

BTO 2016.027 | April 2016

BTO rapport

Haalbaarheidsstudie
stoffendatabase voor de
Nederlandse
drinkwatersector

BTO

Haalbaarheidsstudie stoffendatabase voor de Nederlandse drinkwatersector

BTO 2016.027 | April 2016

Opdrachtnummer

400914

Projectmanager

Kirsten Baken

Opdrachtgever

LMO - Laboratorium Managers Overleg
BTO - Thematisch onderzoek - Nieuwe stoffen

Kwaliteitsborger(s)

Annemarie van Wezel

Auteur(s)

Rosa Sjerps, Erik Emke en Mario Maessen

Verzonden aan

Dit rapport is verspreid onder BTO-participanten en is
openbaar.

Jaar van publicatie
2015

Meer informatie
Rosa Sjerps, MSc
T 030 60 69 704
E rosa.sjerps@kwrwater.nl

Keywords
Database, nieuwe stoffen,
signaleren, in-silco
onderzoeksinstrument

PO Box 1072
3430 BB Nieuwegein
The Netherlands

T +31 (0)30 60 69 511
F +31 (0)30 60 61 165
E info@kwrwater.nl
I www.kwrwater.nl

KWR Watercycle
Research
Institute

BTO 2016.027 | April 2016 © KWR

Alle rechten voorbehouden.

Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

BTO Managementsamenvatting

Uniforme stoffendatabase met bekende en onbekende stoffen voor de Nederlandse drinkwatersector is haalbaar

Auteur(s): Rosa Sjerps MSc, ing. Erik Emke en drs. Mario Maessen

Onderzoekers en stakeholders uit de sector stelden samen een breed gedragen en goed afgestemd pakket van eisen op voor de opzet van een stoffendatabase, waarin gegevens over bekende en onbekende stoffen worden gedeeld tussen de drinkwaterlaboratoria. De database zal bestaan uit twee functionaliteiten: stoffeigenschappen van bekende stoffen (primaire functionaliteit) en screeninggegevens voor het uitwisselen van gegevens over onbekende stoffen (secundaire functionaliteit). De bijbehorende kosten voor een dergelijke opzet zijn geschat op 50 k€. De definitieve raming kan pas worden opgesteld na een systeemanalyse. De meerwaarde van deze database is tweeledig. Enerzijds voegt deze database aan de bekende stoffen veel gevraagde maar moeilijk vindbare stoffeigenschappen toe. Anderzijds beoogt deze database een actueel overzicht te geven van nieuwe en onbekende stoffen voor de bedrijfstak en het onderzoek dat naar dergelijke stoffen is uitgevoerd (SPO en BTO).

Belang: beschikbare stofgegevens uniform opslaan en doorzoekbaar maken voor de drinkwaterwereld

Binnen diverse projecten en instituten is een schat aan gegevens over chemische stoffen in de watercyclus beschikbaar. Het gaat om stoffeigenschappen, analysetechnieken, emissies, lotgevallen in het milieu zoals sorptie en degradatie, concentraties in diverse milieumatrices, effecten op de humane gezondheid en het milieu, vorming van bijproducten en mate van verwijdering bij waterbehandeling. Bestaande gegevens zijn vastgelegd in de literatuur, databases of per persoon of organisatie. Gebrek aan uniformiteit en betrouwbaarheid maken het lastig om bestaande gegevens in retrospectief efficiënt te doorzoeken en te hergebruiken. Er is in de drinkwaterwereld daarom behoefte aan een centraal toegankelijke, web-gebaseerde en gevalideerde database die zorgt voor een uniforme manier om data op te slaan, te doorzoeken en te delen met vakgenoten. Dit project had tot doel te komen tot een breed gedragen en goed afgestemd pakket van eisen voor de opzet van een dergelijke stoffendatabase.

Aanpak: enquête expertteam, workshops, samenvoegen eisen en opstellen kostenplaatje

Het pakket van eisen voor de stoffendatabase is in stappen vastgesteld. Eerst is een enquête uitgevoerd onder een team van experts (onderzoekers, stakeholders en eindgebruikers) om hun voorkeuren te identificeren. Vervolgens zijn twee workshops gehouden waarin de wensen en eisen nader zijn gespecificeerd. Op basis van deze resultaten is het definitieve pakket van eisen vastgesteld en zijn globaal de haalbaarheid en het kostenplaatje uitgewerkt. Bij het uitvragen van het pakket van eisen ging het onder meer over de keuzes ten aanzien van het format, de gewenste functionaliteit van de database en de vereiste informatie inclusief validatievereisten. Omdat het een sector-brede database betreft worden deze specificaties goed afgestemd met de verschillende stakeholders uit de sector.

Resultaten: uniforme databasestructuur met gegevens over bekende en onbekende stoffen

De opzet van de database is onderverdeeld in vijf kennisdomeinen voor stofgegevens: stoffeigenschappen, analytische chemie, monitoringsgegevens (welke laboratorium meet

wat), toxicologie en zuiveringsefficiëntie. Elk van deze kennisdomeinen bevat relevante stofparameters voor opname in de database.

Bekende én onbekende stoffen

Volgens het pakket van eisen moet de databasestructuur bestaan uit twee delen, met twee functionaliteiten: stofeigenschappen van bekende stoffen en screeninggegevens voor het uitwisselen van gegevens over onbekende stoffen. De parameters die gewenst of vereist zijn voor opname in de database zijn uitgezet in een format. De meerwaarde van deze database is tweeledig. Enerzijds voegt deze database aan de bekende stoffen veel gevraagde maar moeilijk vindbare stofeigenschappen toe. Anderzijds beoogt deze database een actueel overzicht te geven van nieuwe- en onbekende stoffen.

Kosten

De kosten voor het opzetten van een dergelijke database zijn geschat op 50k€. De definitieve raming kan pas worden opgesteld na een systeemanalyse. De systeemanalyse kan als eerste fase worden opgezet (10 k€), waarna een go or no go beslissing gemaakt kan worden voor de op te zetten database.

Bestaande databases gebruiken

Het onderzoek heeft zich ook gericht op de mogelijkheid dat ervoor wordt gekozen geen databasestructuur op te zetten, maar gebruik te maken van bestaande databases. In dat geval zijn er talrijke opties, in het rapport is een niet-uitputtende opsomming opgenomen. Stofeigenschappen zijn direct of indirect op te zoeken met platforms zoals Chemspider. Monitoringsgegevens van stoffen, die niet onder

een wettelijke verplichte verplichting vallen, kunnen mogelijk worden toegevoegd aan REWAB.

Verwijderingsefficiëntie is momenteel onderdeel van de watersharetool AbatES. Screeninggegevens worden gedeeld met behulp van platform zoals EMPODAT. Verder zijn accurate MSMS spectra beschikbaar via Massbank en zijn er initiatieven om deze ook voor onbekende stoffen op te slaan. Het Europese NORMAN netwerk stelt via STOFF-Ident analytische informatie beschikbaar van voor water-relevante stoffen om hun identificatie te ondersteunen.

