

Doorontwikkeling zoekmachine Overheid.nl

Projectplan in het kader van de burgerthesaurus

In opdracht van:
Advies Overheid.nl

Zenc, mei 2004

Auteurs:
Ko Mies
John Oldenhuizing
Roel Titulaer

Management Samenvatting

Achtergrond/aanleiding

Burgers gaan steeds meer gebruik maken van internet als informatiebron, ook voor overheidsinformatie. De ontsluiting via Overheid.nl is daarbij een belangrijk hulpmiddel. Gelijk met deze toename van het gebruik van internet door burgers neemt ook de hoeveelheid informatie die de overheid op internet wordt aangeboden sterk toe. Tevens ontwikkelt de zoek- en ontsluitingstechnologie op internet zich steeds verder en groeit het wensenpatroon met betrekking tot de manier waarop men informatie wenst te kunnen vinden mee. Tot slot blijkt de authenticiteit van bronnen (is dit informatie van de overheid of van een commerciële partij) voor veel internetters een groeiend probleem te zijn.

Deze drie ontwikkelingen:

- toename gebruik internet;
- toename aanbod online informatie van de overheid;
- ontwikkeling en acceptatie door het publiek van nieuwe technologieën
- toenemende onduidelijkheid t.a.v. authenticiteit van informatie

gecombineerd met de opgedane ervaringen met Overheid.nl tot nu toe, leiden tot de noodzaak om de centrale ontsluiting van online overheidsinformatie in het algemeen, en die via Overheid.nl in het bijzonder sterk te verbeteren. Deze noodzaak is extra sterk aanwezig omdat particuliere zoekmachines en portals overheidsinformatie niet voldoende volledig ontsluiten en aan gebruikers zelf de verantwoordelijkheid over laten om te bepalen of de gevonden informatie wel echt van de overheid is of niet.

Het voorliggende plan verschaft inzicht in de stappen die gezet moeten worden in de periode tot eind 2005/begin 2006 (en de jaren daarna) om met een sterk verbeterde ontsluiting van overheidsinformatie de dienstverlening aan burgers en bedrijven te verbeteren. Een specifiek voor de overheid ontwikkeld zoekmechanisme is noodzakelijk om goede en betrouwbare ontsluiting van (alle) overheidsinformatie en dienstverlening mogelijk te maken. Het brede domein van de overheid wordt onvoldoende ontsloten door marktpartijen, die daar voor zover het gebeurt vaak ook nog commerciële bedoelingen bij hebben. Algemene zoekmachines vinden vaak maar 30 tot 40 procent van de pagina's van een website en ze doorzoeken niet alle websites. Daarnaast is ook het kunnen aanbieden van (het gebruik van) de (zelf ontwikkelde) zoekmachine aan andere partijen (zowel overheid als niet overheidspartijen) een belangrijk perspectief om de ontsluiting van overheidsinformatie een extra impuls te geven.

De zoekmachine is een concreet toegepast (deel)project, waarvan (gezien bovenstaande) maatschappelijk(e) nut en noodzaak volledig worden onderschreven. Het innovatieve element van de zoekmachine zit in de breedte van de toepassing (organisatiekundig): de ontsluiting van overheidsinformatie en -diensten op deze manier is nog niet eerder gedaan. Het aansluiten van de zoektermen van de zoekmachine bij het begrippenkader van de gebruiker (de burger) is daarbij een belangrijk innovatief element. De zoekmachine dient gebruik te maken van bestaande technieken en is daarmee niet gericht op technische innovatie. Door de concrete toepassingsmogelijkheden van de zoekmachine in het brede overheidsdomein wordt een relatief hoge impact op de maatschappelijke innovatie bereikt.

Probleemstelling

De huidige ontsluiting via Overheid.nl levert langzamerhand, als gevolg van de vier geschetste recente ontwikkelingen, onvoldoende volledige toegang tot overheidsinformatie en -diensten en er zijn onvoldoende betrouwbare alternatieven voor handen voor de burger. De vindbaarheid (en daarmee de kenbaarheid) voor burgers van overheidsinformatie en -diensten is daardoor nog sterk te verbeteren. Ook levert de huidige zoekmachine een onvoldoende duidelijke presentatie van de gevonden informatie, waardoor de hanteerbaarheid onvoldoende is. Het aanbod van informatie en diensten zal daarnaast in de komende jaren sterk stijgen, doordat steeds meer overheidsorganisaties (conform de bestuurlijk geformuleerde doelstellingen) hun informatie en diensten via internet gaan aanbieden. De hanteerbaarheid (maar ook de vindbaarheid) zal daardoor nog verder onder druk komen te staan.

Aan de vraagkant gaan burgers steeds hogere eisen stellen aan de ontsluiting omdat in het dagelijks gebruik het internet een steeds beter gebruikt kanaal voor informatievoorziening en dienstverlening wordt. Ook de technische mogelijkheden zijn in hoog tempo in ontwikkeling. De huidige zoekmachine is om deze redenen toe aan een nieuwe impuls. De doorontwikkeling van de bestaande ontsluitingsmechanismen levert een product op dat mogelijk ook door andere organisaties (zowel overheid als niet overheidsorganisaties) kan worden gebruikt (denk aan uitgevers, informatiebank-beheerders, gemeenten en andere overheden, etc.).

Doelstellingen

De doelstellingen van het (deel)project zijn, afgeleid van aanleiding en probleemstelling:

- Voor burgers (en bedrijven) verbeteren van de toegankelijkheid (vindbaarheid) van informatie over overheidsdiensten;
- Voor burgers (en bedrijven) vergroten van de hanteerbaarheid van gevonden informatie over overheidsdiensten;
- Waar mogelijk stimuleren van de beschikbaarheid van overheidsinformatie.

Deze doelstellingen ondersteunen de volgende doelen uit het Actieplan Andere Overheid:

- Het verbeteren van de kwaliteit van de publieke dienstverlening van de overheid.
- Het transparanter maken van de overheid.

Kenmerken overheidsinformatie

Op hoofdlijnen geven we hier de kenmerken van overheidsinformatie die voor de zoekmachine van belang zijn:

- heterogeniteit van bronnen
- heterogeniteit van domeinen
- ongestructureerde gegevens
- weinig invloed op content
- context niet altijd duidelijk

Door deze kenmerken is niet één technologie of mechanisme beschikbaar dat alles oplost. Een goede afweging van mogelijkheden en voor- en nadelen is noodzakelijk en zal moeten worden gebaseerd op gedegen nader onderzoek.

Uitgangspunten

De ontwikkeling van de vernieuwde zoekmachine wordt gebaseerd op een aantal technische, organisationele en financiële uitgangspunten:

- 'proven technology': het gaat niet om technologische, maar om maatschappelijke innovatie;
- aansluiting bij bestaande initiatieven;
- minimale beheerlast voor Advies Overheid.nl;
- meer bronnen ontsluiten en bestaande bronnen vollediger ontsluiten;
- vraaggerichter ontsluiten;
- dienstverlening volgens 'no wrong door'-principe;
- zo mogelijk voor gebruik beschikbaar stellen aan zowel overheids- als niet-overheidsorganisaties.

Relevante ontwikkelingen

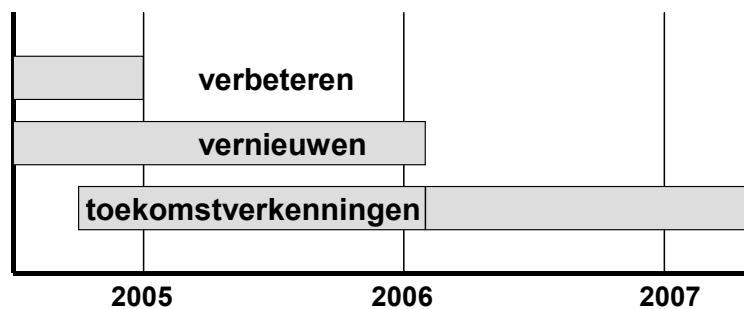
Een aantal ontwikkelingen op het gebied van zoekmachines is voor de Nederlandse situatie van belang. We hanteren daarbij de volgende indeling:

- keyword handling
- zoektechnologieën
- andere ontsluitingsmechanismen
- taxonomieën/thesauri/semantische indelingen
- marketing/distributie

Aanpak/activiteiten

De invulling van de functionaliteiten, positionering en organisatie van de nieuwe zoekmachine is veelomvattend. Om deze reden volgt het (deel)project Doorontwikkeling zoekmachine drie ontwikkellijnen:

- Verbeteren (Quick Wins):
In deze ontwikkellijijn worden snelle, gemakkelijk te realiseren veranderingen in de huidige zoekmachine aangebracht, waardoor al op de korte termijn voor de burger verbeteringen merkbaar zijn.
- Vernieuwen:
In deze meer structurele ontwikkellijijn worden de verschillende aspecten van de zoekmachine nader onderzocht en afgewogen in een toekomstvisie. Vanuit de ontwikkelde visie worden de specificaties vastgesteld voor en de realisatie gewaarborgd van de nieuwe zoekmachine.
- Toekomstverkenningen.
Omdat de ontwikkelingen wat betreft de technologie en de behoeften van burgers niet stil staan zal parallel aan de realisatie en de operationalisatie van de nieuwe zoekmachine in de ontwikkellijijn Toekomstverkenningen verder gezocht worden naar mogelijke vernieuwingen voor de verdere toekomst.



Begroting/planning

In de begroting/planning van het (deel)project is uitgegaan van het volledige activiteitenplan. Afhankelijk van keuzes die gemaakt worden tijdens het (deel)project zal de realisatie zich bewegen binnen een bandbreedte ten opzichte van de aangegeven verwachte kosten. De invulling van de personele kosten is mede afhankelijk van keuzes t.a.v. zelfdoen, uitbesteden of samenwerking. Vooralsnog zijn wij uitgegaan van de aanname dat 1 projectmedewerker gedurende anderhalf jaar benodigd is.

In het (deel)project moet ook rekening worden gehouden met kosten van marketing en distributie. Maar omdat het de voorkeur verdient deze kosten te delen met andere (deel)projecten zijn ze voor dit (deel)project niet nader gespecificeerd. Bij de uitvoering zal, met name in de doorontwikkeling naar andere domeinen, gezocht moeten worden naar een goede verdeling van de hiervoor benodigde kosten.

Onderstaand geven wij een overzicht van kosten per kostengroep en vervolgens een verdeling per jaar.

Kostengroep	Verwacht €
bouw	345,000
management	92,000
onderzoek	193,500
ontwerp	105,000
software	350,000
voorbereiding	67,500
Totaal	1,153,000

Personeel	157,500
Eindtotaal	1,310,500

	Totaal
2004	533,000
2005	777,500
Eindtotaal	1,310,500

Resultaten

In de periode medio 2004 – 2007 worden drie ontwikkelijnen gevolgd, met verschillende resultaten, activiteiten en dynamiek:

- **Verbeteren**

De korte termijn ontwikkeling, zal eind 2004 afgerond worden en leidt tot een verbeterde zoekmachine op basis van de huidige technologie. Met name de kwaliteit en presentatie van de zoekresultaten worden verbeterd.

Resultaten:

- Versnelling zoekproces burgers

- Kwalitatief betere zoekresultaten

- **Vernieuwen**

Van medio 2004 tot eind 2005/begin 2006, leidt tot een sterk verbeterde zoekmachine (inclusief beheerorganisatie en vernieuwde interface) waarin verschillende ontsluitingsmechanismen zijn verwerkt, met vooralsnog een focus op dienstverlening vanuit de overheid.

