

DE VROEGE IDENTIFICATIE VAN RISICO'S VOOR DE VOEDSELVEILIGHEID

Fred van de Brug

T +31 30 694 41 27, F +31 30 694 40 70
E fred.vandebrug@tno.nl, WWW.TNO.NL

Ineiding

De veiligheid van ons voedsel is de afgelopen jaren op een zeer acceptabel niveau gekomen. Helaas worden we zo nu en dan toch geconfronteerd met een geval van contaminatie in het voedselproductienetwerk. Voorbeelden van recente incidenten zijn “*verdacht carcinogeen acrylamide ontstaat bij verhitten van sommige voedingsmiddelen (2002)*” en “*carcinogeen dioxine via klei aan aardappelschillen als veevoeder in melk (2004)*”.

De volgende trends zijn van belang voor de ontwikkeling en de verspreiding van zowel bekende als onverwachte risico's:

- Toename van complexiteit en dynamiek van de voedselproductie door de globalisering en open grenzen.
- De toenemende behoefte aan diversiteit van het voedselaanbod onder meer als gevolg van de multiculturele samenstelling van de bevolking.
- De risico's van gecombineerde blootstelling van stoffen en “nieuwe” stoffen.
- Het ontstaan van nieuwe microbiële gevaren als gevolg van resistentie.
- Het gebruik van nieuwe technologie om voedingsmiddelen te produceren.

De grootste zorg met betrekking tot incidenten van contaminaties van voeding ligt op het gebied van eventuele gezondheidsschade. Daarnaast is er ook de maatschappelijke onrust en kan de economische schade voor bedrijven en soms voor gehele sectoren in de agro-food industrie oplopen tot vele miljoenen Euro's per incident.

In een aantal gevallen van voedselveiligheid incidenten bleek achteraf dat er informatie aanwezig was geweest voorafgaand aan het feitelijke incident. Indien bedrijven en overheid vóór het moment van het ontstaan van het incident, over de relevante informatie hadden kunnen beschikken was men in staat geweest om prioriteiten te stellen en opties voor beheersmaatregelen te ontwikkelen. Hierdoor was het mogelijk geweest om de schade voor de maatschappij te beperken en mogelijk zelfs te voorkomen.

Het onderliggende informatieprobleem is dat het aantal informatiebronnen zeer groot is en dat deze bronnen divers van aard zijn. Daarnaast is er het probleem naar wat voor soort informatie gezocht moet worden en hoe deze informatie vervolgens in verband moet worden gebracht met voedingsmiddelen, grondstoffen of processen.

In de uitgangssituatie bij aanvang van het Food Informatics

project (zie www.vl-e.nl) had TNO niet de beschikking over een informatiesysteem welke in staat was om de identificatie van opkomende risico's goed te ondersteunen. Hier lag de uitdaging om informatiesystemen te ontwikkelen om op kosteneffectieve wijze de relevante informatie over mogelijke risico's te extraheren uit de enorme hoeveelheid beschikbare informatie.

Requirements

Een aantal historische cases, waaronder de al eerder genoemde cases acrylamide (2002) en dioxine (2004), zijn geanalyseerd en hieruit zijn generieke requirements afgeleid. Deze requirements zijn generiek van aard om ze te kunnen valideren met behulp van historische cases en om er zeker van te zijn dat, met behulp van deze requirements, de kans toeneemt ten opzichte van de uitgangssituatie dat ook in de toekomst opkomende risico's tijdig geïdentificeerd kunnen worden.

De relevante concepten uit de historische cases zijn bepaald en deze zijn gebruikt om een prototype ontologie te ontwerpen en te vullen met begrippen. Daar nieuwe en vooral onverwachte gevaren voor de voedselveiligheid zich kenmerken door gebrek aan a priori kennis van begrippen is het ook nodig om gebruik te kunnen maken van context informatie, dat wil zeggen informatie rondom de nog onbekende begrippen.

Wat databronnen betreft is het beeld ontstaan dat relevante informatie met betrekking tot opkomende risico's aanwezig kan zijn in wetenschappelijke literatuur en in media berichten.