Implementatie: opzet database met go/no go of aanvulling bestaande databases

De (drink)watersector moet zich beraden over de keuze om wel of niet een uniforme stoffendatabase op te zetten voor de sector op te zetten, die zorgt voor één structuur met opslag van noodzakelijke gegevens en die actueel en bereikbaar is voor de watersector. Een dergelijke database is bruikbaar als opslag van gegevens, het delen van informatie en het signaleren en prioriteren van nieuwe bedreigingen of als tool voor modelleringen. De eindgebruikers zijn medewerkers van drinkwaterlaboratoria, drinkwaterbedrijven en KWR. Het alternatief kan zijn om informatie te (blijven) verzamelen uit een reeks niet uniforme databases. Ontwikkelen van een uniforme database wordt aanbevolen, waarbij kan worden gestart met een systeemanalyse om daarna een go/no go voor het opzetten van een database te geven.

Rapport

Dit onderzoek is beschreven in rapport *Haalbaarheidsstudie stoffendatabase voor de Nederlandse drinkwatersector* (BTO 2016.027).

Jaar van publicatie
2015

Meer informatie
Rosa Sjerps, MSc
T 030 60 69 704
E rosa.sjerps@kwrwater.nl

Keywords

Database, nieuwe stoffen, signaleren, in-silco onderzoeksinstrument

PO Box 1072
3430 BB Nieuwegein
The Netherlands

T +31 (0)30 60 69 511
F +31 (0)30 60 61 165
E info@kwrwater.nl
I www.kwrwater.nl

KWR *Watercycle
Research
Institute*

BTO 2016.027 | April 2016 © KWR

Alle rechten voorbehouden.

Niets uit deze uitgave mag worden veelevoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Samenvatting

Onderzoekers en stakeholders uit de sector stelden samen een breed gedragen en goed afgestemd pakket van eisen op voor de opzet van een stoffendatabase, waarin gegevens over bekende en onbekende stoffen worden gedeeld tussen de drinkwaterlaboratoria. Elk Nederlands drinkwaterlaboratorium is vertegenwoordigd binnen het expertteam. De database zal bestaan uit twee delen met twee functionaliteiten:

- 1) Stofeigenschappen van bekende stoffen
- 2) Screeninggegevens voor het uitwisselen van gegevens over onbekende stoffen

De gewenste op te nemen parameters vallen onder zeven hoofdstukken:

- 1) Unieke identiteit
- 2) Stofeigenschappen
- 3) Stofeigenschappen uit BTO onderzoek
- 4) Makkelijk vindbare stofeigenschappen
- 5) Monitoring (welk laboratorium meet welke stof)
- 6) Analytische gegevens onbekenden
- 7) Monitoring onbekenden

Alle parameters in deze hoofdstukken zijn gewenst of vereist voor opname. De makkelijk vindbare stofeigenschappen zijn met weinig moeite op te zoeken in bestaande databases en zijn geen prioriteit voor het opnemen in de database. Deze eigenschappen zullen in een later stadium aan de database worden toegevoegd.

De belangrijkste aandachtspunten met betrekking tot de database liggen in de beveiliging van de vertrouwelijke data, en de vulling van de dataset. De gebruikers willen niet heel veel tijd kwijt zijn aan het vullen van de database. Dat moet mogelijk zijn door de import en export van data zo eenvoudig mogelijk te maken. Dit betekent dat er bijvoorbeeld delen uit de LIMS systemen via een vast format moeten kunnen worden geïmporteerd. De vulling van de database vindt plaats door de gebruikers. KWR zal de database bij oplevering vullen met stofgegevens die KWR de laatste drie jaar heeft verzameld. Van de deelnemende bedrijven wordt hetzelfde verwacht.

De meerwaarde van deze database is tweeledig. Enerzijds voegt deze database aan de bekende stoffen veel gevraagde maar moeilijk vindbare stofeigenschappen toe. Anderzijds beoogt deze database een actueel overzicht te geven van nieuwe- en onbekende stoffen.

De kosten voor de opzet van een dergelijke database zijn geschat op 50 k€. De definitieve raming kan pas worden opgesteld na een systeemanalyse. De systeemanalyse kan als eerste fase worden opgezet (10 k€), waarna een go or no go beslissing gemaakt kan worden voor de op te zetten database.

Onderdeel	Gepaarde kosten
Opzet structuur database	50 k€ (+/-50%)
Invulling database	Door gebruikers in te vullen
Beheer database	5 k€ per laboratorium per jaar

Het beheer van de database door KWR zal ongeveer 5 k€ per laboratorium per jaar gaan kosten, omdat alle gegevens toegankelijk zijn voor alle deelnemers is er uitgegaan van een standaard bedrag per lab. De laboratoria zijn zelf verantwoordelijk voor het bijhouden en up-to-date houden van de ingevulde gegevens. KWR voert een jaarlijkse kwaliteitsborging uit op de ingevoerde data.

Als men ervoor kiest om geen database structuur op te zetten, maar gebruik te maken van bestaande databases, zijn er talrijke opties waarvan in het onderliggende rapport een niet uitputtende opsomming is gegeven. Stofeigenschappen zijn direct of indirect op te zoeken met platforms zoals Chemspider. Binnen de REACH-wetgeving is de database IUCLID opgezet voor het opslaan, indienen en uitwisselen van informatie over stoffen volgens het OECD format. Verwijderingsefficiëntie in drinkwaterzuiveringen is momenteel onderdeel van de watersharetool AbatES. Monitoringsgegevens van stoffen, die niet onder een wettelijke verplichte verplichting vallen, kunnen worden toegevoegd aan REWAB. Screeninggegevens worden momenteel gedeeld met behulp van platforms zoals EMPODAT (waarvan de matrix water een onderdeel is). Verder zijn accurate MSMS spectra beschikbaar via Massbank en zijn er initiatieven om deze ook voor onbekenden op te slaan. Analytische informatie van water-relevante stoffen om identificatie te ondersteunen zijn binnen het Europese NORMAN netwerk beschikbaar via STOFF-Ident. Bij deze keuze zullen de resultaten van de bedrijfstak m.b.t. screeningsonderzoek minder goed gedeeld worden.

Inhoud

Samenvatting	2	
Inhoud	4	
1	Introductie	5
1.1	Introductie	5
1.2	Leeswijzer	6
2	Aanpak	7
2.1	Globale aanpak	7
2.2	Het team	7
2.3	Enquête	7
2.4	Workshop I	7
2.5	Workshop II	8
2.6	Pakket van eisen en wensen	8
2.7	De uitwerking, haalbaarheid en kostenplaatje	8
3	Pakket van eisen en wensen	9
3.1	Uitkomst enquête	9
3.2	Uitkomst workshop	9
3.3	Uitkomst tweede workshop	12
3.4	Samengevat	13
4	Functionele eisen en opzet database	16
4.1	De functionele eisen	16
4.2	De haalbaarheid en globale kostenraming	19
5	Referenties	21
Bijlage I		22
•	Samenstelling expertteam	22
Bijlage II		23
•	Enquête & respons	Error! Bookmark not defined.
•	Enquête & respons	Error! Bookmark not defined.
Bijlage II		22
•	Template parameters gewenst voor opname in database en kostenraming voor inhoud database	22

1 Introductie

1.1 Introductie

Binnen diverse (Europese) projecten, onderzoeksinstituten en waterbedrijven is een schat aan gegevens over chemische stoffen in de watercyclus beschikbaar. Het gaat om stofeigenschappen, analysetechnieken, emissies, lotgevallen in het milieu zoals sorptie en degradatie, concentraties in diverse milieumatrices, effecten op de humane gezondheid en het milieu, vorming van bijproducten en mate van verwijdering bij waterbehandeling. Bestaande gegevens zijn vastgelegd in de literatuur en in bestaande databases in nationaal (o.a. REWAB, Bestrijdingsmiddelenatlas) en internationaal verband (o.a. Chemspider, TOXNET). Ook zijn de data vaak vastgelegd in niet gedeelde datadragers specifiek per persoon of organisatie. Omdat de bronnen van de gegevens variabel zijn, is ook de kwaliteit van deze gegevens wisselend. Gebrek aan uniformiteit en betrouwbaarheid maken het lastig om bestaande gegevens in retrospectief efficiënt te doorzoeken en te hergebruiken.