Resultaten:

- Visie vastgesteld
- Keuze van nieuwe (proven) technologieën en navigatiemethoden bepaald
- Nieuwe zoekmachine gerealiseerd
- Samenwerkende catalogi gerealiseerd
- Beheerorganisatie ingericht

- **Toekomstverkenningen** (nieuwe techiek, mogelijk nieuwe organisatie, mogelijk nieuw beleid)

Waarbij de zoekmachine kan worden ingezet voor alle domeinen van overheidsinformatie en waarmee Overheid.nl ook de daarop volgende jaren de aansluiting behoudt bij nationale behoeften en internationale ontwikkelingen. In 2006 worden een aantal experimenten met veelbelovende zoektechnologieën uitgevoerd, geëvalueerd en eventueel in de zoekmachine verwerkt.

Resultaten:

- Groot bereik burgers (doelstelling): 75 % (van de bezoekers)
- Alle bronnen ontsloten
- Breed bereik: dienstverlening, overheidsinformatie en organisaties
- Experimenten uitgevoerd en verwerkt
- Projectplan/visie vervolg na 2006/2007

Inhoudsopgave

Doorontwikkeling zoekmachine Overheid.nl	1
Management Samenvatting.....	2
Achtergrond/aanleiding	2
Probleemstelling.....	3
Doelstellingen	3
Kenmerken overheidsinformatie.....	3
Relevante ontwikkelingen.....	4
Aanpak/activiteiten	4
Begroting/planning	5
Resultaten.....	5
1. Projectachtergrond.....	9
1.1. Inleiding	9
1.2. Huidige situatie	10
1.3. Opzet projectplan.....	11
2. Projectdefinitie.....	12
2.1. Doelstellingen	12
2.2. Bereik	12
2.3. Resultaten	13
2.4. Uitsluiting	15
2.5. Beperkingen	15
2.6. Relaties met andere (deel)projecten	16
3. Business case	17
4. Mogelijke technologieën en ontsluitingsmechanismen	19
4.1. Inleiding	19
4.2. Kenmerken overheidsinformatie en uitgangspunten.....	19
4.3. Relevante ontwikkelingen	20
4.4. Samenvattend.....	25
5. Kwaliteitsverwachtingen	26
6. Acceptatiecriteria.....	29
7. Organisatiestructuur	31
8. Risicologboek.....	32
9. Projectaanpak	34
9.1. Inleiding	34
9.2. Verbeteren: medio 2004 – eind 2004	34
9.3. Vernieuwen: medio 2004 – eind 2005.....	35
9.4. Toekomstverkenningen: 2005 tot 2007	36
10. Begroting / Planning	37

Bijlage 1: Bronnenoverzicht.....	39
Bijlage 2 : Quick scan technologieën en ontsluitingsmechanismen en aanbevelingen onderzoek	40
Bijlage 3: Query-analyse en Quick wins MZD.....	46
Keyword handling	46
Overige aanbevelingen	47
Bijlage 4: Detailbegroting en -planning	48

1. Projectachtergrond

1.1. Inleiding

Burgers gaan steeds meer gebruik maken van internet als informatiebron, ook voor overheidsinformatie. De verbetering van de ontsluiting via Overheid.nl is daarbij een belangrijk hulpmiddel. De huidige ontsluiting levert echter voor de moderne burger onvoldoende bevredigende antwoorden op, zowel wat betreft directe informatie als wat betreft (doorverwijzigings)informatie over dienstverlening door de overheid. Aan de andere kant wil de overheid haar dienstverlening en informatievoorziening steeds meer via het internet gaan aanbieden en heeft daarvoor ook concrete doelen geformuleerd. Het is om deze redenen dat de (door)ontwikkeling van de zoekmachine van Overheid.nl noodzakelijk is in het teken van de verbetering van de elektronische dienstverlening door de overheid aan burgers en bedrijven.

Voor de (door)ontwikkeling van de zoekmachine is door het programmabureau Advies Overheid.nl ¹ de opdracht gegeven een (deel)projectplan op te stellen. Het doel van dit plan is inzicht te verschaffen in de stappen die gezet moeten worden in de periode tot eind 2005/begin 2006 (en de jaren daarna) om met behulp van de zoekmachine de dienstverlening aan burgers en bedrijven te verbeteren. Een specifiek voor de overheid ontwikkelde zoekmachine is noodzakelijk om goede en betrouwbare ontsluiting van (alle) overheidsinformatie en dienstverlening mogelijk te maken. Algemene zoekmachines vinden vaak maar 30 tot 40 procent van de pagina 's van een website en ze doorzoeken niet alle websites. Het brede domein van de overheid wordt daarom onvoldoende ontsloten door marktpartijen, die daar voor zover het gebeurt vaak ook nog commerciële bedoelingen bij hebben. Daarnaast is ook het kunnen aanbieden van (het gebruik van) de (zelf ontwikkelde) zoekmachine aan andere partijen (zowel overheid als niet overheidspartijen) een belangrijk perspectief om de ontsluiting van overheidsinformatie een extra impuls te geven.

De zoekmachine is een concreet toegepast (deel)project, waarvan (gezien bovenstaande) maatschappelijk(e) nut en noodzaak volledig worden onderschreven. Het innovatieve element van de zoekmachine zit in de breedte van de toepassing (organisatiekundig): de ontsluiting van overheidsinformatie en -diensten op deze manier is nog niet eerder gedaan. Het aansluiten van de zoektermen van de zoekmachine bij het begrippenkader van de gebruiker (de burger) is daarbij een belangrijk innovatief element. De zoekmachine dient gebruik te maken van bestaande technieken en is daarmee niet gericht op technische innovatie. Door de concrete toepassingsmogelijkheden van de zoekmachine in het brede overheidsdomein wordt een relatief hoge impact op de maatschappelijke innovatie bereikt.

De ontsluiting via Overheid.nl is één van de bouwstenen die de burger ten dienste staan om van de dienstverlening en informatievoorziening van de overheid gebruik te maken. Daarmee is ook meteen gezegd dat bij de doorontwikkeling van de bestaande zoekmachine rekening gehouden moet worden met de andere bouwstenen die in ontwikkeling zijn. We denken daarbij aan de ontsluiting van de productinformatie via verschillende productencatalogi, aan de invoering van metadatering van

¹ Advies Overheid.nl <<http://www.advies.overheid.nl/>>

overheidsinformatie, etc.. Dit zijn allemaal projecten die via verschillende wegen leiden tot hetzelfde doel: het leveren aan burgers (en bedrijven) van diensten en overheidsinformatie.

Bij de doorontwikkeling van de zoekmachine zullen we intensief zoeken naar mogelijkheden om te komen tot synergetische effecten met de andere lopende (deel)projecten. In ieder geval zullen we er naar streven dat de technologie koppelingen en uitwisselingen mogelijk maakt.

1.2. Huidige situatie

Het (deel)project omvat de dóórontwikkeling van de zoekmachine. De basis is al gelegd in 1999 met de website Overheid.nl. Op Overheid.nl hebben burgers vanaf het begin op basis van een startpagina principe online informatie van overheden kunnen vinden. Deze website is steeds verder uitgebreid. Tegenwoordig kunnen bezoekers informatie zoeken op basis van:

- **doelgroep:** particulier, ondernemers of organisaties, politici of ambtenaren, jongeren, guests;
- **thema:** Landbouw en visserij, Natuur en milieu, Onderwijs, Onderzoek en wetenschap, Ruimtelijke ordening en grote stedenbeleid, Veiligheid en defensie, Verkeer en waterstaat, Werk en sociale zekerheid, etc.
- **geografische positie:** op basis van postcode
- **type gezochte informatie:** contactgegevens, wet- en regelgeving, (officiële) publicaties, productinfo, etc.

Sinds maart 2003 beschikt Overheid.nl ook over een zoekmachine. Deze zoekmachine doorzoekt (zo goed mogelijk) meer dan 1.000 websites en kan bevroegd worden op:

- zoektermen;
- type organisatie waarvan men informatie zoekt;
- postcode;
- ouderdom van de informatie.

Vrijwel alle onderdelen van Overheid.nl zijn tijdens hun ontwikkeling en/of daarna onderworpen aan gebruikersonderzoeken met als doel gebruikersgemak te optimaliseren. Deze gebruikersonderzoeken vormen mede de basis bij de uitwerking van dit projectplan. De verdere doorontwikkeling gebeurt daardoor met bestaande ervaringen, naast aspecten als nieuwe technische mogelijkheden, relaties met andere (deel)projecten en natuurlijk de uitgangspunten van Andere Overheid.

1.3. Opzet projectplan

In dit projectplan worden alle elementen beschreven die van belang zijn voor enerzijds de besluitvorming over de aanpak en anderzijds voor de uitvoering van de gewenste activiteiten. We gebruiken bij de beschrijving op hoofdlijnen een indeling conform Prince2 om nauw aan te sluiten bij de eisen die gesteld worden aan een plan voor de uitvoering. In hoofdstuk 2 beschrijven we de doelstellingen en grenzen van het (deel)project. Hoofdstuk 3 bevat de business case. In hoofdstuk 4 worden kenmerken van overheidsinformatie beschreven, de uitgangspunten en relevante ontwikkelingen voor de nieuwe zoekmachine. In de hoofdstukken 5 en 6 geven we inzicht in respectievelijk de kwaliteitsverwachtingen en de acceptatiecriteria voor de drie ontwikkellijnen. Hoofdstuk 7 geeft een beeld van de gewenste organisatiestructuur op hoofdlijnen. De risico's zijn opgesomd in hoofdstuk 8. De stappen in de aanpak zijn op hoofdlijnen in hoofdstuk 9 weergegeven. Op basis hiervan zijn ook de planning en begroting opgesteld. De hoofdlijn van de begroting is opgenomen in hoofdstuk 10. De detailplanning en -begroting is opgenomen als een losse bijlage bij dit document.

In de bijlagen 1, 2 en 3 zijn de gebruikte bronnen bij het schrijven van het (deel)projectplan weergegeven, wordt een analyse gedaan van te onderzoeken/te gebruiken technologieën en worden aanbevelingen gedaan voor quick wins.

2. Projectdefinitie

2.1. Doelstellingen

Doelstellingen van het (deel)project

- Voor burgers (en bedrijven) verbeteren van de toegankelijkheid (vindbaarheid) van informatie over overheidsdiensten;
- Voor burgers (en bedrijven) vergroten van de volledigheid van gevonden informatie over overheidsdiensten;
- Waar mogelijk stimuleren van de beschikbaarheid van overheidsinformatie.

Deze doelstellingen ondersteunen een aantal doelen:

- Het verbeteren van de kwaliteit van de publieke dienstverlening van de overheid.
- Het transparanter maken van de overheid.

2.2. Bereik

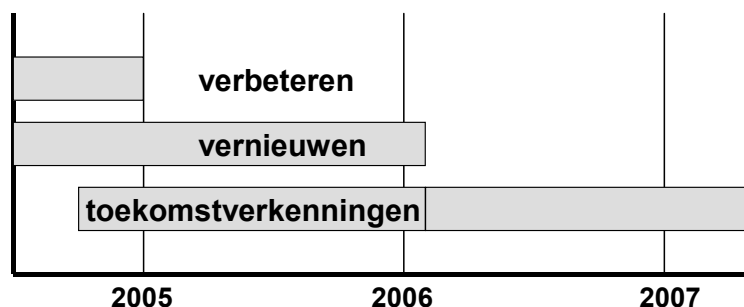
Bij dit (deel)project zijn verschillende partijen betrokken. Onderstaand geven we een niet limitatieve opsomming hiervan:

- Advies Overheid.nl: als opdrachtgever en positiebepaler voor de zoekmachine;
- Gemeentelijke overheden: als leverancier van (informatie over) diensten, als beheerder van eigen websites;
- Provinciale overheden: als leverancier van (informatie over) diensten, als beheerder van eigen websites;
- Rijksoverheid: als leverancier van (informatie over) diensten, als beheerder van eigen websites (informatiebronnen), als opdrachtgever (Min BZK) voor het (deel)project;
- Waterschappen: als leverancier van (informatie over) diensten, als beheerder van eigen websites;
- ZBO's: als leverancier van (informatie over) diensten, als beheerder van eigen websites

Voor al deze betrokken organisaties geldt dat hun websites met een eigen zoekmachine zijn uitgerust. Dit kan mogelijkwijs (afhankelijk van in het traject te maken keuzes) veranderen met de ontwikkeling van de nieuwe zoekmachine.

