

# Project plan gestructureerde dataopslag vermissingen

10.2.e. - 10.2.e.

10 april 2020

---

# Inhoudsopgave

<b>1</b>	<b>Inleiding</b>	<b>3</b>
<b>2</b>	<b>Context</b>	<b>4</b>
10.2	Bedrijf . . . . .	4
2.2	Sarea . . . . .	4
<b>3</b>	<b>Opdracht</b>	<b>5</b>
3.1	Probleem . . . . .	5
3.2	Opdracht . . . . .	5
3.3	Oplossingsrichting . . . . .	6
<b>4</b>	<b>Project resultaten</b>	<b>7</b>
4.1	Project objectieven . . . . .	7
4.2	Scope . . . . .	7
<b>5</b>	<b>Stakeholders</b>	<b>8</b>
<b>6</b>	<b>Project activiteiten</b>	<b>9</b>
6.1	Project management . . . . .	9
6.2	Planning . . . . .	10
6.3	HBO competenties . . . . .	10
<b>7</b>	<b>Onderzoek</b>	<b>12</b>
7.1	Onderzoeksstrategie . . . . .	12
7.2	Dataverzamelmethode . . . . .	12
7.3	Analyse . . . . .	13
<b>8</b>	<b>Theoretisch kader</b>	<b>14</b>
8.1	Algoritme . . . . .	14
8.2	Database soorten . . . . .	14
8.3	PWA . . . . .	15
<b>9</b>	<b>Randvoorwaarden</b>	<b>16</b>
<b>10</b>	<b>Risico analyse</b>	<b>17</b>
	<b>Literatuur</b>	<b>18</b>

---

# 1 Inleiding

De politie is begonnen met het verder laten ontwikkelen van een initiatief genaamd Sarea. Sarea helpt burgers met zoektochten naar vermiste personen opzetten. Een van de gewenste functionaliteiten van Sarea is het weergeven van een gebied waar de persoon waarschijnlijk kan worden teruggevonden. Hier is een algoritme voor ontwikkeld door TNO. De verder ontwikkeling van dit algoritme word gehinderd door gebrek aan een databron om nodige informatie over individuele vermissingen te vinden. In dit project wordt een systeem opgezet om de nodig informatie in op te slaan. In dit projectplan zal verder de context van de opdracht, project resultaten, planning, stakeholders en risico's worden besproken.

---

## 2 Context

In dit hoofdstuk wordt de context gegeven voor de opdracht.

### 2.1 Bedrijf

De politie Noord-Nederland wil (technologische) ontwikkelingen benutten om het politiewerk effectiever te organiseren en nieuwe verbindingen te maken met de samenleving. Daarbij is in de inbreng van ideeën vanuit de politieorganisatie en samenwerking met externe partners en burgers in het veiligheidsdomein van groot belang.

Het innovatiehuis van de Politie Noord-Nederland maakt gebruik van de ruimte in de Digital Society Hub. Daar werken ze samen met studenten, onderzoekers en bedrijven.

### 2.2 Sarea

Sarea is een hulpmiddel voor het zoeken naar vermiste mensen. Sarea is een idee van oorsprong van een politie agent. Het idee is sinds verder uitgedacht en doorontwikkeld. Het bedrijf TNO is bezig geweest met het testen van een prototype en is bezig met het ontwikkelen van een algoritme om waarschijnlijke locaties te bepalen waar de vermiste kan worden teruggevonden. In september 2019 is er een project begonnen met de Hanzehogeschool Groningen waar negen CMD en zeventien ICT studenten werkten aan het realiseren van Sarea.

De huidige versie van Sarea heeft een app en website waarop zoekopdrachten kunnen worden opgezet en beheerd. Gebruikers voeren bij het aanmaken een aantal eigenschappen van de vermiste in. In deze versie wordt het zoekgebied bepaald door het laatst geziene punt te nemen en met dit als center een cirkel te trekken met een variabele straal. De straal van de cirkel is enkel afhankelijk van de leeftijd van de vermiste. Tijdens het zoeken naar een vermiste wordt de locatie van de zoekers bijgehouden door de Sarea app. De locatie van iedereen is te zien door de deelnemers wat helpt inzicht te krijgen in waar al wel en niet gezocht is.