Een binnen de drinkwaterwereld centraal toegankelijke, web-gebaseerde en gevalideerde database lijkt een goede optie om beschikbare gegevens op een uniforme manier op te slaan en goed doorzoekbaar te maken. Toekomstige informatie over nieuwe stoffen kan gemakkelijk aan de database toegevoegd worden. Indien het vóórkomen van nieuwe en onbekende stoffen worden opgenomen, zijn dergelijke databases ook bruikbaar als tool voor het signaleren en prioriteren van nieuwe bedreigingen of als tool voor modelleringen. Voor de bouw van een dergelijke database is niet gekozen voor de uitbreiding van een bestaande database structuren. Om bestaande database structuren aan te passen aan de nieuwe wensen zouden deze geheel moeten worden omgebouwd. Het aanpassen van een bestaande database structuren veel complexer wat betreft organisatie en inhoud. Dit zou dermate kostbaar worden dat dit niet binnen het beschikbare budget (en tijdspad) zou passen, buiten de vraag of de huidige eigenaren hieraan zouden willen meewerken.

Vooraf aan het realiseren van de stoffendatabase voor de waterketen moeten keuzes gemaakt worden ten aanzien van het format, de gewenste functionaliteit van de database en de vereiste informatie inclusief validatievereisten. Omdat het een sector-brede database betreft moeten deze specificaties goed worden afgestemd met de verschillende stakeholders uit de sector. Dit project beoogt om gezamenlijk met de sector tot een breed gedragen en goed afgestemd pakket van eisen te komen voor een dergelijke database. Het pakket van eisen en wensen vormt zo de basis voor een mogelijk op te zetten stoffendatabase voor de drinkwatersector. Vervolgens wordt een inschatting gemaakt van de haalbaarheid en de kosten gemoeid met de opzet van een mogelijke database.

Deze rapportage geeft een wensbeeld van de op te stellen database. Op basis van de globale kostenraming zal nog een definitieve keuze worden gemaakt welke functionele eisen in de database worden opgenomen. Zowel vanuit het Labmanagersoverleg (LMO) als het bedrijfstakonderzoek van de drinkwaterbedrijven (BTO) is de wens geuit om te onderzoeken of het mogelijk is om een database met web-based interface voor stoffen op te zetten, waarbij alle relevante gegevens voor de bedrijfstak van deze stoffen makkelijk voor alle laboratoria en waterbedrijven toegankelijk zijn en er een dynamische factsheet kan worden gecreëerd. De database zal met name bedoeld zijn om de schat aan gegevens die door de bedrijfstak zelf zijn verzameld op een makkelijk manier beschikbaar maken.

1.2 Leeswijzer

In Hoofdstuk twee wordt de aanpak beschreven om tot een pakket van eisen en wensen te komen voor het opzetten van een stoffendatabase. In Hoofdstuk drie zijn de resultaten beschreven met betrekking tot de inhoud en de toepassing binnen het pakket van eisen en wensen voor de op te zetten database. In Hoofdstuk vier wordt de technische uitwerking van de beoogde opzet beschreven en worden de gepaarde kosten berekend.

2 Aanpak

2.1 Globale aanpak

Het opstellen van het pakket van eisen is uitgevoerd in een aantal stappen. Als eerste stap is een enquête uitgevoerd bij een team van experts dat bestaat uit onderzoekers en stakeholders (zie paragraaf 2.2). In deze enquête is onderzocht welke voorkeuren bestaan in het expert team (zie paragraaf 2.3). Vervolgens zijn twee workshops gehouden waarin de wensen en eisen nader zijn gespecificeerd (zie paragrafen 2.4 en 2.5). Op basis van deze resultaten is het definitief pakket van eisen vastgesteld (paragraaf 2.6) en is globaal de haalbaarheid en het kostenplaatje uitgewerkt (paragraaf 2.7).

De opzet van de database hebben we onderverdeeld in vijf kennisdomeinen voor stofgegevens: stofeigenschappen, analytische chemie, monitoringsgegevens, toxicologie en zuiveringsefficiëntie. Elk van deze kennisdomeinen bevatte relevante stofparameters voor opname in de database.

2.2 Het team

In het expertteam zijn alle Nederlandse drinkwaterlaboratoria vertegenwoordigd, met 5 deskundigen van de drinkwaterlaboratoria en 5 onderzoekers van KWR. De samenstelling van het expertteam is weergegeven in Bijlage I.

2.3 Enquête

Ten eerste is een enquête opgesteld door KWR en verzonden aan de leden van het expertteam. De enquête betrof de inhoud en functie van de database. De enquête bestond uit:

- Algemene vragen
- Relevante parameters voor vijf onderscheiden kennisdomeinen: stofeigenschappen, analytische chemie, monitoringsgegevens, toxicologie en zuiveringsefficiëntie
- In gebruik zijnde bestaande databases

De enquête is toegevoegd in Bijlage II. De antwoorden zijn geanalyseerd aan de hand van vier categorieën met bijbehorende probleemstellingen:

- Welke partijen zijn verantwoordelijk voor het opzetten en onderhouden van de database?
- Welke gegevens moeten er in de database worden opgeslagen?
- Wat is de gewenste output van de database?
- Wat zijn de belangrijkste toepassingen van de database?

De respons op de enquête is geëvalueerd en diende mede als input voor workshop I.

2.4 Workshop I

De eerste workshop werd gehouden op 23 juni 2015. Naast de bespreking van de enquête had de workshop als doel gezamenlijk tot een pakket van eisen en wensen te komen. Deelnemers brainstormden over de opbouw van de stoffendatabase, de gewenste output en de taken van de stakeholders, aan de hand van probleemstellingen ingedeeld in vier categorieën omtrent de omzet van de database: verantwoordelijkheden en rollen van de

beheerders en gebruikers, input en gegevensopslag in de database, output en toepassingen van de database. Aan de hand van de stellingen (zie tekst box 1) discussieerden de deelnemers in groepjes van twee over het onderwerp. De resultaten werden gepresenteerd aan en bediscussieerd met de groep. Zo probeerden we zo veel mogelijk visies mee te nemen in het pakket van eisen en wensen voor de opzet van een stoffendatabase.