2.3. Resultaten

In de periode medio 2004 – 2007 worden drie ontwikkellijnen (verbeteren, vernieuwen en toekomstverkenningen) gevolgd, met verschillende resultaten, activiteiten en dynamiek. De ontwikkellijnen lopen voor een deel parallel, maar zullen elkaar aanvullende resultaten opleveren. In schema:



- **Verbeteren**, de korte termijn ontwikkeling, zal eind 2004 afgerond worden en leidt tot een verbeterde zoekmachine op basis van de huidige technologie. Met name de kwaliteit en presentatie van de zoekresultaten worden verbeterd.

Syndicatie-mogelijkheden (voor web developers)
<ul style="list-style-type: none">• HTML-code zoekbox en banner aanbieden voor gebruik op andere sites
Quick wins (zie bijlage 3)
<ul style="list-style-type: none">• Afvangen veel gebruikte zoekwoorden• Weergave meest gebruikte zoekwoorden• Verbeterde presentatie zoekresultaten• Verbeterde relevantie zoekresultaten• Koppeling zoekresultaten aan thematische startpagina's• Korte termijn mogelijkheden webmarketing/distributie• Beheermodule
Instrument voor het meten van (be)zoekgedrag
Koppeling of integratie wet- en regelgeving
Minder spider-problemen

- **Vernieuwen**, van medio 2004 tot eind 2005/begin 2006, leidt tot een sterk verbeterde zoekmachine (inclusief beheerorganisatie en vernieuwde interface) waarin verschillende ontsluitingsmechanismen zijn verwerkt, met vooralsnog een focus op dienstverlening vanuit de overheid.

Visie-ontwikkeling
<ul style="list-style-type: none">• Indeling in (meer) doelgroepen, rollen of profielen, voorselectie content-typen• Hoeveelheid (eigen) content• Mate van samenwerking (ministeries, portals, derden)• Links naar derden/bronnen buiten de overheid• Personalisering• Multi-channeling (UMTS, GPRS, GSM, iMode)

<ul style="list-style-type: none"> • Onderscheid in informatietypen: op de zoekinterface of op de resultatenpagina
<p>Toepasbaarheid zoektechnologieën op basis van content-analyse</p> <ul style="list-style-type: none"> • Gestructureerde/ongestructureerde informatie • Domeinen • Databases • Activiteitenindices • Bronnen met niet-tekst (afbeeldingen, audio, video, etc.)
<p>De interesse voor een zoekmachine als 'shared service' / distributiemogelijkheden</p> <ul style="list-style-type: none"> • Kostenmodel • Functionaliteiten huidige zoekmachines • Functionaliteiten nieuwe zoekmachine • Mogelijkheden distributie • Mogelijke partners die content aanleveren
<p>De zoekmachines op buitenlandse overheidsportals en nationale sites</p> <ul style="list-style-type: none"> • Toepasbaarheid • Licentiemodellen/Kosten • Open-source oplossingen • Gebruik taxonomie en thesaurus
<p>Gebruik van natuurlijke taaltechnologie voor ontsluiting van catalogi en almanak</p> <ul style="list-style-type: none"> • Toepasbaarheid • Kosten integratie • Gebruik burgerthesaurus (spreektaal → specifieke terminologie / jargon)
<p>Toepasbaarheid andere ontsluitingsmechanismen op basis van contentanalyse</p> <ul style="list-style-type: none"> • Indeling doelgroepen • GIS • Thema's • Narrow down mechanismen/related items • Personalisering (met name alerts) • Externe contentleveranciers
<p>De bruikbaarheid en beschikbaarheid van Nederlandstalige thesauri</p> <ul style="list-style-type: none"> • Huidige initiatieven • Toepasbaarheid/bruikbaarheid • Eigendom van een publieke organisatie • Beheer • Kosten integratie
<p>Een taxonomie t.b.v. de creatie trefwoordenboom en automatische dossiervorming</p> <ul style="list-style-type: none"> • Bestaande initiatieven / taxonomieën • Toepasbaarheid/bruikbaarheid • Wijze van termen verzamelen (top down en/of bottom-up) • Wijze van categorisatie (handmatig of automatisch) • Gebruik in verschillende ontsluitingsmechanismen • Onderhoud taxonomie • Kosten integratie • Afspraken en regels voor toekennen van termen en introductie nieuwe termen
<p>Koppeling of integratie productencatalogi gemeenten, provincies, waterschappen, Rijk</p>

- **Toekomstverkenningen**, waarbij de zoekmachine kan worden ingezet voor alle domeinen van overheidsinformatie en waarmee Overheid.nl ook de daarop volgende jaren de aansluiting behoudt bij nationale behoeften en internationale ontwikkelingen. In 2006 kunnen met veelbelovende zoektechnologieën uitgevoerd, geëvalueerd en eventueel in de zoekmachine verwerkt. Uiteraard zijn dit nog optionele activiteiten, maar wij willen ze hier toch alvast noemen.

Experiment met community-vormen, bijv. een forum voor gepersonaliseerde antwoorden
Experiment met een 'topic map' of ontologie voor een bepaald domein
Experiment met 'social search'
Ontwikkeling van een projectplan met vervolgstappen en visie
Inventarisatie beschikbaarheid nieuwe bronnen

2.4. Uitsluiting

Om te voorkomen dat (deel)projecten elkaar overlappen, zal in dit (deel)project niet losstaand worden gewerkt aan:

- a) realiseren van transacties;
- b) verbeteren van samenwerkende productencatalogi;
- c) realiseren van metadata;
- d) verbeteren van de kwaliteit van informatiebronnen (websites van andere organisaties dan Advies Overheid.nl);

Wel zullen in dit (deel)project de eventuele samenhangen worden bewaakt en waar van toepassing worden afgestemd met de (deel)projecten die bovenstaande doelen beogen.

Toegankelijkheid kent een aantal aspecten: vindbaarheid en kenbaarheid, beschikbaarheid, hanteerbaarheid, betaalbaarheid, betrouwbaarheid en duidelijkheid. De zoekmachine verbeterd de aspecten vindbaarheid en kenbaarheid en hanteerbaarheid. Dit (deel)project heeft geen invloed op de ander aspecten, beschikbaarheid, betaalbaarheid, betrouwbaarheid en duidelijkheid, van de ontsloten informatie.

2.5. Beperkingen

Van invloed op de uitkomst van het (deel)project is uiteindelijk de kwaliteit van de onderliggende informatiebronnen. Op de kwaliteit hiervan heeft Advies Overheid.nl geen invloed. We denken hierbij aan de begrijpelijkheid van aangeboden informatie, de actualiteit hiervan, het aanbod van informatie in andere dan de Nederlandse taal, de kosten die voor de informatie in rekening worden, etc.. Dit alles is de verantwoordelijkheid van de bronbeheerders en dienstenleveranciers. Overheid.nl heeft uiteraard wel een belangrijke rol in het ontsluiten van de juiste informatie, het garanderen van de authenticiteit en het stimuleren van de kwaliteit van de informatie voor de juiste doelgroep.

2.6. Relaties met andere (deel)projecten

We noemen in dit stadium de belangrijkste gerelateerde (deel)projecten, maar kunnen daarin niet limitatief zijn, omdat gedurende de looptijd ongetwijfeld nieuwe (deel)projecten worden opgestart en andere beëindigd. In de aanpak zijn acties opgenomen om andere (deel)projecten te inventariseren en hiermee af te stemmen.

Project metadatering

Dit project ontwikkelt een open standaard voor publicatie en metadatering van overheidsinformatie. De verwachting is dat eind 2004 deze standaard wordt vastgesteld en dat in de loop van 2005 de eerste sites de standaard geïmplementeerd zullen hebben. De zoekmachine zal dan ook van de metadata volgens deze standaard gebruik kunnen maken om deze sites te doorzoeken.

Project databanken

Veel informatie van de overheid is opgeslagen in databanken. Sommigen van deze databanken staan al online. Advies Overheid.nl start een apart project om meer databanken online te brengen.

(Deel)project samenwerkende catalogi

Dit (deel)project heeft als doel om het voor gebruikers mogelijk te maken om in de verschillende productencatalogi van gemeenten, provincies, waterschappen en Rijk heen naar productinformatie te kunnen zoeken alsof het één loket lijkt te zijn.

Project voorgenomen besluiten en vergunningen

Dit project streeft er naar om overheden voorgenomen besluiten en verstrekte vergunningen te laten publiceren op het Internet. Deze publicaties zullen voldoen aan de hierboven genoemde publicatie- en metadatastandaard. In de loop van 2005 zullen de eerste structurele publicaties op sites van overheden plaatsvinden. Deze publicaties zullen ook door de zoekmachine gevonden moeten gaan worden.

Project decentrale wettenbanken

Dit project streeft ernaar om verordeningen van gemeenten, waterschappen en gemeenten in geconsolideerde overzichten online te krijgen, net als dat al gerealiseerd is voor de nationale wet- en regelgeving². Ook deze publicaties zullen zich conformeren aan de standaardmetadatering zoals hierboven genoemd. De eerste wettenbanken komen al in 2004 online.

Wet- en regelgeving op Overheid.nl

Momenteel is Wet- en Regelgeving van Overheid.nl nog niet doorzoekbaar vanuit de centrale zoekmachine van Overheid.nl. De applicatie zal worden aangepast zodat dat als nog kan. De schatting is dat vanaf juli 2005 de koppeling gerealiseerd kan worden.

Officiële Publicaties

De huidige Officiële Publicaties van Overheid.nl³ zijn niet gekoppeld aan de zoekmachine. Momenteel wordt gewerkt aan een nieuwe versie van Officiële Publicaties, compleet met standaard metadata. De nieuwe versie van de zoekmachine is waarschijnlijk klaar eind 2005 en kan dan worden aangesloten op de zoekmachine.

² Wet- en regelgeving <<http://www.wetten.nl/>>

³ Officiële Publicaties <<http://www.overheid.nl/op>>

3. Business case

De overheid stelt zich ten doel via ICT de kwaliteit van de publieke dienstverlening te verhogen en de elektronische toegankelijkheid van overheidsinformatie voor de burger te vergroten.

“In 2007 zal alle overheidsinformatie waarvan bekendmaking wettelijk voorgeschreven is, ook via Internet toegankelijk zijn. Niet alleen via gerichte vragen, maar automatisch: actieve openbaarmaking. Een gebruiksvriendelijke zoekmachine gaat burgers helpen om de gewenste informatie op te zoeken.”⁴

Minister de Graaf noemt hier een aantal aspecten van toegankelijkheid waarvoor een zoekmachine oplossingen biedt.

- 1) **Kenbaarheid en vindbaarheid:** Een zoekmachine moet de gebruiker informeren welke informatie bij welk overheidsorgaan online beschikbaar is en hoe en waar die kan worden gevonden, met andere woorden het aanbod van informatie en diensten transparant maken.
- 2) **Hanteerbaarheid:** Een zoekmachine moet de gebruiker helpen zijn weg te vinden in de groeiende hoeveelheid informatie, met andere woorden het zoekproces faciliteren.

De huidige zoekmachine op Overheid.nl voldoet op dit moment onvoldoende aan de eisen van deze tijd. De kwaliteit en presentatie van de zoekresultaten laat steeds meer te wensen over. Uit het gebruikersonderzoek van december 2003 blijkt dat burgers niet geheel tevreden zijn over de huidige zoekmachine. Vooral de relevantie en presentatie van de zoekresultaten laten te wensen over. Bovendien zullen de komende jaren meer bronnen worden ontsloten (zie paragraaf 2.6), waardoor extra eisen aan de zoekmachine worden gesteld. Voor de burger zijn er op geen alternatieven (andere zoekmachines gericht op het terrein van de overheid) beschikbaar voor het brede domein van overheidsinformatie en -dienstverlening.