---

## 3 Opdracht

In dit hoofdstuk wordt het de opdracht besproken. Als eerst wordt het probleem uitgelicht. Hierop volgend wordt de te ontwikkelen oplossing beschreven.

### 3.1 Probleem

TNO heeft een algoritme ontwikkelt voor het bepalen van een gebied waar een vermiste persoon waarschijnlijk kan worden teruggevonden. Dit algoritme is vooral gebaseerd op expert kennis, welke gebaseerd is op vaste regels. Voor het verder ontwikkelen van het zoekalgoritme is door de politie de wens geuit om het algoritme meer datagedreven te maken, omdat datagedreven voorspellingen vaak specifiekere voorspellingen geven. Voor het maken van een dergelijke applicatie is een dataset nodig met de tooling om deze dataset gemakkelijk te maken en onderhouden. Op dit moment is er geen bruikbare gestructureerde dataset beschikbaar van Nederlandse vermissingen. De beschikbare informatie is vaak niet gestructureerd of onvolledig. In een onderzoek over langdurig vermiste Nederlanders was het volgende te vinden over werken met de bestaande databases. “De registraties worden niet bewaard voor analysedoeleinden en de systemen zijn daar dan ook niet op ingericht” (van Leiden & Hardeman, z. j.).

### 3.2 Opdracht

Om het eerder omschreven probleem op te lossen dient een gestructureerde dataset te worden aangelegd. In dit project worden de gereedschappen gemaakt voor het invoeren en opslaan van gegevens rond vermissingen. De invoer van deze gegevens dient verricht te kunnen worden door een persoon zonder technische achtergrond. 10.2.g



---

### 3.3 Oplossingsrichting

10.2.g



---

## 4 Project resultaten

In dit hoofdstuk worden de uitkomsten van het project besproken. Ook zal worden aangegeven wat buiten de scope van het project valt.

### 4.1 Project objectieven

De te behalen beroepsproducten zijn een gerealiseerd systeem voor het invoeren, opslaan en ophalen van gegevens over vermissingen, een verantwoordingsverslag en een advies over het verder verzamelen van gegevens. Ter ondersteuning van deze producten zal een ontwerp worden gemaakt van het systeem. Dit ontwerp wordt weer ondersteund door een opgestelde eisen en een vergelijkingsonderzoek verslag naar verschillende databasen.

10.2 g



---

## 5 Stakeholders

Tijdens dit project zijn er een aantal stakeholders. Deze stakeholders staan hieronder weergegeven met naam, email, rol binnen dit project en bedrijf wanneer van toepassing.

Naam 10.2.e.  
Rol Afstudeerstagiair  
Email 10.2.e. @st.hanze.nl

Naam 10.2.e.  
Rol Bedrijfsbegeleider/ Opdrachtgever  
Bedrijf Politie  
Email 10.2.e. @gmail.com

Naam 10.2.e.  
Rol Stagebegeleider/ Tafelvoorzitter  
Bedrijf Hanzehogeschool Groningen  
Email 10.2.e. @pl.hanze.nl

Naam 10.2.e.  
Rol Tafelvoorzitter  
Bedrijf Hanzehogeschool Groningen  
Email 10.2.e. @pl.hanze.nl





Naam 10.2.e.  
Bedrijf TNO  
Email 10.2.e. @tno.nl



---

## 6 Project activiteiten

### 6.1 Project management

Voor project management wordt gebruik gemaakt van <sup>10.2.g</sup> . <sup>10.2.g</sup>  geeft de flexibiliteit om snel veranderingen te maken bij veranderingen van eisen en prioriteiten. De toevoeging van sprints geeft een serie aan vaste feedbackmomenten voor de nodige stakeholders om sturing te geven aan de eisen en prioriteiten. Verder geeft <sup>10.2.g</sup>  een overzicht in de huidige progressie door middel van het <sup>10.2.g</sup>  bord waar alle taken op staan met de backlog.