Tekst box 1: Vier probleemstellingen ter discussie tijdens de brainstormsessie:

1. Rollen

Welke partijen (denk aan laboratoria, KWR, etc.) spelen welke rol (bijv. aanvullen van gegevens, kwaliteitsborging/validatie, etc.) m.b.t. de stoffendatabase?

2. Input

Welke informatie of welke combinaties van informatie zijn nodig om de database een waardevolle aanvulling te laten zijn op bestaande databases (zoals REWAB, OECD)?

3. Output

Wie zijn de eindgebruikers en op welke manier zouden zij informatie willen ontsluiten uit de stoffendatabase?

4. Toepassingen

Waar ga je de opgevraagde informatie uit de database voor gebruiken?

2.5 Workshop II

De tweede workshop werd gehouden op 24 november 2015. Het doel van deze workshop was vaststelling van het definitieve pakket van eisen en wensen en de definitieve opzet voor de database. Het resultaat van de workshop is een template met de gewenste parameters in de database.

2.6 Pakket van eisen en wensen

De resultaten van de enquête en de resultaten van de workshops leiden tot het opstellen van een pakket van eisen en wensen voor de stoffendatabase. Dit pakket van eisen en wensen vormt de basis voor een opzet van een stoffendatabase. Uit de inventarisatie kan tevens als resultaat komen dat het niet gewenst is een stoffendatabase op te zetten.

2.7 De uitwerking, haalbaarheid en kostenplaatje

De functionele eisen in ogenschouw nemend, is een voorstel gedaan voor de technische uitwerking van de stoffendatabase. Vervolgens is samen met een KWR medewerker die ervaren is in de ICT-branche, op basis van expert judgement de uitvoerbaarheid getoetst en is een zeer globale kostenraming gemaakt.

3 Pakket van eisen en wensen

In dit hoofdstuk worden de resultaten gepresenteerd van de enquête (3.1) en de workshops (3.2 en 3.3). Vervolgens zijn de resultaten samengevat in een pakket van eisen en wensen (3.4). Dit pakket van eisen en wensen gaat over de invulling en de gewenste toepassingen van de database. Vervolgens wordt in Hoofdstuk 4 de opzet van de database verder uitgewerkt, de haalbaarheid getoetst en de gepaarde kosten berekend.

3.1 Uitkomst enquête

Alle vijf stakeholders retourneerden een ingevulde enquête. De respons is te vinden in Bijlage II. In Tabel 3-1 zijn de resultaten van de enquête samengevat naar de vier gedefinieerde categorieën omtrent de database: rollen, input, output en toepassingen.

Tabel 3-1 Resultaten van de enquête samengevat. De respons is te vinden in Bijlage II.

Categorie	
Rollen	Het merendeel van de respondenten vindt het gewenst dat zowel KWR als de laboratoria gegevens kunnen in- en aanvullen. Tevens vindt het merendeel dat KWR de validatie procedure voor de ingevoerde content jaarlijks moet uitvoeren ter controle van foutief of verkeerd ingevulde gegevens en stoffeigenschappen.
Input	Er is verdeeldheid over het wel of niet opnemen van chemische eigenschappen, monitoringsdata, toxicologische data en analytische data. Veelvuldig komt het argument naar voren dat gegevens kunnen worden opgezocht in andere databases. Het zou een goede aanvulling op bestaande databases zijn om een lijst stoffen op te nemen die worden geanalyseerd bij de drinkwaterlaboratoria, en een lijst met stoffen die worden aangetroffen in Nederlandse (drink)wateren en de bronnen daarvoor. Het delen van monitoringsgegevens is een gevoelig punt. Een koppeling naar REWAB/RIWA en BTO gegevens wordt genoemd als pluspunt. Sterk naar voren komt het belang van het opnemen van specifieke zuiveringsrendementen (per techniek, matrix en procescondities). Deze informatie is over het algemeen lastig te ontsluiten, deels kunnen we de informatie opnemen uit de watershare tool AbatES. Een koppeling naar bestaande QSAR modellen zoals bv. Zuivering en Toxiciteit. Zowel een opzet in het Nederlands en in het Engels worden geprefereerd. De meerwaarde van deze database is tweeledig. Enerzijds voegt deze database aan de bekende stoffen veel gevraagde stoffeigenschappen toe. Anderzijds beoogt deze database een actueel overzicht te geven van nieuwe- en onbekende stoffen.
Output	Voor de gebruiker is de gewenste (lokale) output flexibel, waarbij parameters kunnen worden aangevinkt en gesorteerd naar wens. Het gewenste format van de output is Excel of CSV. Respondenten zijn verdeeld of er een korte factsheet moet komen van elke stof.
Toepassingen	De belangrijkste toepassing van de database is het ontsluiten van gegevens, met name gegevens uit SPO en BTO onderzoek, zoals de zuiveringsrendementen.

3.2 Uitkomst workshop

Op 23 juni 2015 zijn zeven experts bij elkaar gekomen om te discussiëren over de gewenste opzet van een database. Hieronder is de *word cloud* weergegeven van de notulen. De woorden (stoffen) database, (bestaande) databases, gegevens, informatie en initieerder KWR

zijn de meest voorkomende woorden. Andere opvallende woorden zijn ontsluiten, belang en onbekende (stoffen).



Figuur 3-1 Word cloud van de notulen van de workshop.

Algemene opmerkingen van de deelnemers zijn:

1. Het opslaan van waar welke gegevens zijn gevonden voorkomt herhaling van werk.
2. Er is geen behoefte aan het opzetten van een database, voor gegevens die elders zijn te vinden.
3. De database ontsluit informatie voor onervaren collega's.
4. Het delen en ontsluiten van gegevens is een basis voor goede samenwerking tussen verschillende partijen.
5. De combinatie van gegevens uit verschillende disciplines is nuttig voor prioriteringsdoelenden, mits alle benodigde gegevens opgenomen zijn.

Welke partijen hebben een rol in de database en wat is hun taak?

1. Waterbedrijven en waterlaboratoria krijgen de taak doelstoffen aan te leveren. Daarnaast kunnen zij nieuw gesignaleerde stoffen invoeren, waarna gegevens kunnen worden aangevuld.
2. De taken van KWR zijn: het beheer en onderhoud van de database, de validatie (correctheid) van gegevens en het signaleren van nieuwe stoffen. KWR houdt tevens het gebruik van de database bij.

Welke informatie is nodig om de database een waardevolle aanvulling te laten zijn op bestaande databases?

1. De invoer bestaat uit stoffen gekoppeld aan een unieke 'identificier', de InChI-code met daarbij een naam/synoniem en een CAS nummer.
2. De focus ligt op de stoffen onderzocht in BTO-verband en in het analysepakket van laboratoria. Daarnaast kunnen nieuwe stoffen aangetoond in (bronnen van) drinkwater worden toegevoegd. Deze stoffen zijn nog niet opgenomen in REWAB.

3. Voor bestaande gegevens wordt een koppeling gemaakt waar de informatie te vinden is.
4. Het is belangrijk om bronvermelding te geven bij alle gegevens.
5. De koppeling tussen stoffen met gelijke eigenschappen is nuttig (via bijvoorbeeld een QSAR), zodat gidsstoffen kunnen worden geïdentificeerd die bijvoorbeeld wel goed meetbaar zijn t.o.v. de vergelijkbare stof.
6. De verwijzing bij welk laboratorium de stof wordt/werd gemeten is een waardevolle aanvulling. Op deze manier kunnen de waterlabs zelf contact opnemen met juiste contactpersoon.