De door dit (deel)project vernieuwde zoekmachine zal in eerste instantie worden gebruikt om de dienstverlening te verbeteren door het koppelen of integreren van productencatalogi van gemeenten, provincies, waterschappen en het Rijk. Indien een burger (als klant) een product in de verkeerde catalogus zoekt, wordt hij via de zoekmachine doorverwezen naar de juiste organisatie ('no wrong door'). Bovendien kunnen productbeschrijvingen automatisch verrijkt worden met informatie van gerelateerde producten van andere overheidsorganisaties.

De zoekmachine van Overheid.nl kan mogelijk een generieke component of bouwsteen zijn voor het zoeken van en ontsluiten van overheidsinformatie en -diensten. Andere overheidsorganisaties kunnen er voor kiezen om niet zelf te investeren in zoektechnologie, maar de zoekmachine van Overheid te gebruiken voor het ontsluiten van hun eigen informatie. Dit geldt uiteraard ook voor niet-overheidsorganisaties die geïnteresseerd zijn in de ontsluiting van overheidsinformatie (uitgevers, informatiebank-beheerders, etc.).

⁴ Actieprogramma 'Andere Overheid' en toespraak minister de Graaf bij de afsluiting van de kennismarkt 'Elektronische Andere Overheid', 19-2-2004.

We kijken uiteraard niet alleen naar de zoekmachine, maar ook naar andere gebruiksvriendelijke manieren om overheidsinformatie en -diensten te ontsluiten (zie hoofdstuk 4). Elke applicatie (bijv. wettenbank, almanak, etc.) en website heeft een eigen ontsluitingsmechanisme. **De zoekmachine is de enige mogelijkheid voor burgers om in alle bronnen tegelijkertijd te zoeken.** Het vormt de ondersteuning van het (virtuele) een-loket voor de overheid.

4. Mogelijke technologieën en ontsluitingsmechanismen

4.1. Inleiding

Om te komen tot de juiste keuzes voor een zoekmachine voor de Nederlandse situatie, dient een complexe afweging gemaakt te worden van verschillende mogelijke aspecten en mechanismen ook tegen de achtergrond van de al bestaande zoekmachine(s) en de ontwikkeling van andere componenten. Daarnaast zijn voor de doorontwikkeling de kenmerken van belang van de overheidsinformatie in Nederland en de uitgangspunten die worden gekozen voor en door Advies Overheid.nl. In dit hoofdstuk beschrijven we op hoofdlijnen de daarbij belangrijke ontwikkelingen, als illustratie van de in het (deel)project (in de ontwikkellijn Vernieuwen) uit te voeren onderzoeks- en visieontwikkelingsactiviteiten. Het is tevens een illustratie van de mogelijke functionaliteiten van de zoekmachine en context waarin deze moet kunnen functioneren.

4.2. Kenmerken overheidsinformatie en uitgangspunten

We realiseren ons dat we in dit bestek niet alomvattend kunnen zijn in de beschrijving van overheidsinformatie, maar op hoofdlijnen geven we hier de kenmerken die voor de zoekmachine van belang zijn:

- **heterogeniteit van bronnen:** In Nederland zijn talloze bronnen beschikbaar met verschillende soorten informatie. Niet alle informatie kan op dezelfde manier worden benaderd, ingedeeld en gepresenteerd.;
- **heterogeniteit van domeinen:** Het gaat om informatie over hele verschillende domeinen (een adressengids is iets heel anders dan een productencatalogus, bij de waterschappen gaat het over hele andere zaken dan bij het ministerie van Economische Zaken). Daardoor is informatie moeilijker te ontsluiten, dan in het geval van een enkele homogene bron (bijv. over belastingen);
- **ongestructureerde gegevens:** Gegevens zijn veelal ongestructureerd (bijvoorbeeld teksten zonder duidelijke structuur, websites). Dit maakt het zoeken en ophalen van goede resultaten moeilijker;
- **weinig invloed op content:** Overheid.nl heeft weinig invloed op het maken en redigeren van de informatie die de zoekmachine moet ontsluiten. Daardoor zal niet alle informatie op dezelfde manier gestructureerd of begrijpelijk zijn;
- **context niet altijd duidelijk:** Informatie kan in verschillende contexten verschillende betekenissen en verschillende relevantie hebben (de wet op de huursubsidie is voor juristen gesneden koek, maar voor (sommige) burgers vermoedelijk totaal onbegrijpelijk). “Goede” zoekresultaten hangen dus af van de doelgroep en de vraag.

Door deze kenmerken is niet één technologie of mechanisme beschikbaar dat alles oplost. Een goede afweging van mogelijkheden en voor- en nadelen is noodzakelijk en zal moeten worden gebaseerd op gedegen nader onderzoek (uit te voeren in de fase Vernieuwen). Besluitvorming hierover kan mede worden gebaseerd op technologie ontwikkelingen en vergelijkbare internationale ontwikkelingen, die wij ter illustratie in dit hoofdstuk kort beschrijven. We pretenderen geen volledigheid, maar geven een doorsnede van huidige ontwikkelingen en veel gebruikte technologieën.

Uitgangspunten

De ontwikkeling van de vernieuwde zoekmachine wordt gebaseerd op een aantal technische, organisationele en financiële uitgangspunten, die tijdens de visie-ontwikkeling aangescherpt en/of geformaliseerd kunnen worden.

- De doorontwikkeling van de zoekmachine wordt gebaseerd op 'proven technology': het gaat niet om technologische, maar om maatschappelijke innovatie.
- We zoeken daarom ook zoveel mogelijk aansluiting bij bestaande initiatieven.
- Op dit moment is het uitgangspunt dat de beheerlast voor Advies.overheid.nl minimaal moet zijn en de investeringen in lijn moeten zijn met de verwachte baten.
- De vernieuwde zoekmachine moet meer bronnen ontsluiten en bestaande bronnen vollediger ontsluiten.
- De zoekresultaten moeten beter aansluiten bij de "vraag" van de gebruiker.
- De zoekmachine heeft niet alleen een zoekfunctie, maar dient ook als "doorverwijzingsmechanisme". Bezoekers die een product in een verkeerde productencatalogus zoeken, worden via Overheid.nl naar de juiste productbeschrijving geleid, het zogenaamde 'no wrong door'-principe.
- De zoekmachine zal zo mogelijk voor gebruik beschikbaar gesteld worden aan andere organisaties. Zowel overheids- als niet-overheidsorganisaties kunnen met het gebruik van de vernieuwde zoekmachine bijdragen aan het verbeteren van de ontsluiting van overheidsinformatie. Uiteindelijk kan dat ook nog een kostenbesparing opleveren omdat niet elke (overheids)organisatie meer hoeft te investeren in een zoekmachine.

4.3. Relevante ontwikkelingen

Een aantal ontwikkelingen is voor de Nederlandse situatie van belang. Wij beschrijven ze hier globaal om aan te geven welke mogelijkheden er zijn voor de doorontwikkeling van de zoekmachine.

We hanteren daarbij de volgende indeling, waarvan ieder onderdeel toegelicht wordt en wordt geïllustreerd met buitenlandse voorbeelden:

- **keyword handling:** het afvangen van veelgebruikte zoekwoorden door het meten van zoekgedrag
- **zoektechnologieën:** technologieën die het zoekproces ondersteunen en de relevantie van resultaten verbeteren
- **andere ontsluitingsmechanismen:** mogelijkheden van ontsluiting buiten het "zoekscherf" om
- **taxonomieën/thesauri/semantische indelingen:** mechanismen om informatie beter doorzoekbaar te maken of te structureren, veelal door het vooraf uitwerken van een indeling en toekennen van verrijkte informatie aan documenten of aan zoekvragen.
- **marketing/distributie:** het breder beschikbaar maken van de zoekmachine, door functionaliteiten beschikbaar te maken voor anderen en de zichtbaarheid van de zoekmachine te vergroten

Keyword handling

Veel zoekmachines vangen veelgebruikte trefwoorden af, om gesponsorde resultaten te kunnen weergeven of om de gebruiker snel verder te helpen naar wat hij vermoedelijk zoekt. In de commerciële wereld wordt bijvoorbeeld veel voor termen als "autoverzekering" betaald. Gebaseerd op het zoekgedrag binnen Overheid.nl kan dit vertaald worden naar woorden als euthanasie of varkenspest, die kunnen verwijzen naar de gelijknamige dossiers bij respectievelijk de ministeries BZK en LNV. De bekende zoekmachine Google ⁵ (Amerikaanse versie) vangt ook gestructureerde vragen af (bijvoorbeeld adressen en nummers van patenten). In Nederland kan dit worden vertaald naar bijvoorbeeld nummers van kamerstukken. Een veel gebruikt mechanisme is ook het plaatsen van de meest gebruikte zoektermen op de homepage. Dit gebeurt nu al in het overheidsloket. De grotere portalen meten continu bezoek en zoekgedrag en verwerken dit in hun aanbod en presentatie.

Deze mechanismen worden als quick win gedefinieerd voor Overheid.nl en komen terug in de ontwikkellijn Verbeteren (zie ook bijlage 3).

Zoektechnologieën

De huidige zoekmethode is gebaseerd op trefwoorden en operatoren (booleaans en 'stemming'). Deze veelgebruikte zoekmethoden zijn goedkoop en eenvoudig te gebruiken. Meer geavanceerde zoekmethoden maken gebruik van een thesaurus. Op dit begrip komen we later nog terug. Door ook op synoniemen en homoniemen (vocabulaire) te zoeken wordt de 'recall' verbeterd. De gebruiker kan zoeken op gerelateerde termen, nauwere en bredere termen. Internationaal zien we een beperkte toepassing van andere zoekmechanismen:

- *Associatief zoeken*
- *Zoeken via concepten/patronen/natuurlijke taal*
- *Automatische classificatie*
- *Natuurlijke taaltechnologie*
- *Social search*

Tenslotte zijn er allerlei experimentele technologieën, waarmee bijvoorbeeld in beeldbestanden kan worden gezocht, of die gebaseerd zijn op rss-feeds.

De toepasbaarheid van bovenstaande technologieën zal moeten worden beoordeeld op basis van een analyse van de bronnen en content van Overheid.nl. De genoemde technologieën hebben vaak grote financiële consequenties. De toegevoegde waarde moet kunnen worden aangetoond. In de conclusies doen we een aanzet voor de te onderzoeken technologieën.

