Auteur: Max Huiskes (2151960)

Klas: 32022INF1

Vak Beroepsproduct 5/6

School: Avans Stratenlaan te Den Bosch

Datum:

Versie: 1.0

Technisch ontwerp Data

Inhoudsopgave

[Inleiding 2](#_Toc149069574)

[Data architecture 3](#_Toc149069575)

[Beschrijving structuur 3](#_Toc149069576)

[Data dictionary 5](#_Toc149069577)

[Technisch ontwerp 6](#_Toc149069578)

[Bijlage A – Data pipeline 7](#_Toc149069579)

[Bijlage B – Technisch ontwerp Data 7](#_Toc149069580)

# Inleiding

In dit bestand wordt een brede erkenning getoond van het werk dat een student heeft voltooid om een robuuste gegevensarchitectuur te maken, compleet met een gedetailleerd gegevenswoordenboek en een technisch ontwerp voor gegevensopslag. De studie van de student richt zich op de kritieke stappen binnen de datapijplijn, zoals de gegevensinname en gegevensopslag, en laat zien hoe deze elementen zijn ontworpen en getest om waardevolle inzichten en analyses te bieden.

De data-architectuur is de basis van elk data gedreven project en de mogelijkheid om efficiënt gegevens te verzamelen, op te slaan en op te halen is van cruciaal belang voor het succes ervan. Door de aandacht voor detail en technische competentie van de student te combineren, worden de fundamenten voor effectieve gegevensverwerking in dit document gelegd.

In dit artikel wordt u door de belangrijkste aspecten van de datapijplijn geleid, waarbij de leerling begint met de identificatie van gegevens en vervolgens doorgaat naar de behandeling van gegevens. Het datawoordenboek dient als een cruciale stap in het begrijpen en documenteren van de betekenis en structuur van gegevens, wat bijdraagt aan betere communicatie en het juiste gebruik van informatie.

Deze collectie biedt een gedetailleerde kijk op de technische en conceptuele vooruitgang die de student heeft geboekt tijdens het ontwikkelen van de data-architectuur. Het toont de complexiteit en complexiteit waarmee een moderne data-engineer wordt geconfronteerd, evenals de innovatieve oplossingen die zijn ontwikkeld om deze complexiteiten aan te pakken. Samenvattend bevestigt deze verklaring de toewijding aan data gedreven uitmuntendheid en de voortdurende zoektocht naar efficiëntie en inzicht in de wereld van data science en analytics.

# Data architecture

A diagram of a network diagram

Description automatically generated

Zie bijlage A als voor het bestanden van de afbeelding.

## Beschrijving structuur

Dit gegevenssysteem beschrijft een end-to-end gegevensverwerkingsstroom die begint met gegevensinvoer en eindigt met gegevensvisualisatie en berekeningen. Hier is een uitsplitsing van de structuur en componenten van het systeem:

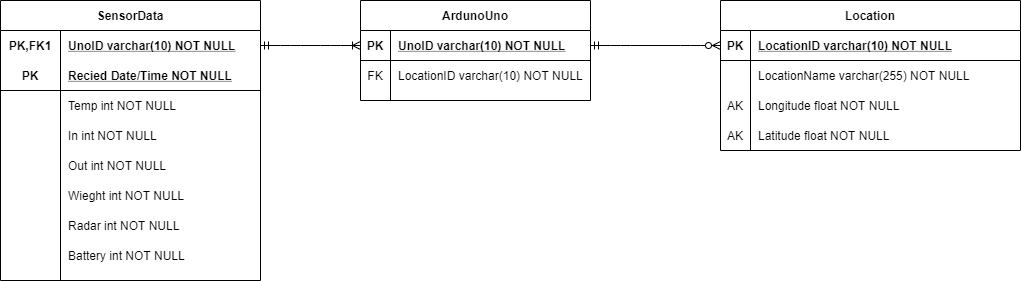
1. Consumption (Gegevensinname):
   1. Sensoren: Op een Arduino Uno met een LoRa-schild zijn vier sensoren geïnstalleerd: een gewichtssensor, een temperatuursensor, een IR-breakersensor en een radarsensor. Deze sensoren verzamelen gegevens met betrekking tot gewicht, temperatuur, beweging en radarinformatie.
   2. Gegevenstransformaties: De sensorgegevens worden verwerkt en opgeslagen in een specifiek payload-formaat door de Arduino. De temperatuurnormen worden hier toegepast (met 127 toegevoegd) en de gegevens worden gecombineerd in een byte-array.
   3. Verzending: De gegevens worden verzonden via het LoRa-protocol en de functie 'do\_send'. Als de zender niet wordt gebruikt (er vindt geen andere gegevensuitwisseling plaats), worden de gegevens in een payload verpakt en naar de LoRa Gateway verzonden.
2. LoRa Gateway/ The Things Network: De gegevens die door de Arduino worden verzonden, worden ontvangen door een LoRa Gateway en verzonden naar The Things Network (TTN).
3. Payload Formatter: Bij TTN worden de ontvangen gegevens opgeslagen in hexadecimaal formaat. Een payload-formatter wordt gebruikt om deze gegevens toegankelijker te maken. De functie 'decodeUplink' zet gegevens om in leesbare waarden, zoals temperatuur, gewicht, uitgangssignalen (Out), ingangssignals (In), radarinformatie en batterijniveaus.
4. FIWARE (Gemeente's-Hertogenbosch): Gedecodeerde gegevens worden verzonden naar een FIWARE-database die wordt onderhouden door de gemeente 's-Hertogenbosch. Het is mogelijk om binnen FIWARE een REST API te maken voor verbeterde gegevenstoegang en verwerking.
5. Node-Red: Node-Red wordt gebruikt om data aan te passen voordat deze wordt opgeslagen in een MongoDB database. Wiskundige formules en algoritmes kunnen hier aangepast worden, en het biedt de mogelijkheid om machine learning toe te passen voor voorspellingen en analyses. Bijvoorbeeld. Het systeem genereert voorspellingen gebaseerd op de verzamelde data. Zoals gewicht in het vogelhuisje in relatie tot het weer.
6. Dashboard: Het dashboard toont alle sensorgegevens en de resulterende acties. Dit stelt gebruikers in staat om gegevens en analyses op een uitgebreidere manier te onderzoeken en te begrijpen.

Deze datapijplijn begint met dataverzameling via sensoren en eindigt met de presentatie van data en inzichten via een dashboard. Het biedt de mogelijkheid om gegevens te verzamelen, converteren, opslaan, analyseren en visualiseren, waardoor waardevollere inzichten en betere besluitvorming mogelijk zijn, met name in contexten zoals vogeltrek en milieumonitoring.

# Data dictionary

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Field Name | Data Type | Data Format | Field Size | Description | Example |
| Recieved | DateTime | YYYY-mm-ddTHH:MM:ss | 19 | Datum en tijd van het ontvangen van de data in de database | 2023-10-24T17:34:35 |
| In | Integer |  | 3 | Hoeveel vogels erin zijn gevlogen in de afgelopen 5 minuten | 10 |
| Out | Integer |  | 3 | Hoeveel vogels eruit zijn gevlogen in de afgelopen 5 minuten | 11 |
| Temp | Integer |  | 3 | Wat de temperatuur in het nestje is | 5 °C |
| Weight | Integer |  | 8 | Het gewicht van de vogels in het vogelnestje | 150 g |
| Radar | Integer |  | 3 | De aantal vogels die in de afgelopen 5 minuten | 30 |
| UnoID | Varchar | xxxxxxxxxx | 10 | Dit is een uniek nummer voor een Arduino | 0987654321 |
| LocationID | Varchar | xxxxxxxxxx | 10 | Dit is een uniek id voor een locatie | 0123456789 |
| LocationName | Varchar |  | 255 | Dit is de naam van de locatie waar het vogelhuisje staad | Casino Park |
| Longitude | Float |  | 10 | De longitude van waar het vogelhuisje staat | 51.736252 |
| Latitude | Float |  | 10 | De latitude van waar het vogelhuisje staat | 5.329714 |

# Technisch ontwerp



Zie Bijlage B voor het bestand.

# Bijlage A – Data pipeline



# Bijlage B – Technisch ontwerp Data