Het gaat dus om het toevoegen van stoffeigenschappen die vaak gebruikt worden maar lastig te verzamelen zijn en het toevoegen van nieuwe stoffen die nieuw of nog niet geïdentificeerd zijn. Bestaande database structuren zijn niet bedoeld om dergelijke informatie op te slaan en zijn voor deze toepassing niet geschikt.

Wie zijn de eindgebruikers, en op welke manier willen zij informatie ontsluiten?

1. De primaire eindgebruikers zijn de waterlaboratoria, de drinkwaterbedrijven en KWR. RIVM en waterschappen zijn secundaire gebruikers.
2. De drinkwaterlaboratoria en KWR hebben verschillende belangen. Het belang van de laboratoria ligt vooral bij locatie-specifieke gegevens voor hun eigen bedrijf. KWR heeft belang bij de vergaring en analyse van landelijke gegevens.
3. De meeste interesse ligt bij een ontsluitingsstool in plaats van een lokale database. De ontsluitingsstool is een combinatie van stoffen ('identifiers') en bronvermelding en/of koppelingen naar gegevens.
4. Het voordeel van een database met bronvermelding en/of koppelingen naar gegevens ligt in het feit dat de informatie zelf opgeslagen ligt in externe valide databases. Het beheer van de informatie is dan ook in handen van externe partijen en bespaart KWR en de drinkwaterlaboratoria veel tijd en geld die anders gestoken zou zijn in het onderhoud van een lokale database.
5. Het ontsluiten van informatie kan via links naar BTO rapporten op BTO-net en andere bronnen die handmatig moeten worden bekeken of via koppelingen naar bestaande databases die automatisch informatie ontsluiten. Automatische koppelingen kost echter (ook) onderhoud (dit lijkt geen grote kostenpost) en kan lastig zijn met privacy en rechten.
6. De database moet web-based, flexibel en uitbreidbaar zijn.
7. Daarnaast komt het belang naar voren voor het uitwisselen van screeningsgegevens tussen de laboratoria over onbekende stoffen. Van deze stoffen is hun identiteit en daarmee hun unieke 'identificer' echter niet bekend, zodat deze stoffen met hun accurate massa, retentietijd en MS fragmentatie in een aparte tabel in de database moeten worden opgenomen.

Waar ga je de opgevraagde informatie uit de database voor gebruiken?

Deze probleemstelling is niet uitvoerig bediscussieerd tijdens de workshop. Tussen de regels door ontdekten wij drie toepassingen:

1. Snel en up-to-date informatie ontsluiten waar anders meer tijd en moeite zou worden ingestoken.
2. Onderzoeksarchief van chemische stoffen
3. Inzicht in welke antropogene stoffen gemeten moeten worden in het kader van het drinkwaterbesluit. De database biedt de mogelijkheid voor het toepassen van een (prioriterings) model, zoals de meetstrategie ontwikkeld door HWL (Houtman 2015)

waarmee drinkwaterbedrijven kunnen bepalen welke antropogene stoffen specifiek voor hen relevant zijn en in het monitoringsprogramma opgenomen dienen te worden.

3.3 Uitkomst tweede workshop

In de tweede workshop had als doel het definitieve pakket van eisen en wensen en een definitieve opzet voor de database vast te stellen. De uitkomst van deze workshop is het template in Bijlage III met de gewenste parameters voor opname in de database. Deze opzet is gebruik bij de kostenraming (paragraaf 4.2).

Het doel van de database is enerzijds het maken van een overzicht van enkele chemische eigenschappen van bekende chemische stoffen en anderzijds het uitwisselen van informatie met betrekking tot het voorkomen onbekende stoffen. De inspanning om de database te vullen moet beperkt blijven.

De op te nemen parameters volgens Bijlage III zijn onderverdeeld in zeven secties:

- 1) Unieke identifiers
- 2) Stofeigenschappen
- 3) Stofeigenschappen uit BTO onderzoek
- 4) Makkelijk vindbare stofeigenschappen
- 5) Monitoring
- 6) Analytische gegevens onbekenden
- 7) Monitoring onbekenden

Sectie 1-5 heeft betrekking op de bekende stoffen, secties 6 en 7 op de onbekenden. Alle parameters in deze hoofdstukken zijn gewenst of vereist voor opname. Enkel de stofeigenschappen met een lagere prioriteit zijn gemakkelijk op te zoeken in bestaande database en zijn niet noodzakelijk voor opname maar kunnen later worden toegevoegd.

In sectie 1 van de template zijn de stof identifiers opgenomen. Deze zijn belangrijk voor opname in de database. Dit zijn: naam, CAS, INCHI, IUPAC naam, lab-specifieke AQUON-naam, REWAB naam en REWAB code en REACH code. Het is de bedoeling dat de gebruiker één van deze identifiers invoert en dat de andere namen worden opgezocht.

In sectie 2 van de template zijn de stofeigenschappen opgenomen die gewenst zijn voor opname in de database. Dit zijn: brutoformule, molecuulstructuur, stofgroep, humaan-toxicologisch relevantie en wel/geen prioritaire stof. Een stof kan als meerdere stofgroepen worden getypeerd. Gewenst is een linkje naar waar deze informatie te vinden is. Daarnaast moet worden aangegeven of een stof is opgenomen in een bepaalde lijst met prioritaire stoffen (HWL, KWR, watchlist).

In sectie 3 van de template zijn enkele stofeigenschappen opgenomen uit BTO en sector onderzoek zoals de provisional drinking water quality guideline value, en zuiveringsrendementen. Deze eigenschappen zijn gewenst voor opname in de database. De provisional drinking water quality guideline value wordt opgenomen als link naar een referentierapport. Indien de gebruiker dat wil kan deze de waarde in de database worden ingevoerd. De zuiveringsrendementen worden opgenomen als referentie naar het rapport met betreffende gegevens. Zonder context kunnen zuiveringsrendementen niet worden geduid (afhankelijk van condities).

In sectie 4 in de template zijn enkele stofeigenschappen opgenomen die makkelijk op te zoeken zijn en niet noodzakelijk zijn voor opname in de database, zoals log Kow, log D,

dampspanning, halfwaardetijd en TTC. Deze kunnen echter wel weer van belang zijn voor het identificeren van onbekenden.

Sectie 5 in de template hoort bij het kopje monitoring. De laboratoria zien graag voor welke stoffen andere laboratoria over een analysemethode beschikken. In de database wordt dit opgenomen met een kruisje per lab (wel niet gemeten) of de rapportagegrens. De laboratoria krijgen hiervoor een template waarin ze aangeven welke stoffen ze kunnen monitoren. De laboratoria zijn zelf verantwoordelijk voor het updaten van deze gegevens.

Monitoringsgegevens kunnen wanneer beschikbaar uit REWAB worden geïmporteerd. Hierbij kan de REWAB locatie worden vermeld. Wanneer de stof niet in REWAB is opgenomen dan kan de naam van de locatie en de x, y, z coördinaten worden opgenomen. Het opnemen van de normstelling van de stoffen is mogelijk, doordat dit eenvoudig uit REWAB kan worden geïmporteerd.