Andere ontsluitingsmechanismen

Informatietypen

Het ontsluitingsmechanisme is vaak afhankelijk van het informatietype. Overheidsorganisaties worden ontsloten door te zoeken op bestuursniveau (Rijk, provinciaal of lokaal ⁶), op macht (wetgevend, uitvoerend, rechterlijk ⁷), geografisch (postcode, GIS⁸), thematisch of alfabetisch. Producten kunnen stapsgewijs gevonden

⁵ Google <<http://www.google.com/>>

⁶ FirstGov, <<http://www.firstgov.gov/Agencies.shtml>>

⁷ FirstGov, <<http://www.firstgov.gov/Agencies.shtml>>

⁸ Bund.de, <<http://www.bund.de/Service/Geosuche-.5565.htm>>, Add-wijzer <<http://www.addwijzer.org/>>

worden: via levensgebeurtenis en criteria of via thema, vraagpatroon en doelgroep⁹. Vaak werken bepaalde mechanismen beter voor deelverzamelingen. Zo komt GIS op als ontsluitingsmechanisme voor bijvoorbeeld locatiegerelateerde zaken, zoals milieu-eisen en bestemmingsplannen. Veelgebruikte mechanismen zijn ook het verwijzen naar gerelateerde items (o.a.) Amazon en narrow down mechanismen, waarbij de gebruiker een aantal criteria opgeeft en daarmee de zoekresultaten afbakt (bijvoorbeeld Vacanselect¹⁰

De meest bekende zoekmachines maken het onderscheid in informatietypen (webpagina's, afbeeldingen, nieuwsberichten, etc.) op de interface en niet op de resultatenpagina¹¹. Ze laten dus de gebruiker kiezen naar welke soort informatie hij op zoek is. Overheid.nl maakt dit onderscheid momenteel niet. De resultaten uit het Overheidsloket (producten), de Staatsalmanak (personen en organisaties) en de Nieuwsbank (persberichten) en webpagina's worden op dit moment samengevoegd. De resultaten kunnen in de verbeterde zoekmachine per informatietype op de resultatenpagina geclusterd worden. Indien het aantal applicaties toeneemt, zal de resultatenpagina onoverzichtelijk worden. Aanpassing van de interface (tabbladen, links of radio buttons) is dan noodzakelijk.

Momenteel wordt alleen tekst ontsloten. In de toekomst moeten misschien ook afbeeldingen, audio, video, grafieken, tabellen en kaarten ontsloten worden. Dit moet uit een content-analyse blijken.

Personalisatie

Een portal biedt diverse mogelijkheden voor pro-actieve dienstverlening. Op AmsterdamMail en Google News Alerts¹² kan een gebruiker zich op nieuwsberichten abonneren. In Canada kan dit met updates van de websites¹³. In Singapore gaat men nog verder. Gebruikers van My.ecitizen¹⁴ kunnen voor diverse elektronische diensten notificaties ontvangen per SMS of e-mail. Bezoekers van Denmark.dk kunnen de zoekresultaten opslaan. Mijn.overheid.nl is nog niet mogelijk (geen eigen content) of ongewenst (taak van de markt).

Communityvormen

Via bijvoorbeeld een forum kunnen bezoekers elkaar helpen bij het vinden van informatie of antwoorden op vragen¹⁵. Op commerciële sites wordt dit middel bijvoorbeeld ingezet zodat gebruikers lijstjes van hun favoriete producten kunnen bijhouden (Amazon ListMania¹⁶). Veel gebruikt is ook het laten stemmen op de waarde van informatie door bezoekers (bijvoorbeeld de kwaliteit van links, reviews).

De meeste van deze technieken zijn relatief goedkoop. Daartegenover staat de mogelijkheid tot misbruik en de onmogelijkheid van het controleren op de kwaliteit van

⁹ VIND of OPUS, bijv. op websites van resp. gemeente Breukelen <<http://www.breukelen.nl/>> en gemeente Zoetermeer <<http://www.zoetermeer.nl/>>

¹⁰ Vacanselect – zoeken op bestemming <<http://www.vacanselect.nl/>>

¹¹ Nieuw-Zeeland <<http://www.govt.nz/>>

¹² AmsterdamMail <<http://www.amsterdammail.nl/>> en Google News <<http://www.google.com/newsalerts?hl=en>>

¹³ Canada <<http://www.canada.gc.ca/canada/SCS?l=1&en=yes>>

¹⁴ My.ecitizen <<http://www.ecitizen.gov.sg/MyeCitizen/index.htm>>

¹⁵ Experts-exchange <<http://www.experts-exchange.com/>>

¹⁶ Amazon <<http://www.amazon.com>>

informatie. Het gaat uiteindelijk wel om de overheid als betrouwbare informatiebron en om Overheid.nl als neutrale partij in het ontsluiten van informatie. Desalniettemin is het interessant om verder te onderzoeken en experimenten op te zetten.

Directgov: de inzet van content en doelgroepen

Directgov¹⁷ is/wordt de opvolger van UK Online. We noemen deze site afzonderlijk, omdat dit portal fundamenteel van andere overheidsportals verschilt. Het portal kan 90% van de zoekopdrachten direct (op de site zelf) afhandelen. De site biedt veel eigen content in plaats van externe links. De departementen vullen ieder een deel van het portal ('franchising'). Een bijkomend voordeel is de consistentie ('Common Look and Feel').

Opvallend is het grote aantal doelgroepen (automobilisten, ouders, senioren, etc.). De indeling in levensfasen is vervallen. Verder heeft het portal een context-gevoelige zoekmachine (die ook vragen in natuurlijke taal begrijpt), FAQ, een helpdesk, links naar derden, een alfabetische trefwoordenlijst en een 'subject directory'. De informatie is ook beschikbaar via digitale TV ('multi-channeling').

Uit een recente survey bleek dat het merendeel van de ondervraagden zeer tevreden is over Directgov¹⁸.

Advies overheid.nl dient een aantal van bovengenoemde overwegingen in de visie-ontwikkeling mee te nemen: de indeling in doelgroepen, de hoeveelheid content, de mate van samenwerking, links naar derden en multi-channeling. Het is om deze reden dat deze activiteit in de ontwikkelijn Vernieuwen is opgenomen.

Marketing/Distributie

De meest eenvoudige vorm van distributie, is het plaatsen van een klein zoekvenster op een andere site. Het aantal gemeenten¹⁹ en portals dat nu al een zoekbox van Overheid.nl op hun site heeft, is klein en kan verder uitgebreid worden. Deze distributievorm kan als marketinginstrument gezien worden.

Een andere beperkte vorm van distributie is het plaatsen van RSS-feeds (bijv. met koppen van recente persberichten of kamerstukken)²⁰. Voor professionele gebruikers kan een browser-toolbar ontwikkeld worden. Dit zijn kleine, eenmalige investeringen.

Bij een verdergaande vorm kan de zoekmachine als Shared Service worden opgezet. Hierdoor hoeven overheidsorganisaties geen eigen zoekmachine meer te hebben, maar gebruiken zij de zoekmachine van Overheid.nl. Dan kan ingesteld worden dat de zoekresultaten van de eigen site als eerste worden getoond. Op de interface kan de gebruiker aangeven of hij/zij op de site wil zoeken of "in de hele overheid". Overheid.nl is dan een soort application service provider. De zoekmachine van FirstGov (die gebruik maakt van FAST) is bijvoorbeeld ook op de sites van het ministeries van Landbouw en Buitenlandse Zaken te vinden²¹. Verdere integratie is mogelijk door bijvoorbeeld binnen dossiers van ministeries rechtstreeks naar zoekresultaten van de zoekmachine te verwijzen, die continu geïndexeerd worden. Een deel van het dossier wordt hiermee automatisch samengesteld.

¹⁷ Directgov <<http://www.direct.gov.uk/>>

¹⁸

<http://europa.eu.int/ISPO/ida/jsp/index.jsp?fuseAction=showDocument&documentID=2494&parent=chapter&preChapterID=0-140-194>

¹⁹ Gemeente Roermond <<http://www.roermond.nl/index8653.htm>>, MSN <<http://www.msn.nl/>>

²⁰ NU.nl <<http://www.nu.nl/rss/>>

²¹ FirstGov <<http://www.firstgov.gov/>>, USDA <<http://www.usda.gov/>>, US Dept. of State <<http://www.state.gov/>>, FAST <<http://www.fast.no/>>

Metadatastructuren

Veel landen en overheidsorganisaties werken aan structuren gebaseerd op metadata zoals thesauri, taxonomieën en topic maps. We gaan hier kort in op een aantal vormen.

Soorten metadatastructuren

Gartner onderscheidt vijf soorten metadatastructuren geordend op toenemende complexiteit en mate van semantiek.

1. woordenlijsten: indices, glossaries, dictionaries
2. classificatiesystemen: taxonomieën, subject heading, hiërarchieën
3. semantische representaties: knowledge maps, semantische netwerken, thesauri, topic maps
4. kennisgebaseerde representaties: case-based reasoning, causale modellen, KRL
5. ontologieën: SHOE, RDF, DAML+OIL, Ontology Exchange Language

Voor de onderzoekswerkzaamheden in dit (deel)project dienen de verschillende metadatastructuren die in Nederland worden gebruikt in het licht van de zoekmachine te worden bekeken. Wij noemen hier:

- *Taxonomie*

Een taxonomie is een concept waarbij op basis van een hiërarchische classificatie (met gebruikmaking van onderlinge relaties) en applicatie (ICT-tool) de mogelijkheid wordt gecreëerd om interne en externe informatie vanuit het gebruikersperspectief (of corporate gezichtspunt) te ontsluiten en waarbij hulpmiddelen voor navigatie zorgdragen voor een snelle retrieval (bron: Reekx) of simpel gezegd een hiërarchische structuur van termen en categorieën.

- *Thesaurus*

Een thesaurus is een alfabetisch geordende gestructureerde en gecontroleerde lijst van termen met hun onderlinge (hiërarchische, associatieve en/of semantische) relaties, die wordt gebruikt bij het toegankelijk maken en raadplegen van (geautomatiseerde) informatie (bron: WODC). De gebruiker kan zoeken op nauwere (meer specifieke),

- *Topic Map*

Een topic map kan dienen als navigatielaag bovenop een informatieverzameling. Topic maps zijn geschikt voor portals zoals Overheid.nl, omdat ze informatie uit een grote hoeveelheid heterogene bronnen kunnen combineren. Een topic map identificeert de concepten die in de informatie liggen opgeslagen, omschrijft de tussenliggende relaties (context) en legt een verbinding tussen de concepten en onderliggende informatie. Topic maps vormen een conceptuele laag waarin kennis uit en over documenten onafhankelijk van die documenten kan worden vastgelegd.

- *Ontologie*

Een ontologie is een classificatiesysteem voor begrippen en hun onderlinge verbindingen binnen een bepaald kennisdomein. Op middellange termijn zijn ontologieën interessant om mee te experimenteren. De ontwikkelingen zijn nog in een pril stadium.

4.4. Samenvattend

De invulling van de functionaliteiten, positionering en organisatie van de nieuwe zoekmachine is veelomvattend. Om deze reden zijn in het (deel)project drie ontwikkellijnen benoemd. In de ontwikkellijn Verbeteren (Quick Wins) worden snelle, gemakkelijk te realiseren veranderingen in de huidige zoekmachine aangebracht, waardoor al op de korte termijn voor de burger verbeteringen merkbaar zijn. In de meer structurele ontwikkellijn Vernieuwen worden de in dit hoofdstuk benoemde aspecten nader onderzocht en afgewogen in een toekomstvisie. Vanuit de ontwikkelde visie worden de specificaties vastgesteld voor en de realisatie gewaarborgd van de nieuwe zoekmachine. Aangezien de ontwikkelingen wat betreft de technologie en de behoeften van burgers niet stil staan zal parallel aan de realisatie en de operationalisatie van de nieuwe zoekmachine in de ontwikkellijn Toekomstverkenningen verder gezocht worden naar mogelijke vernieuwingen voor de verdere toekomst.

Daarnaast moet onderzocht worden welke zoekmachines (inclusief licentiemodellen) gebruikt worden op buitenlandse overheidsportals, evenals welke zoekmachines er al in omloop zijn binnen de (Rijks-)overheid. Wellicht kan gebruik gemaakt worden van bestaande licenties. Tenslotte moeten de ontwikkelingen op het gebied van open source zoekmachines worden onderzocht.

5. Kwaliteitsverwachtingen

In het vorige hoofdstuk zijn kenmerken van de overheidsinformatie, de uitgangspunten voor zoekmachine en een aantal relevante ontwikkelingen beschreven. Deze zijn van belang bij de uitvoering van de verschillende activiteiten in het (deel)project. Daarnaast is van belang wat de verwachtingen zijn van verschillende betrokken doelgroepen voor de zoekmachine. Met andere woorden het (deel)project kan vanuit meerdere perspectieven bekeken worden:

- de burger,
- de content,
- de content-leverancier,
- de beheerder van de zoekmachine,
- de technologie en het groter geheel.