---

## 6.2 Planning

10.2.g



---

### 6.3 HBO competenties

De HBO competenties zullen in de volgend momenten worden ingevuld:

- Analyse Voor het voldoen van de analyze competentie wordt er gekeken naar bestaande methoden van invoer en opslag van vermissingszaken.
- Onderzoek: Voor het ontwerp zal een onderzoek worden gedaan naar verschillende database mogelijkheden.
- Ontwerpen: Voor het maken voor van het product wordt een ontwerp gemaakt met daarin de technologie en architectuur keuzes.
- Realiseren: Het bouwen van het eerder ontworpen product.
- Adviseren: Aan het einde een schriftelijk advies geven over de verdere ontwikkeling en gebruik van het gerealiseerde product.

---

## 7 Onderzoek

Voor het onderzoek wordt gekeken wat voor soort database het meest geschikt is voor het opslaan van vermissingsdata. Voor dit onderzoek wordt er gekeken naar de verschillende types databasen die bestaan en dus niet naar verschillende implementaties van hetzelfde concept (bijvoorbeeld relationele databasen). Als deelvragen zijn er de volgende:

- Wat zijn er voor verschillende typen databasen die gebruikt kunnen worden?
- Welke databasen kunnen de nodige data opslaan?
- Welke databasen kunnen data zoeken op de nodige criteria?
- Zijn er performance implicaties bij de databasen wanneer er meer data word opgeslagen?

### 7.1 Onderzoeksstrategie

Voor onderzoeksstrategie is er gekozen voor een experiment. In dit experiment is de database de variabele die veranderd gaat worden om de gevolgen op de datastructuur te kunnen observeren. De afhankelijke variabelen zijn de datastructuur en de prestaties en limieten bij het gebruik van de database.

10.2.g



### 7.3 Analyse

Voor data-analyse wordt eerst vergeleken op welke zoekopdrachten ze wel of niet konden uitvoeren. Als tweede worden de databases op de verschillende executie snelheden vergeleken en door middel van een t-test bepaald of het resultaat significant is of niet.

---

## 8 Theoretisch kader

In dit hoofdstuk wordt al beschikbare informatie over het onderwerp besproken.

### 8.1 Algoritme

Er is al een versie ontwikkeld voor het vinden van vermiste personen op basis van data. Dit algoritme maakt gebruik van een Bayesiaans Netwerk om voorspellingen te doen. Omdat de data niet beschikbaar was is er gebruik gemaakt van gegenereerde gegevens (Sternheim, 2019).

### 8.2 Database soorten

Er zijn verschillende methoden om data langdurig op te slaan. Deze verschillende soorten databases worden verdeeld over twee categorieën SQL en NoSQL. Als eerst wordt de SQL type besproken.

SQL databases zijn relationele databases. Een relationele database is een type database dat data punten opslaat die gerelateerd zijn tot een ander. Relationele databases zijn gebaseerd op het relationeel model, een intuïtieve manier voor het representeren van data in tabellen. In een relationele database is elke rij in de tabel een vermelding met een uniek id dat de sleutel wordt genoemd. De kolommen van de tabel houden de attributen van de data, en elke vermelding heeft meestal een waarde voor elk attribuut, wat het maken van de relatie tussen datapunten makkelijk maakt. (Oracle, 2020)

Dan de andere categorie NoSQL databases. De term NoSQL betekend dat de database geen relationele database is maar er zijn meerdere soorten die hieronder vallen. NoSQL databases zijn gebouwd voor specifieke datamodellen en zijn flexibel om de vorm van de data die wordt opgeslagen aan te passen. In de NoSQL categorie kunnen de we volgende typesd weer omschrijven: key-value, document, graph, in-memory en search (Amazon, 2020f).