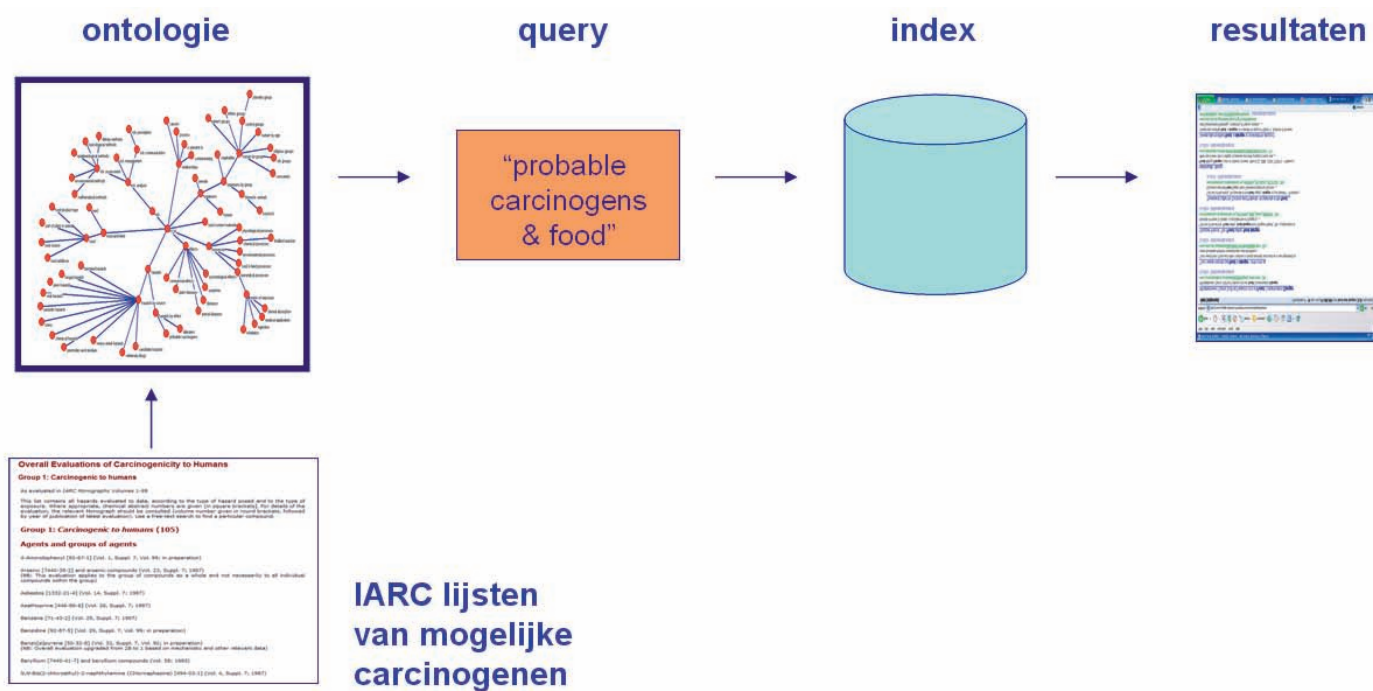
De belangrijkste requirements zijn als volgt:

- Het kunnen gebruiken van eigen document collecties waarin gezocht moet kunnen worden.
- Het kunnen inladen van ontologieën.
- Deze ontologieën op eenvoudige wijze kunnen gebruiken bij de query constructie.
- Het kunnen extraheren uit documenten van nieuwe begrippen met behulp van context woorden.

Om de binnen het vl-e ontwikkelde tools te kunnen ontwikkelen en te evalueren zijn test data verzameld. De volgende twee use case's zijn gebruikt om de toepassing van de binnen het vl-e ontwikkelde tools te demonstreren.

Use case 1: zoeken met bekende namen van carcinogene substanties

De International Agency for Research on Cancer (IARC),



Figuur 1

onderdeel van de WHO, heeft als een van haar taken het classificeren van factoren welke een risico kunnen vormen voor het ontstaan van kanker bij de mens. De factoren omvatten ondermeer chemische stoffen, maar ook biologische substanties en complexe mengsels. Inmiddels zijn ongeveer 400 factoren (waaronder acrylamide en dioxine) geassocieerd als carcinogene of als mogelijk carcinogene voor de mens. Voor dit soort stoffen en andere factoren hanteert men op dit moment geen maximum concentratie waaronder er geen negatief gezondheidseffect zal optreden, omdat men er vanuit uit gaat dat een enkel molecuul de oorzaak kan zijn van het uiteindelijke ontstaan van kanker. Het moge duidelijk zijn dat deze factoren ongewenst zijn in voeding.

De lijsten van stoffen en substanties welke de IARC heeft geassocieerd als zijnde carcinogene of als mogelijk carcinogene voor de mens worden gepubliceerd op de website van de IARC en kunnen dus gebruikt worden bij het zoekproces in documentcollecties of deze stoffen worden genoemd in relatie tot voeding. De betreffende lijsten van stoffen, inclusief de andere factoren, zijn gedurende het project opgenomen in de ontologie. Om gebruik te kunnen maken van de ontologie is er behoefte aan een zoekstelsel waarmee in indexen van de documentcollecties kan worden gezocht.

Binnen het vl-e is een informatiesysteem ontwikkeld waarmee:

- Eigen document collecties kunnen worden geïndexeerd.
- In deze indexen kan worden gezocht.
- De queries kunnen worden uitgebreid met informatie uit een eigen ontologie.
- Een eigen ontologie kan worden geladen in een repository.

Met behulp van dit informatiezoekstelsel is het mogelijk om gericht te zoeken in indexen van documentcollecties met behulp van de in de ontologie opgeslagen namen van factoren welke een risico kunnen vormen voor het ontstaan van kanker bij de mens in combinatie van namen van voedingsmiddelen, evenals opgeslagen in de ontologie.