Vanuit deze verschillende perspectieven zullen ook de verwachtingen ten aanzien van de zoekmachine verschillen. In dit hoofdstuk zetten wij deze verwachtingen op hoofdlijnen uiteen en vertalen die naar activiteiten in de drie ontwikkellijnen (in de toegevoegde figuur).

Perspectief van de burger

De burgers wil zo snel en eenvoudig bij de juiste informatie terechtkomen. Overheid.nl kan het zoekproces faciliteren, o.a. met een portal en zoekmachine en in de toekomst mogelijk via meerdere kanalen. Gebruikers zullen de zoekmachine alleen gebruiken indien de presentatie van de resultaten duidelijk is en de resultaten relevant zijn. Door webmarketing moet de naamsbekendheid van Overheid.nl onder burgers vergroot worden.

Perspectief van de content

Overheid.nl streeft naar een volledige ontsluiting van alle bronnen met overheidsinformatie. De focus ligt primair op het ontsluiten van producten en diensten. Later worden ook andere bronnen ontsloten. Zo komen bijvoorbeeld decentrale wettenbanken en productencatalogi online beschikbaar. De beschikbaarheid van nieuwe bronnen en informatietypen moet geïnteriseerd worden.

Ontsluiting van alle overheidsinformatie is niet altijd mogelijk (bij betaalde toegang) of wenselijk (geen overheid). Overheid.nl is relatief aanbodgericht. Meer vraaggerichte wijzen van ontsluiting worden onderzocht.

Perspectief van de content-leverancier

De content-leverancier (bronbeheerder) heeft een belang bij een goede ontsluiting van zijn informatie, zeker indien de zoekmachine van Overheid.nl als zoekmachine voor zijn site is te gebruiken. Hij zal eerder geneigd zijn metadata toe te kennen en problemen met de indexering te verhelpen.

Perspectief van de beheerder (het beheer)

De beheerlast voor Advies Overheid.nl zal toenemen, maar moet minimaal zijn. Het beheer omvat ook de analyse en monitoring van het zoekgedrag en zoekproces.

Met het succes van de zoekmachine als spil in de informatievoorziening en dienstverlening aan burgers zal de benodigde beheercapaciteit toenemen. Dat blijkt ook bij succesvolle zoekmachines (b.v. Google, Ilse, DirectGov, etc.) in de markt het geval te zijn. Bij bepaalde oplossingen kan in de kosten gedeeld worden.

Perspectief van de technologie

Bij het realiseren van de vernieuwde zoekmachine wordt gebruik gemaakt van bewezen technologie en open standaarden. Ook bij een toenemend aantal queries, een aantal nieuwe bronnen en een toenemende hoeveelheid informatie per bron moeten de prestaties en het gebruiksgemak gelijk blijven. Metadata biedt burgers mogelijkheden om specifieker te zoeken dan alleen op trefwoorden. Mogelijk zijn naast de zoekmachine andere ontsluitingsmechanismen (navigatiemethoden) gewenst.

Perspectief van de zoekmachine als component in een groter geheel

Internet is slechts één van de kanalen om overheidsinformatie te verkrijgen. Overheid.nl is slechts één van de portalen voor overheidsinformatie op Internet. De zoekmachine is slechts één van de ontsluitingsmechanismen op Overheid.nl. Door samenwerking, integratie en distributie kan het belang van de zoekmachine in het geheel worden vergroot. Advies Overheid.nl gaat het wiel niet opnieuw opvinden en zal zich indien mogelijk bij bestaande initiatieven aansluiten om bijvoorbeeld een taxonomie, thesaurus of andere e-government bouwstenen te ontwikkelen. De zoekmachine zelf kan ook een component zijn die door andere overheidsorganisaties gebruikt kan worden.

Van kwaliteitsverwachtingen naar activiteiten

In de onderstaande tabel zijn de kwaliteitsverwachtingen vanuit de bovengenoemde perspectieven vertaald in de activiteiten die in de verschillende ontwikkellijnen worden uitgevoerd. De ontwikkellijn Toekomstverkenningen valt voor een belangrijk deel buiten de scope van dit (deel)project maar biedt een doorkijkje naar de verdere toekomst van de zoekmachine.

IN DE ONDERSTAANDE TABEL GEVEN DE GRIJNSGEDRUKTE CELLEN MOGELIJKE ONTWIKKELINGEN AAN.

Perspectief	Ontwikkellijn Aspect	Verbeteren	Vernieuwen	Toekomstverkenningen
Burger	<i>vindbaarheid informatie en diensten</i>	Naamsbekendheid vergroten / presentatie resultaten verbeteren	Zoekproces faciliteren / relevantie resultaten verbeteren	Integratie van portals
	<i>Bereik (doelgroepen)</i>	Meer bezoekers	Meer burgers	Meer studenten, automobilisten, senioren, etc.
	<i>Kanalen</i>	Internet	Internet	Internet, GSM, e-mail
Content	<i>Beschikbaarheid</i>	Bestaande bronnen beter ontsluiten (indexeren)	Nieuwe bronnen ontsluiten (volledigheid)	Content-analyse: Nieuwe bronnen zoeken en ontsluiten
	<i>Focus</i>	Producten en diensten	Alle overheidsinformatie	Multimedia, feeds
	<i>Oriëntatie</i>	Aanbodgericht (bronnen)	Vraaggericht ('no wrong door')	Gepersonaliseerde antwoorden
Content-leverancier	<i>distributie</i>	Zoekbox op website plaatsen (syndicatie)	Integratie van metazoek-machine in website	Shared Service of Overheid.nl als ASP
	<i>samenwerking</i>	Ad hoc	Periodiek	Contractueel
	<i>indexering</i>	Meer pagina's indexeren	Pagina's beter indexeren (metadata)	Controlled term indexing
Beheer(der) zoekmachine	<i>kosten</i>	Projectkosten	Gezamenlijk investeren	Totale kosten-besparing
	<i>analyse / monitoring zoekgedrag, zoek-proces en vraag</i>	Incidentele, handmatige analyses	Periodieke, semi-automatische analyses	Real-time inzicht door automatische analyse
	<i>organisatie technisch beheer</i>	Onveranderd	Zelf of ASP (nieuwe beheer-organisatie)	Gezamenlijk / centraal
Technologie	<i>acceptatiegraad</i>	Bestaande technologie	Nieuwe, bewezen technologie (nieuwe zoekmachine)	Nieuwe, veelbelovende technologie
	<i>zoekmethode</i>	Trefwoorden (o.a. booleaans)	Metadata-waarden en evt. natuurlijke taal	Concepten / nauwere, bredere en gerelateerde termen
	<i>ontsluitings-mechanismen</i>	+ top 5 van producten op homepage	+ thema-pagina's	+ trefwoorden-boom en evt. 'social search'
Zoekmachine als component	<i>ontwikkelingswijze</i>	Bestaande initiatieven inventariseren	Aansluiten bij bestaande initiatieven	Nieuwe projecten initiëren / experimenteren
	<i>e-government bouwstenen</i>	Metadata	Taxonomie	Thesaurus
	<i>integratie</i>	Banners plaatsen (link)	Specifiek doorverwijzen (deeplinks)	Eén zoekmachine voor alle overheidsportals

6. Acceptatiecriteria

In dit hoofdstuk worden de acceptatie criteria genoemd die voor de 3 ontwikkellijnen zijn te onderscheiden.

Verbeteren

- Op de homepage van Overheid.nl wordt een top 5 met veelgevraagde producten en diensten geplaatst;
- Er is een beheermodule om zoekwoorden aan bronnen te koppelen;
- Veel gebruikte zoekwoorden worden afgevangen;
- De zoekbox van Overheid.nl of banner is ook op andere overheidswebsites te vinden;
- De presentatie van de zoekresultaten is verbeterd. De resultaten worden gesplitst per informatietype (personen, organisaties, producten, webpagina's);
- De relevantie van zoekresultaten is zichtbaar verbeterd: geen error-pagina's, geen dubbele pagina's, zoekwoord in URL scoort hoger, etc;
- De thematische startpagina's worden gekoppeld aan de zoekresultaten;
- Er is een meetinstrument om het zoekgedrag in kaart te brengen;
- Webmarketing (waaronder Google Adwords) en syndicatie van de zoekbox, zie Actieplan Promotiecampagne Overheid.nl najaar 2004.

Vernieuwen

- Advies Overheid.nl heeft een visie ontwikkeld op de toekomst van Overheid.nl;
- Advies Overheid.nl heeft de toepasbaarheid van zoektechnologieën, taxonomieën en thesauri en andere ontsluitingsmechanismen laten onderzoeken;
- De mogelijkheden en wenselijkheid van een shared service zijn onderzocht;
- De zoekmachine op Overheid.nl wordt vernieuwd;
- Er is een nieuwe interface ontwikkeld waarbij de zoekmachine een prominentere rol heeft;
- Ontwikkeling van een nieuw navigatiemechanisme (naast de zoekmachine) is niet uitgesloten;
- Er heeft afstemming plaatsgevonden met andere (deel)projecten rondom de ontsluiting van overheidsinformatie;
- Met behulp van de zoekmachine kunnen burgers specifiek zoeken, d.w.z. niet alleen op trefwoorden / booleaans;
- Alle beschikbare productencatalogi (rijk, gemeenten, provincies en waterschappen, waar mogelijk ZBO's) worden door de zoekmachine ontsloten;
- Met behulp van de zoekmachine worden burgers die een product in de verkeerde productencatalogus zoeken naar de juiste organisatie verwezen;
- Advies Overheid heeft standaarden voor overheidsorganisaties vastgesteld en gepubliceerd om de toegankelijkheid te verbeteren;
- Spider-problemen worden gereduceerd;
- De logbestanden van de zoekmachine worden maandelijks geanalyseerd. Veel gebruikte zoekwoorden worden afgevangen;
- Advies Overheid.nl heeft een overzicht van nieuwe bronnen en informatietypen en een beeld van hoe deze te ontsluiten;
- Advies Overheid.nl is nieuwe samenwerkingsverbanden aangegaan op het gebied van distributie en marketing;
- Er is een nieuwe beheerorganisatie opgezet.

Toekomstverkenningen

- Er zijn experimenten met veelbelovende zoektechnologieën uitgewerkt, uitgevoerd en geëvalueerd;
- De beschikbaarheid van nieuwe bronnen is geïnventariseerd;
- Via de zoekmachine en bijbehorende ontsluitingsmechanismen wordt waar mogelijk alle overheidsinformatie ontsloten, tenminste alle informatie rondom producten en diensten en de democratische rechtsstaat;
- De vernieuwde zoekmachine is geëvalueerd en aanpassingen zijn verricht;
- Er ligt een projectplan met vervolgstappen (o.a. op basis van een onderzoek naar internationale ontwikkelingen).

7. Organisatiestructuur

Opdrachtgever:

- BZK/DIOS (Rob Evelo)

Opdrachtnemer:

- Advies Overheid.nl

Projectteam (Advies Overheid.nl)

- Senior projectleider
- Projectmedewerkers (1 à 2, afhankelijk van de verhouding uitbesteden vs. zelf doen)
- Projectondersteuning: via Advies Overheid.nl

De werving en selectie van medewerkers is niet in de planning en begroting meegenomen. De verwachting is dat medewerkers per augustus starten. In de planning en begroting is meegenomen dat bepaalde werkzaamheden tijdelijk door externen worden uitgevoerd. Om de planning in de tijd te kunnen realiseren zullen de quick wins door externen moeten worden gerealiseerd (er is nog geen projectleider beschikbaar).