Tot slot is tijdens deze workshop de zorg uitgesproken dat het vertrouwelijke data betreft en dat deze alleen binnen de drinkwaterwereld beschikbaar mogen komen. Er moet dus een goede beveiliging op de toegang tot de database komen.

3.4 Samengevat

Door middel van de resultaten van de enquête en de workshop is er een basis gelegd voor de opzet van een stoffendatabase vanuit de sector. Ten eerste is er de behoefte aan een ontsluitingstool voor het vergemakkelijken voor het opzoeken van gegevens. De ontsluitingstool maakt koppelingen tussen relevante stoffen (ingevoerd door waterlaboratoria en KWR) en bestaande locaties waar informatie is opgeslagen zie Figuur 3-2). De drinkwaterlaboratoria zien het belang al bij een bronvermelding of een link naar waar de data eventueel handmatig kan worden opgezocht. KWR hecht vooral waarde aan de koppeling met bestaande databases met kwalitatief goede gegevens zodat de gegevens zo veel mogelijk automatisch worden ontsloten. De te maken koppelingen zijn uitgezet in Figuur 4-1.



Figuur 3-2 Opzet van een ontsluitingstool.

Praktisch gezien houdt een ontsluitingstool in dat er geen gegevens worden opgeslagen in een nieuwe database maar enkel koppelingen worden gemaakt naar bestaande database. Echter, dit is niet haalbaar voor de enkele gewenste parameters, zoals zuiveringsrendementen, 'welke laboratorium deze stof kan meten' of het vastleggen van gegevens over onbekende stoffen gevonden binnen de bedrijfstak. Voor deze parameters zal wel een database worden gebouwd.

Koppelingen worden wel gemaakt voor het opzoeken van de unieke identifiers. Wanneer één van de identifiers wordt ingevuld kunnen de andere identifiers worden ontsloten uit andere databases. Tevens kan via het CAS-nummer monitoringsgegevens uit de REWAB database worden ontsloten. Mogelijk kunnen ook later koppelingen worden toegevoegd aan de Estimation Program Interface (EPI) Suite (US EPA, 2015) en toxicologische eigenschappen van

Hazardous Substances Data Bank (HSDB) van TOXNET door gebruik te maken van de Inchi code. De EMPODAT database, ontwikkeld binnen het Europese NORMAN netwerk, zal ook uitgebreid worden met Inchi codes zodat cross referentie mogelijk wordt. De ontwikkeling van databases voor het identificeren van onbekenden en suspects gaat hard. In 2016 zal de NORMAN Massbank uitgebreid worden met de mogelijkheid voor het opslaan van onbekende verbindingen. Mogelijk kunnen we in de toekomst een koppeling maken of aansluiten bij bestaande databases ontwikkeld voor de identificatie van gemonitorde onbekende stoffen. Een volledige lijst met de meest relevante bestaande databases is te vinden in Tabel 3-2. Deze zullen niet allemaal worden gekoppeld aan de database, maar zijn voornamelijk bedoeld ter informatie.

Tabel 3-2 Bestaande databases.

Database	Omschrijving
Analytische chemie	
NIST MS database	Standaard referentie data voor massaspectrometrie
NIH/CADD database	Omzetting van en naar identifiers voor chemische structuren
Wiley	10 databases over chemische stoffen
(NORMAN) Massbank	Hoge kwaliteit massaspectrometrie database (open source)
STOF-IDENT	Analytische informatie om identificatie van water-relevante stoffen te ondersteunen
Metlin	Hoge kwaliteit massaspectrometrie database voornamelijk gericht op metabolomics (open source)
MZcloud	Hoge kwaliteit massaspectrometrie database (open source)
Monitoring	
REWAB	REgistratie WAterkwaliteitsgegevens Bedrijven (REWAB). Monitoringsgegevens drinkwaterbedrijven
RIWA database	Meetgegevens microverontreinigingen in oppervlaktewater rivierwaterbedrijven
WATSON	Microverontreinigingen in influent en effluent van RWZI's
Waterbase (RWS)	Meetgegevens microverontreinigingen rijkswateren
Bestrijdingsmiddelenatlas CML	Meetgegevens van bestrijdingsmiddelen in oppervlaktewater
Waterkwaliteitsportaal	Waterkwaliteitsgegevens
EMPODAT	Europese monitoringsdata van 'emerging substances'
Emissiegegevens	
Emmissieregistratie website	Geregionaliseerde uitstoot (emissie) van circa 350 verontreinigende stoffen
Stofeigenschappen	
Chemspider	Database chemische structuren
Epi Suite	Fysisch-chemische eigenschappen en programma's die het gedrag in het milieu voorspellingen
Chemaxon	Cheminformatica software platformen, applicaties en diensten
The pesticide properties database (IUPAC)	Fysisch-chemische en ecotoxicologische data van bestrijdingsmiddelen
Pharmaceutical Explanatory database (PaDEL)	Voorspelling van 1400 chemische descriptoren
Virtual Computational Chemistry Laboratory	Voorspelling van logP, wateroplosbaarheid en pKa(s)
Zuiveringsrendementen	

AbatES	Eigenschappen van Emerging Substances (Watershare Tool)
Toxicologische eigenschappen	
EU Echem portal van de OECD	Allerhande toxicologische databases te benaderen via het EU Echem portal van de OECD
TOXNET	Portaal naar toxicologische databases
Hazardous Substances Data Bank (HSDB)	Wetenschappelijke toxicologische data voor >5000 stoffen
Integrated Risk Information System (IRIS)	Risicobeoordelingen humane gezondheid voor >500 stoffen
International Agency for Research on Cancer (IARC)	Carcinogene stoffen

Het is gewenst om in de database zowel stofeigenschappen als gegevens over onbekende stoffen op te nemen. Alle gewenste parameters opgenomen in het template (Bijlage III). De database kan op kleine schaal worden opgezet. Vervolgens zijn er mogelijkheden voor uitbreiding. De eerste set-up bevat tenminste de volgende informatie:

- Welk laboratorium kan de stof meten
- Specifieke zuiveringsrendementen voor bepaalde zuiveringstechnieken

Het delen van informatie over aangetroffen (onbekende) stoffen is een waardevolle aanvulling op bedrijfsspecifiek screeningsonderzoek. Deze stoffen hebben echter geen bevestigde identiteit en daardoor ook geen unieke identifier. Deze stoffen worden ondergebracht in de database, in een specifiek gedeelte voor onbekende stoffen, in dit gedeelte wordt specifiek informatie opgeslagen over de analytische methode, de meetwaarde en de meetlocatie.