Use case 2: zoeken met voorheen onbekende namen van mogelijk carcinogene substanties

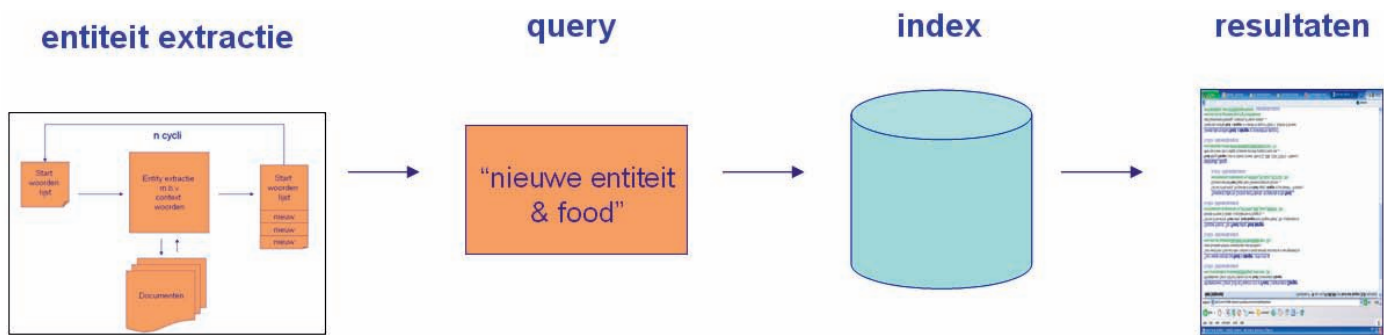
Naast de door de IARC organisatie geassocieerde stoffen zijn er stoffen welke wel mogelijk verdacht zijn en door de IARC mogelijk in overweging zijn genomen gedurende de evaluatie procedures, maar uiteindelijk niet zijn opgenomen in de IARC lijsten. Ook van deze stoffen welke geassocieerd zijn met het ontstaan van kanker bij mens of bij dier willen we weten of deze in verband kunnen worden gebracht in relatie tot voeding.

Over deze stoffen en hun effecten wordt gepubliceerd in de wetenschappelijke literatuur. Hierdoor is de wetenschappelijke literatuur tevens een bron om de namen van dit soort van stoffen (entiteiten) te identificeren, te extraheren en vervolgens te gebruiken bij een query. Het volume van de wetenschappelijke literatuur is echter enorm groot en om de relevante informatie er door experts uit te laten halen is kostbaar. Daarom is er behoefte aan een geautomatiseerde wijze om dit te doen.

Binnen het vl-e is een software tool ontwikkeld waarmee:

- Nieuwe entiteiten uit documenten worden geëxtraheerd.
- De extractie vindt plaats met behulp van context informatie.
- Als input kunnen eigen data worden gebruikt.

Met behulp van deze software tool is het mogelijk om de



Figuur 2

in de "use case 1" genoemde namen van carcinogenen of mogelijk carcinogenen voor de mens als input voor de tool te gebruiken. De tool zoekt vervolgens in een aantal iteraties in een document collectie met behulp van context informatie naar nieuwe entiteiten welke genoemd worden in dezelfde context als waarin de stoffen op de input lijst voorkomen.

De nieuw gevonden entiteiten kunnen vervolgens door experts worden beoordeeld op relevantie en, indien relevant, kunnen deze entiteiten worden gebruikt als input voor het informatiesysteem genoemd in de "use case 1".

Resultaten en conclusies

Gedurende het project is samengewerkt binnen de gebruikersgroep Food Informatics en de ontwikkelingsgroepen binnen vl-e. Dit heeft geleid tot ontwikkeling van kennis over hoe het beste kan worden gezocht naar relevante informatie naar opkomende voedselveiligheidsgevaaren en het heeft geleid tot de ontwikkeling van software tools, waardoor deze kennis ook toegepast kan worden. De ontwikkelde ontologie is gebruikt bij het maken van queries voor het gericht zoeken naar informatie in document collecties. Ook is een tool ontwikkeld en getest waardoor het

mogelijk is om nieuwe entiteiten te extraheren uit documenten.

De ontwikkelde kennis in combinatie met de ontwikkelde software tools blijken een belangrijke verbetering te zijn ten opzichte van de uitgangssituatie bij aanvang van het project.

Een ander resultaat is dat binnen het project is aangetoond dat door samenwerking tussen de vele academische en niet-academische partners tot concrete resultaten is gekomen en dat door deze samenwerking een brug is geslagen tussen state of the art onderzoek en technologische innovatie in het voedingsonderzoek.

Gegeven de goede resultaten op gebied van kennis, software tools en samenwerking binnen vl-e zal TNO in 2009 doorgaan met de ontwikkeling, implementatie en toepassing van het informatiesysteem ten behoeve van de voedselveiligheid. De behoefte aan dergelijke systemen is ook nu nog steeds aanwezig. Recente voedselveiligheidsincidenten, zoals "melamine in huisdierenvoeding (2007)" en "melamine in melkpoeder (2008)" illustreren dat continue alertheid is geboden om de voedselveiligheid te kunnen waarborgen.

Joint International Agricultural Conference
6, 7 and 8 July 2009, Wageningen