Er wordt een klankbordgroep opgezet van betrokken partijen/mensen die het projectteam adviseren en uitvoering geven aan mogelijk lokale implementaties:

- Vertegenwoordigers van de verschillende overheidslagen: Rijksoverheid, E-provincies, EGEM, gemeenten (b.v. Dordrecht, Voorst, Groningen), etc.
- Projectleiders e-dienstverlening bij implementerende partijen
 - Gemeenten
 - Waterschappen
 - Provincies

Gebruikersgroep ²² die input en feedback kan geven aan projectteam en lokale implementatie teams tijdens ontwikkeling en implementatie zoekmachine:

- Gerelateerde zoekmachines (Postbus 51, MSN, Ilse, etc.);
- Burgers (Consumentenbond, Burger@Overheid);
- Webpanel;

Operationeel beheer van de zoekmachine bestaat uit een aantal lagen/rollen. Onderstaand geven we in grote lijnen aan hoe de invulling functioneel in elkaar zit. De kwantiteit in termen van fte's is sterk afhankelijk van de keuzes die gemaakt zullen worden t.a.v. de positionering in de toekomst. In de ontwikkelijn vernieuwen zal dit verder worden uitgewerkt. Keuzes voor bijvoorbeeld een Shared Service Centre hebben op de uiteindelijke invulling een grote invloed.

Regievoering:	Advies Overheid.nl
Communicatie:	Redactie Advies overheid.nl
Redactie:	Redactie Advies overheid.nl
Dagelijks beheer:	Webmaster
Technisch beheer:	Ontwikkelorganisatie
Financieel beheer:	Advies Overheid.nl
Hosting/exploitatie:	(nader te bepalen)

²² Deze gebruikersgroep kan gedeeld worden met het project Transparantie.

8. Risicologboek

Risico	Kans (0-1)	Impact (1-3)	Prioriteit (kans x impact)	Oplossing(en)
project				
gebrek aan aansluiting bij andere projecten (ICTU, BZK, Rijk, provincies, gemeenten)	0,5	3	1,5	Overleg, communicatie en samenwerking Omgevingsanalyse aan de start van elke actielijn
gebrek aan resultaten op korte termijn	0,75	3	2,25	Indeling aanpak in drie ontwikkellijnen met eigen resultaten
communicatie				
concurrentie / overlap met andere portals (o.a. doelgroepen, positionering)	0,75	2	1,5	Samenwerken en afstemmen. Duidelijke communicatie.
concurrentie in webmarketing	0,5	1	0,5	Samenwerking en afstemming
geen toename in gebruik	0,25	2	0,5	Gebruikers- en/of communicatieonderzoek
organisatie				
toenemende beheerlast	1	2	2	Eigen content en redactie beperken. Taken automatiseren. Samenwerking zoeken met contentleveranciers
medewerking eigenaren productencatalogi	0,75	3	2,25	win-winsituatie duidelijk maken
kosten				
kosten nieuwe zoekmachine	0,5	3	1,5	Focussen op belangrijkste resultaatpunten, samenwerking met andere organisaties in de vorm van een shared service
kosten ontwikkeling en beheer taxonomie	0,75	3	2,25	Samenwerking met andere organisaties
kosten ontwikkeling en beheer thesaurus	0,75	3	2,25	Samenwerking met andere organisaties
technologie				
kwaliteit resultaten ondermaats	0,75	3	2,25	gebruik metadata stimuleren, finetuning ranking
complexere zoekopdrachten	0,75	2	1,5	gebruik metadata stimuleren
grotere hoeveelheid content	0,75	2	1,5	specifiekere zoekfunctionaliteit
begrijpbaarheid content voor doorsnee burger	0,5	2	1	doorverwijzen naar Postbus 51, aansluiten bij
vindbaarheid	0,75	3	2,25	'no wrong door'-principe: verwijzen naar andere organisaties
bronnen niet of later	0,5	1	0,5	Afhankelijk van content leverancier

beschikbaar				
indexeerbaarheid van bepaalde bronnen	0,75	1	0,75	overleg met content-eigenaar
geen of verkeerd gebruik metadata	0,5	2	1	overleg met content-eigenaar
huidige leverancier kan of wil niet voldoen aan wensen en eisen (quick wins)	0,25	2	0,5	splitsing zoekfunctionaliteit: indexering en ranking, presentatie, etc.

9. Projectaanpak

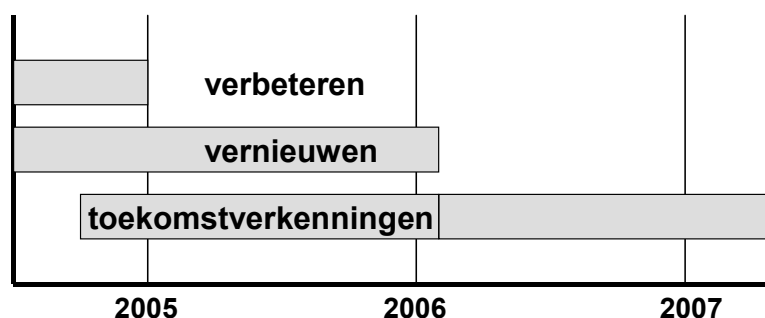
9.1. Inleiding

In dit hoofdstuk wordt de aanpak op hoofdlijnen uitgewerkt. In de paragrafen hierna worden de drie ontwikkellijnen globaal uitgewerkt in opzet, resultaten en acties. In hoofdstuk 9 is een gedetailleerd stappenplan en planning uitgewerkt.

Drie ontwikkellijnen: verbeteren, vernieuwen, toekomstverkenningen

In de periode medio 2004 – eind 2007 worden drie ontwikkellijnen gevolgd, met hun eigen resultaten, activiteiten en dynamiek:

- Verbeteren: de korte termijn ontwikkeling, zal eind 2004 afgerond worden en omvat een aantal "Quick wins".
- Vernieuwen: Tegelijkertijd hiermee start de ontwikkellijn vernieuwen, die naar een vernieuwde en sterk verbeterde zoekmachine zal leiden. Dit wordt eind 2005/begin 2006 afgerond en heeft in eerste instantie dienstverlening als focus.
- Toekomstverkenningen: Overheid.nl moet de komende jaren blijven aansluiten op nationale behoeften en internationale ontwikkelingen en het terrein zal verder verbreden. In deze ontwikkellijn worden hiervoor de voorbereidingen getroffen.



Beslismomenten op basis van visie-ontwikkeling en onderzoek

In de ontwikkellijnen worden de toepasbaarheid en inzet van technologieën, de wenselijkheid van initiatieven en de mogelijkheden van samenwerking verder onderzocht. Hiervoor zijn beslismomenten ingebouwd, die binnen de aangegeven kaders de koers bepalen.

9.2. Verbeteren: medio 2004 – eind 2004

Quick wins in interface, technologie, distributie en samenwerking

Op de korte termijn kan het doorvoeren van een aantal wijzigingen in de zoekinterface, andere ontsluitingsmechanismen en technologie veel opleveren, tegen relatief geringe kosten. Daarnaast kan worden gekeken naar korte termijn mogelijkheden voor samenwerking en distributie. Een analyse van deze zogenaamde quick wins is te vinden in bijlage 2.

In het (deel)project zal een korte analyse van de huidige situatie plaatsvinden en een analyse van de quick wins, waarna realisatie plaatsvindt. Hier wordt ook gekeken naar de mogelijkheden om snel nieuwe bronnen aan te sluiten (bijv. decentrale wet- en regelgeving).

Ontwikkelen en invoeren verbeterde meetinstrumenten

Tegelijkertijd met deze verbetering wordt een meetinstrument (voor de monitoring) van het bezoek en zoekgedrag van Overheid.nl ontwikkeld. Waar mogelijk worden deze gegevens vergeleken met andere overheidssites om een beter beeld te krijgen van het zoekgedrag van de burger.

Op de korte termijn worden geen majeure veranderingen in de technologie of de organisatie doorgevoerd.

Activiteiten op hoofdlijnen

- Analyse quick wins
- Besluitvorming quick wins
- Quick wins in zoekinterface en –technologie worden in overleg met de leverancier gerealiseerd.
- Aanpassen interactie/homepage
- Zorgen voor de ontsluiting van meer bronnen
- Continue monitoring van het zoekproces

9.3. Vernieuwen: medio 2004 – eind 2005

Vernieuwde zoekmachine met focus op dienstverlening

In de ontwikkellijn Vernieuwen wordt een vernieuwde zoekmachine gerealiseerd met aansluiting op andere ontsluitingsmechanismen. De focus ligt hierbij in eerste instantie op het ontsluiten van informatie rondom de diensten en producten van de overheid, waarbij een no-wrong-door principe moet gelden.

Visie-ontwikkeling en onderzoek

Om de zoekmachine van Overheid.nl te laten meegroeien met nationale en internationale ontwikkelingen, moet een aantal basiskeuzes gemaakt worden op de volgende gebieden (zie ook hoofdstuk 4 en bijlage 2):

- mate van samenwerking met andere organisaties: positionering, marketing, inzet functionaliteiten en technologie;
- volledigheid versus hanteerbaarheid van informatie: gebruik bronnen buiten de overheid;
- mate van ondersteuning van de burger door zoekproces te ondersteunen, verschillende ingangen, relaties met andere zoekinterfaces, inzet eigen content en redactie, gebruik van doelgroepen, onderscheid verschillende soorten content, inzet van personalisering

Op basis van deze keuzes wordt bepaald welke technologieën ingezet zullen worden, hoe de interface ondersteunt en welke afspraken met andere partijen dienen te worden gemaakt. Ook kan worden bepaald welk exploitatie- c.q. beheermodel (ASP, eigen beheer, etc.) voor de te ontwikkelen zoekmachine gekozen moet worden. In bijlage 2 is een quick scan opgenomen voor te onderzoeken technologieën en ontsluitingsmechanismen, met een uitwerking hoe deze zaken onderzocht dienen te worden.

Aanbesteding en bouw

Op basis van de eisen en wensen die in het analyse traject worden vastgesteld wordt een selectie- en aanbestedingstraject gestart voor de bouw van de nieuwe zoekmachine en wordt de vernieuwde zoekmachine gerealiseerd.

Inrichting organisatie en samenwerking

Naast de bouw van de vernieuwde zoekmachine wordt een beheerorganisatie opgezet en worden de afspraken met andere partijen ingevuld. Rondom de nieuwe zoekmachine wordt promotie opgezet.

Activiteiten op hoofdlijnen

- Visie-ontwikkeling
- Besluitvorming
- Wensen en eisen uitwerken
- Aanbesteden
- Realisatie bouw
- Organiseren beheer
- Samenwerking met andere overheidsorganisaties
- Meewerken aan standaarden ontsluiting
- Marketing in gang zetten

9.4. Toekomstverkenningen: 2005 tot 2007

Volgen ontwikkelingen en opzetten experimenten

Gezien de ontwikkelingen op het gebied van zoekmachines en overheidsinformatie zal de vernieuwde zoekmachine niet het eindpunt zijn. In hoofdstuk 4 is een analyse opgenomen van veelbelovende technologieën. Overheid.nl kan de ontwikkelingen volgen en in sommige gevallen experimenten opzetten.

Uitbreiding van dienstverlening naar informatie en organisaties

Waar tot eind 2005 de focus zal liggen op dienstverlening: het ontsluiten van product- en dienst informatie, worden vervolgens de ontwikkelingen doorgetrokken naar de twee andere domeinen: overheidsinformatie in brede zin en overheidsorganisaties.

Visie/projectplan vervolg

Aan het einde van de projectperiode zal een plan worden geschreven voor het vervolg.

Activiteiten op hoofdlijnen

- in gang zetten experimenten
- uitbouwen naar ontsluiting overheidsinformatie en –organisaties
- resultaten en visie evalueren
- projectplan schrijven vervolg

10. Begroting / Planning

In dit hoofdstuk wordt een overzicht gegeven van de benodigde financiën voor het (deel)project. De gedetailleerde begroting en planning is opgenomen in bijlage 4 (Excel-bestand).