4 Functionele eisen en opzet database

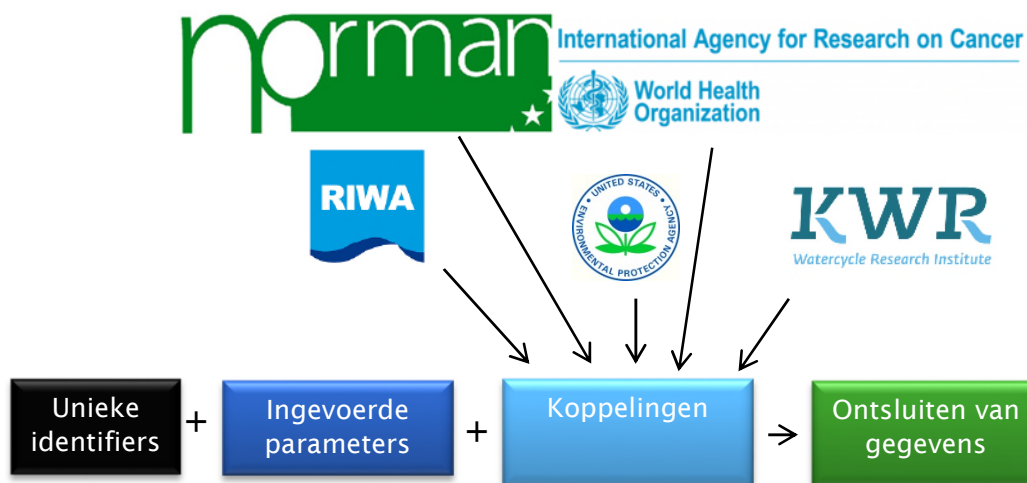
In hoofdstuk drie is de inhoudelijke opzet van de stoffendatabase uiteengezet: het pakket van eisen en wensen. In dit hoofdstuk wordt de technische uitwerking van de stoffendatabase beschreven (paragraaf 4.1). Vervolgens werd samen met een ICT'er van KWR de haalbaarheid en het kostenplaatje ingeschat voor de opzet van de database (paragraaf 4.2).

4.1 De functionele eisen

4.1.1 Algemeen

Uit de enquête en de workshops kwam naar voren dat de ontsluitingstool web-based, flexibel en uitbreidbaar moet zijn:

1. De database is een combinatie van een database met actuele stofeigenschappen, links naar de locatie waar deze eigenschappen zijn op te zoeken en een database met screeningsgegevens voor het delen van informatie over onbekende stoffen.
2. De database is een combinatie van ingevoerde en opgeslagen informatie van gegevens die nog niet zijn opgeslagen in bestaande databases en een ontsluitingsstool (met koppelingen naar bestaande databases), zie Figuur 4-1.
3. Een web-based interface maakt het mogelijk dat er met de ontsluitingstool gewerkt kan worden via een met internet toegankelijke webpagina. De database wordt beveiligd, bijvoorbeeld met een VASCO password.
4. Met een flexibele in- en output kunnen parameters of stoffen worden aangevinkt en gesorteerd naar wens.
5. De tool moet uitbreidbaar zijn. Er wordt gestart met het ontsluiten van informatie volgens Figuur 4-1. In de toekomst kunnen er koppelingen worden toegevoegd waardoor informatie van aanvullende parameters ook gemakkelijk kan worden ontsloten.
6. Wie welke informatie toegevoegd wordt bijgehouden in een audit trail.
7. Het beheer ligt bij KWR en zal daarvoor een beheerder aanwijzen.
8. Mogelijk kan worden bijgehouden hoe vaak de database wordt gebruikt, wat waardevolle informatie geven hoe de database zich wel of niet ontwikkeld.



Figuur 4-1 Globale opzet van de database.

4.1.2 Functionele eisen

Er zijn twee verschillende invoerschermen: Een voor bekende stoffen en een voor onbekende stoffen. In de template in Bijlage III zijn de opgenomen parameters gespecificeerd. Voor een aantal parameters is een koppeling met een externe database vereist. Hieronder volgt een lijst met de te koppelen databases. Sommige zijn eenmalige stand alone databases, andere zijn links naar websites.

- Database met synoniemen InChi, Cas en IUPAC (koppeling Open Babel)
- Database met bruto formules (koppeling via IUPAC-naam (koppeling Open Babel))
- Database met stofstructuren (koppeling via IUPAC-naam (koppeling Open Babel))
- Database met humane toxicologische relevantie (koppeling met Toxnet, HSDB, ITER, IPCS INCHEM of ECHA's eChemPortal)
- Database met REWAB data. Eenmalige download van database.
- Database met daarin een lijst welk lab welke stoffen analyseert (de Lab-database). Deze Lab-database wordt eenmalig aangeleverd, daarna door de gebruikers up to date gehouden.
- Mogelijkheid om hyperlinks op te nemen naar rapportages.
- Database met stoffen in REACH (koppeling naar European Chemicals Agency)

Functionaliteiten van de database voor bekende:

- Controle op correcte invoer en naamgeving (met o.a. CAS checksum, Inchi code) (American Chemical Society, 2015, NIST, 2002, IUPAC 2015)
- Invoer voor meerdere stoffen tegelijk moet mogelijk zijn via een vooraf vastgelegd format. Kan in excel of CSV format zijn. Liefst zo simpel mogelijk.
- Opslag van ingevoerde gegevens in een nieuwe database. Alle wijzigingen worden gelogd met gebruikersnaam.

Functionaliteiten Lab-database:

- Gebruikers moeten de Lab-database kunnen updaten. Alle wijzigingen worden gelogd met gebruikersnaam.

Functionaliteiten van onbekende database:

- Invoer voor meerdere stoffen tegelijk moet mogelijk zijn via een vooraf vastgelegd format. Kan in excel of CSV format zijn. Liefst zo simpel mogelijk. In de input zijn coördinaten, type monster en concentratie equivalenten verplichte velden.
- Controle van lab gegevens en controleren met bekenden. Indien een stof vermoedelijk een match geeft met een bekende stof moet dit worden aangegeven,
- Een onbekenden stijgt naar mate de identificatie vordert van level 5 (accurate massa bekend), Level 4 (eenduidige brutoformule), level 3 (mogelijke kandidaat), Level 2 (kandidaat bevestigd door middel van referentie spectrum) naar Level 1 (kandidaat bevestigd met standaard stof.
- Opslag gevonden stoffen in een nieuw aan te maken database.
- In de nieuwe database moet kunnen worden gezocht op een gewenste combinatie waarvan retentie en accurate massa verplicht zijn. Output liefst in kaart format, maar in tabel kan ook. Gebruiker moet kunnen aangeven hoe exact retentie en accurate massa moeten overeenkomen.

Overige functionaliteiten:

- De nieuwe stoffendatabase is web-based.
- De toegankelijkheid moet goed worden beveiligd. Gegevens zijn zeer vertrouwelijk dus mogen niet publiekelijk beschikbaar komen.
- In- en output worden gelogd, zoals wie voegt wat toe en hoe vaak logt iemand in.

Bijkomende documenten:

- Gebruikershandleiding
- Richtlijn met criteria om onbekende stoffen te selecteren

4.1.3 Invoer, uitvoer en onderhoud

De belangrijkste aandachtspunten met betrekking tot de database liggen in de beveiliging van de vertrouwelijke data en de vulling van de dataset. De gebruikers willen niet heel veel tijd kwijt zijn aan het vullen van de database. Dat moet mogelijk zijn door de import en export van data zo eenvoudig mogelijk te maken. Dit betekent dat er bijvoorbeeld delen uit de LIMS systemen via een vast format moeten kunnen worden geïmporteerd. De vulling van de database vindt plaats door de gebruikers. KWR zal bij het uitkomen van nieuwe rapportage de relevante gegevens toevoegen aan de database. Bij oplevering, zal KWR de database vullen met stofgegevens die KWR de laatste drie jaar heeft verzameld. Oudere rapportages worden door de gebruikers zelf ingevoegd. KWR wijst een trekker aan (operator) die jaarlijks de gegevens evalueert.