Als eerst een key-value database dat op basis van een vooraf gegeven waarde als sleutel waardes kan opslaan en opvragen. Deze waardes kan simpele data zijn maar ook complexe object composities. Een aantal voorbeelden hiervan zijn Amazon DynamoDB en Apache Cassandra (Amazon, 2020c). Dan is er de document database. Hier is data gerepresenteerd in een Object of JSON-like document omdat het een efficiënt en intuïtief data model is voor ontwikkelaars. Voorbeelden van document databases zijn MongoDB en Couchbase (Amazon, 2020a). Als derde is er de graph database. Dit soort maakt het makkelijk voor het maken van applicaties met een dataset met veel connecties. Voorbeelden van een graph database zijn Amazon Neptune en Neo4j

---

(Amazon, 2020b). Als vierde is er de in-memory database. Een in-memory database slaat alles op in het ram en is bedoelt voor zo laag mogelijke reactie tijd. Het brengt het risico dat wanneer de server uitvalt de data verloren gaat. Een voorbeeld hiervan is Aerospike (Amazon, 2020d). Als laatste is er de search database. Een search database is gericht op het zoeken van content. De database is geoptimaliseerd voor het het werken met data dat lang, gestructureerd of ongestructureerd is. Voorbeelden van een search database zijn Amazon Elasticsearch Service en Splunk (Amazon, 2020e).

### **8.3 PWA**

Een progressive web app(PWA) is een web app. Een PWA maakt gebruik van de webbrowser technologieën met een verdere set aan functionaliteiten. Een van de extra functionaliteiten van een PWA is de mogelijkheid om offline te werken. Verder kunnen kan een PWA in Android geïnstalleerd worden waardoor die verschijnt als een normale app (Mozilla, 2020). Een voorbeeld van de PWA is de Twitter website.

---

## 9 Randvoorwaarden

10.1.c





---

## 10 Risico analyse

Tijdens het uitvoeren van het project zijn er de volgende risico's.

Risico	Beheersing
Een of meer stakeholders zijn niet beschikbaar om te antwoorden op vragen	Zo weinig mogelijk afhankelijk zijn van reacties van stakeholders voor het verder kunnen werken
Veranderingen in de eisen en/of prioriteiten	Duidelijk zijn over de eisen zoals ze bekend zijn en verifiëren bij kans op onduidelijkheid om grote veranderingen te vermijden

---

## Literatuur

- Amazon. (2020a). *What is a document database?* Op 2020-02-21 verkregen van <https://aws.amazon.com/nosql/document/>
- Amazon. (2020b). *What is a graph database?* Op 2020-02-21 verkregen van <https://aws.amazon.com/nosql/graph/>
- Amazon. (2020c). *What is a key-value database?* Op 2020-02-21 verkregen van <https://aws.amazon.com/nosql/key-value/>
- Amazon. (2020d). *What is an in-memory database?* Op 2020-02-24 verkregen van <https://aws.amazon.com/nosql/in-memory/>
- Amazon. (2020e). *What is a search-engine database?* Op 2020-02-24 verkregen van <https://aws.amazon.com/nosql/search/>
- Amazon. (2020f). *What is nosql?* Op 2020-02-24 verkregen van <https://aws.amazon.com/nosql/>
- Boicea, A., Radulescu, F. & Agapin, L. I. (2012). MongoDB vs oracle-database comparison. In *2012 third international conference on emerging intelligent data and web technologies* (pp. 330–335).
- Grüner, S., Pfrommer, J. & Palm, F. (2016). Restful industrial communication with opc ua. *IEEE Transactions on Industrial Informatics*, 12(5), 1832–1841.
- MongoDB. (2020). *What is a document database?* Op 2020-02-21 verkregen van <https://www.mongodb.com/document-databases>
- Mozilla. (2020). *Progressive web apps (pwAs)*. Op 2020-02-24 verkregen van [https://developer.mozilla.org/en-US/docs/Web/Progressive\\_web\\_apps/Introduction](https://developer.mozilla.org/en-US/docs/Web/Progressive_web_apps/Introduction)
- Oracle. (2020). *What is a relational database?* Op 2020-02-21 verkregen van <https://www.oracle.com/database/what-is-a-relational-database/>
- Sternheim, A. M. (2019). Vindlocatie van vermiste personen voorspellen met een bayesiaans netwerk.
- van Leiden, I. & Hardeman, M. (z. j.). Vermisten op de kaart.