten.

Het omvat het verzamelen, organiseren en analyseren van klantgegevens om een dieper inzicht te verkrijgen in klantbehoeften, voorkeuren en gedrag.

Het uiteindelijke doel van CRM is het verbeteren van klanttevredenheid, het opbouwen van langdurige relaties en het stimuleren van zakelijk succes door effectieve klantgerichte processen en communicatie.

ERP

ERP staat voor Enterprise Resource Planning.

Enterprise Resource Planning (ERP) is een geïntegreerde softwareoplossing die organisaties helpt bij het stroomlijnen en beheren van diverse bedrijfsprocessen en activiteiten.

Het systeem maakt gebruik van een centrale database om informatie van verschillende afdelingen, zoals boekhouding, voorraadbeheer, productie, human resources en meer, te consolideren en te delen.

Het doel van ERP is het verbeteren van de operationele efficiëntie, het verhogen van de transparantie en het ondersteunen van weloverwogen besluitvorming door real-time toegang tot gegevens.

Query

Een query kan eenvoudig omschreven worden als een zoekopdracht in een database.

Semantisch model

Een semantisch model vertaalt een technische database-implementatie in een voor eindgebruiker leesbaar en begrijpelijk model dat bruikbaar is zonder over (veel) technische kennis te beschikken.

Inhoudsopgave

Voorwaarden voor het begrippenkader	1
Algemene begrippen	2
Datageletterdheid	2
Datamaturiteit	2
Datageletterdheid vs Datamaturiteit	3
Aspecten van Data Governance	4
Data Governance	4
Data Management	4
Datalevenscyclus	5
Informatie-architectuur	5
Applicatie-architectuur	5
Data-architectuur	5
Business Informatie Model	6
Datamodellering en -ontwerp	6
Datamodel	6
Dataopslag en -beheer	7
Data-integratie en interoperabiliteit	7
Informatiebeheer	8
Informatieveiligheid	8
Documentbeheer	9
Contentmanagement	9
Referentiegegevens	9
Stamgegevens (Master Data)	9
Data opslaan en structureren	10
IOT Data	10
Business Intelligence	11
Datakwaliteitsmanagement	11
Technische begrippen	12
ETL / ELT	12
API	12

CRM.....	12
ERP	13
Query.....	13
Semantisch model	13
Inhoudsopgave.....	14