Mogelijk stofeigenschappen voor opname in de database zijn weergegeven in Bijlage III.

Het is de bedoeling dat bij een nieuwe invoer van een bekende stof de gebruiker ten minste één van de unieke identifiers invoert en dat de andere namen worden opgezocht.

Bij de start van de database kunnen organische microcontaminanten uit REWAB worden ingevoerd. Vervolgens kunnen alle niet in de verplichte monitoring gemeten stoffen worden toegevoegd door de laboratoria.

Het gewenste format van de output is een CSV file flexibel in opzet, waarbij van enkele of meerdere stoffen gegevens en een selectie van parameters kan worden uitgevoerd.

4.1.4 Koppelingen

In Tabel 4-1 staan enkele parameters voor opname in de database die zijn op te zoeken in bestaande databases. Een automatische koppeling haalt de gegevens op en ontsluit deze voor eigen toepassingen. Het omzetten van identifiers, zoals CAS-nummer, InChI en SMILES code (gebruikt voor QSAR benaderingen), kan via koppelingen. Daarnaast worden koppelingen gemaakt met REWAB en kunnen zuiveringsrendementen worden ontsloten met de watershare tool AbatES en gemodelleerd via QSARs.

Tabel 4-1 Parameters voor opname in de ontsluitingstool die worden ontsloten via koppelingen.

Parameter	Koppeling	Mogelijkheid om koppeling te maken
Identifiers	Chemical Translation Service (on-line)	Ja, koppeling mogelijk via script
Specifieke zuiveringsrendementen voor bepaalde zuiveringstechnieken	AbatES	In overleg
	Toepassing QSARs KWR (nanomembraanfiltratie en UV- verwijdering)	In overleg
Monitoringsgegevens	RIWA	In overleg
	REWAB	In overleg
	EMPODAT	In overleg
Identificatie van onbekenden	Massbank (MS/MS spectrum)	Koppeling naar URL mogelijk
	Stoff-Ident	In overleg

4.2 Gebruikers

Gebruikers zijn te onderscheiden in twee categorieën de primaire en secundaire eindgebruikers. De primaire eindgebruikers zijn managers, onderzoekers, laboranten en beleidsmakers van de waterlaboratoria, drinkwaterbedrijven, VEWIN en KWR. De secundaire gebruikers zijn, RIVM, Rijkswaterstaat en waterschappen; zij zijn nog niet betrokken in dit stadium.

4.3 De haalbaarheid en globale kostenraming

Op basis van het pakket van eisen en wensen, de functionele eisen is een grove schatting gemaakt van de kosten voor een dergelijke opzet van de database (Figuur 4-2). De totale kosten komen uit op ongeveer 50 k€, voornamelijk bestaande uit de systeemanalyse (7,5 k€), de schil van de database (2,5 k€), de inhoud van de database (10 k€) en de functies (8 k€).

De werkelijke prijs kan 50% afwijken van deze schatting. Een systeemanalyse geeft een nauwkeurige schatting van de werkelijke kosten. De systeemanalysezelf kost 10 k€ en kan een eerste fase zijn waarop een go or no go beslissing gemaakt kan worden. Het resultaat van de systeemanalyse vormt de feitelijke opzet van de database. Dus na de systeemanalyse kan een definitieve raming worden opgesteld.

Tabel 4-2 Kostenraming van de opzet van de database (hierin is geen onderhoud, beheer of de invulling van de database in opgenomen).

Onderdeel	Kosten (k€)
Systeemanalyse	7,5 k€
Schil database	2,5 k€

Inhoud database	10 k€
Functies	8 k€
waarvan invoer + voorwaarden	2 k€
waarvan uitvoer	4 k€
waarvan audit trail	1,6 k€
waarvan checksum	0,4 k€
Begeleiding systeemanalisten	4 k€
Beveiliging	1 k€
Testen afzonderlijke functionaliteiten	3 k€
Communicatie	4 k€
Onvoorzien	10 k€
Totaal	50 k€

De inhoud van de database is geraamd aan de hand van het template, met 48 parameters gewenst voor opname in de database. Elk veld, bijbehorende bij elke parameter heeft een bepaalde structuur. De opzet van een enkel veld is simpeler dan de opzet van een onderliggende tabel. De kostenraming per parameter is weergegeven in Bijlage III. De duurste velden zijn de velden waar er een koppeling benodigd is naar een externe database. De koppeling in een GIS-omgeving is niet opgenomen in de globale kostenraming.

De bovengenoemde kosten zijn enkel de kosten voor het vormgeven van de database. De invulling en het beheer zijn hierin nog niet opgenomen. Een grove schatting voor het beheer is per lab €5 per jaar en een eerste invulling van de database brengt kosten met zich mee afhankelijk van de hoeveelheid gegevens.

5 Referenties

US EPA (2015). Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.11 or insert version used]. United States Environmental Protection Agency, Washington, DC, USA.

Hazardous Substances Data Bank (2015). TOXNET toxicological data network. US National Library of Medicine (US). Available from: <http://toxnet.nlm.nih.gov/newtoxnet/hsdb.htm>.

RIWA database Nieuwegein.

REWAB database.

American Chemical Society, 2015. Check Digit Verification of CAS Registry Numbers. Chemical Abstracts Service (CAS), A division of the American Chemical Society.
<https://www.cas.org/content/chemical-substances/checkdig>

National Institute of Standards and Technology (NIST), 2002. Federal information processing standards publication 180-2. Specifications for the Secure hash standard.
<http://csrc.nist.gov/publications/PubsFIPS.html#fips180-4>

Bijlage I

Samenstelling expertteam

Het team van experts vanuit de drinkwaterlaboratoria en KWR voor de begeleiding van deze haalbaarheidsstudie.

Naam	Bedrijf
Adrie Krom	HWL
Albert de la Mar	Aqualab Zuid
Barry Pieters	HWL
Bernard Bajema	Vitens
Dirk Vries (tot november 2015)	KWR
Erik Emke	KWR
Jan van der Kooi	WLN
Leo Puijker	KWR
Rosa Sjerps	KWR
Ton van Leerdam	KWR
Mario Maessen (i.p.v. Dirk Vries sinds november 2015)	KWR

Bijlage II

Enquête & respons

De enquête samen met de verzamelde respons van de betrokken stakeholders is toegevoegd als een Excel-file.

Bijlage III

Template parameters gewenst voor
opname in database en
kostenraming voor inhoud
database

Zie toegevoegde excel file.

Houtman, C.J., Pieters, B.J. Velzeboer, I., Kroesbergen, J. (2015) Organische stoffen in het Drinkwaterbesluit: 1. Hoe geven we invulling aan de eisen voor 'overige' antropogene stoffen? H2O.