Verwachte kosten van het (deel)project verdeeld naar kostengroep.

Kostengroep	Verwacht €
bouw	345,000
management	92,000
onderzoek	193,500
ontwerp	105,000
software	350,000
voorbereiding	67,500
Totaal	1,153,000

Personeel	157,500
Eindtotaal	1,310,500

In het (deel)project moet ook rekening worden gehouden met kosten van marketing en distributie. Maar omdat het de voorkeur verdient deze kosten te delen met andere (deel)projecten zijn ze voor dit (deel)project niet nader gespecificeerd. Bij de uitvoering zal, met name in de doorontwikkeling naar andere domeinen gezocht moeten worden naar een goede verdeling van de hiervoor benodigde kosten.

Verwachte kosten van het (deel)project verdeeld over de tijd.

	Totaal
2004	533,000
2005	777,500
Eindtotaal	1,310,500

Aanname voor de personeelskosten is dat een projectmedewerker gedurende anderhalf jaar is benodigde.

Bijlage 1: Bronnenoverzicht

Interviews

Erik Hup, programmamanager Overheid.nl
Ferdinand Harmsen, projectleider Overheid.nl
Gerry Bron, coördinator Informatiemanagement RVD/Postbus 51
Hans Dussel, projectleider Overheid.nl
Jacqueline Roesel, redacteur Overheid.nl
Matthijs Sienot, MSN
Peter Waters, projectleider Overheid.nl
Raph de Rooij, adviseur Overheid.nl
Rob Evelo, coördinerend beleidsmedewerker Dienstverlening en Toegankelijkheid, DIOS, Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Analyse

Desk research: internationale sites (zie hoofdstuk 4)
Analyse zoekvragen MetaZoekDienst Overheid.nl

Literatuur

Zuurmond, A. en Lammers, K., *De Kunst van het vinden*, Itafit, 2002
Projectplannen Advies Overheid.nl
Ministerie van BZK, *Actieprogramma 'Andere Overheid'*, 2003
Column Hans Dekker: "Wat wil die burger?", 2004 (geschreven in opdracht van Advies Overheid.nl)

Bijlage 2 : Quick scan technologieën en ontsluitingsmechanismen en aanbevelingen onderzoek

Tijdens het schrijven van het projectplan is een quick scan gedaan naar internationale ontwikkelingen in het ontsluiten van (overheids-)informatie en de technologieën en ontsluitingsmechanismen die dit ondersteunen. In deze bijlage geven we een weergave van deze quick scan. We hanteren daarbij de volgende indeling, waarvan ieder onderdeel toegelicht wordt en wordt geïllustreerd met buitenlandse voorbeelden:

- keyword handling: het afvangen van veelgebruikte zoekwoorden door het meten van zoekgedrag
- zoektechnologieën: technologieën die het zoekproces ondersteunen en de relevantie van resultaten verbeteren
- andere ontsluitingsmechanismen: mogelijkheden van ontsluiting buiten het "zoekscherf" om
- taxonomieën/thesauri/semantische indelingen: mechanismen om informatie beter doorzoekbaar te maken of te structureren, veelal door het vooraf uitwerken van een indeling en toekennen van verrijkte informatie aan documenten of aan zoekvragen.
- marketing/distributie: het breder beschikbaar maken van de zoekmachine, door functionaliteiten beschikbaar te maken voor anderen en de zichtbaarheid van de zoekmachine te vergroten

Keyword handling

Veel zoekmachines vangen veelgebruikte trefwoorden af, om gesponsorde resultaten te kunnen weergeven of om de gebruiker snel verder te helpen naar wat hij vermoedelijk zocht. In de commerciële wereld wordt bijvoorbeeld veel voor termen als "autoverzekering" betaald. Gebaseerd op het zoekgedrag binnen Overheid.nl kan dit vertaald worden naar woorden als euthanasie of varkenspest, die kunnen verwijzen naar de gelijknamige dossiers bij respectievelijk de ministeries BZK en LNV. De bekende zoekmachine Google ²³ (Amerikaanse versie) vangt ook gestructureerde vragen af (bijvoorbeeld adressen en nummers van patenten). In Nederland kan dit worden vertaald naar bijvoorbeeld nummers van kamerstukken. Een veel gebruikt mechanisme is ook het plaatsen van de meest gebruikte zoektermen op de homepage. Dit gebeurt nu al in het overheidsloket. Belangrijk om te weten is dat de grotere portalen continu bezoek en zoekgedrag meten en verwerken in hun aanbod en presentatie.

Deze mechanismen wordt als quick win gedefinieerd voor Overheid.nl en komt terug in de bijlage 3.

Zoektechnologieën

De huidige zoekmethode is gebaseerd op trefwoorden en operatoren (booleaans en 'stemming'). Deze veelgebruikte zoekmethoden zijn goedkoop en eenvoudig te gebruiken. Meer geavanceerde zoekmethoden maken gebruik van een thesaurus. Op dit begrip komen we later nog terug. Door ook op synoniemen en homoniemen (vocabulaire) te zoeken wordt de 'recall' verbeterd. De gebruiker kan zoeken op gerelateerde termen, nauwere en bredere termen. Internationaal zien we een beperkte toepassing van andere zoekmechanismen.

²³ Google.com <<http://www.google.com>>

Associatief zoeken

De Aquabrowser ²⁴ is een voorbeeld van een associatieve zoekmachine die de gebruiker gerelateerde termen (associaties), synoniemen en vertalingen en 'vage' (fuzzy) en spellings-alternatieven toont. Dit wordt in Nederland bijvoorbeeld gebruikt bij de site ehbw.nl (Eerste Hulp bij Werk). Het mechanisme is met name relevant om een bepaald domein te verkennen en vergt veel initiële tijdsinvestering

Zoeken via concepten/patronen/natuurlijke taal

Zoeken op patronen is bijvoorbeeld handig wanneer een woord meerdere spellingswijzen heeft. Dit wordt bijvoorbeeld gebruikt bij de Amerikaanse Sociale Dienst²⁵. Sommige zoekmachines kunnen "concepten" extraheren uit informatie. Technologieën die dit ondersteunen maken statistische analyses van teksten en zijn vaak goed bruikbaar binnen specifieke domeinen en bij ongestructureerde informatie.

Automatische classificatie

Sommige zoekmachines kunnen automatisch resultaten classificeren in (vooraf uitgewerkte of on the fly gegenereerde) categorieën. In sommige gevallen worden deze resultaten gevisualiseerd²⁶. Informatieverzamelingen die van metadata zien voorzien, kunnen specifiek doorzocht worden ²⁷.

Natuurlijke taaltechnologie

Bepaalde zoekmachines kunnen omgaan met "natuurlijke taal". De zoekmachine analyseert de vraag en haar structuur (Hoeveel zetels heeft de PvdA) en genereert op basis hiervan een specifiekere zoekvraag. Een voorbeeld hiervan is al@din van de site bibliotheek.nl²⁸.

Social search

Social search is een technologie die gebruikers elkaar laat helpen. Bij het intikken van een zoekvraag komt bijvoorbeeld naar voren welke vergelijkbare vragen door andere gebruikers zijn gesteld, waardoor de gebruiker wordt geholpen zijn vraag beter te definiëren²⁹. Amazon zet haar technologie in om gebruikers die op een bepaald product zoeken, te laten zien wat andere gebruikers, die ook in dit product geïnteresseerd waren, nog meer zochten. Wat hier speelt is dat alle producten zich in een database bevinden, waardoor deze technologie gemakkelijker kan worden ingezet.

Tenslotte zijn er allerlei experimentele technologieën, waarmee bijvoorbeeld in beeldbestanden kan worden gezocht, of die gebaseerd zijn op rss-feeds.

De toepasbaarheid van bovenstaande technologieën zal moeten worden beoordeeld op basis van een analyse van de bronnen en content van Overheid.nl. De genoemde technologieën zijn vaak behoorlijk prijzig. De toegevoegde waarde moet kunnen worden aangetoond. In de conclusies doen we een aanzet voor de te onderzoeken technologieën.

²⁴ Eerste Hulp Bij Werk <<http://www.ehbw.nl/>>

²⁵ Zoekmethoden <<http://www.ssa.gov/search/>>

²⁶ KartOO <<http://www.kartoo.com/>>

²⁷ Australian Government Entry Point <<http://www.fed.gov.au/>>

²⁸ Al@din <<http://www.bibliotheek.nl/>>

²⁹ Eurekster <<http://www.eurekster.com>>

Daarnaast kan onderzocht worden welke zoekmachines (inclusief licentiemodellen) gebruikt worden op buitenlandse overheidsportals, evenals welke zoekmachines er al in omloop zijn binnen de (Rijks-)overheid. Wellicht kan gebruik gemaakt worden van bestaande licenties. Tenslotte moeten de ontwikkelingen op het gebied van open source zoekmachines worden onderzocht.

Andere ontsluitingsmechanismen

Informatietypen

Het ontsluitingsmechanisme is vaak afhankelijk van het informatietype. Overheidsorganisaties worden ontsloten door te zoeken op bestuursniveau (Rijk, provinciaal of lokaal³⁰), op macht (wetgevend, uitvoerend, rechterlijk³¹), geografisch (postcode, GIS³²), thematisch of alfabetisch. Producten kunnen stapsgewijs gevonden worden: via levensgebeurtenis en criteria of via thema, vraagpatroon en doelgroep³³. Vaak werken bepaalde mechanismen beter voor deelverzamelingen. Zo komt GIS op als ontsluitingsmechanisme voor bijvoorbeeld locatiegerelateerde zaken, zoals milieu-eisen en bestemmingsplannen. Veelgebruikte mechanismen zijn ook het verwijzen naar gerelateerde items (o.a.) Amazon en narrow down mechanismen, waarbij de gebruiker een aantal criteria opgeeft en daarmee de zoekresultaten afbakt (bijvoorbeeld Vacanselect³⁴

De meest bekende zoekmachines maken het onderscheid in informatietypen (webpagina's, afbeeldingen, nieuwsberichten, etc.) op de interface en niet op de resultatenpagina³⁵. Ze laten dus de gebruiker kiezen naar welke soort informatie hij op zoek is. Overheid.nl maakt dit onderscheid momenteel niet. De resultaten uit het Overheidsloket (producten), de Staatsalmanak (personen en organisaties) en de Nieuwsbank (persberichten) en webpagina's worden op dit moment samengevoegd. De resultaten kunnen in de verbeterde zoekmachine per informatietype op de resultatenpagina geclusterd worden. Indien het aantal applicaties toeneemt, zal de resultatenpagina onoverzichtelijk worden. Aanpassing van de interface (tabbladen, links of radio buttons) is dan noodzakelijk.

Momenteel wordt alleen tekst ontsloten. In de toekomst moeten misschien ook afbeeldingen, audio, video, grafieken, tabellen en kaarten ontsloten worden. Dit moet uit een content-analyse blijken.

³⁰ FirstGov, <<http://www.firstgov.gov/Agencies.shtml>>

³¹ FirstGov, <<http://www.firstgov.gov/Agencies.shtml>>

³² Bund.de, <<http://www.bund.de/Service/Geosuche-.5565.htm>>, Add-wijzer <<http://www.addwijzer.org/>>

³³ VIND of OPUS, bijv. op websites van resp. gemeente Breukelen <<http://www.breukelen.nl/>> en gemeente Zoetermeer <<http://www.zoetermeer.nl/>>

³⁴ Vacanselect – zoeken op bestemming <<http://www.vacanselect.nl>>

³⁵ Nieuw-Zeeland <<http://www.govt.nz/>>

Personalisatie

Een portal biedt diverse mogelijkheden voor pro-actieve dienstverlening. Op AmsterdamMail en Google News Alerts³⁶ kan een gebruiker zich op nieuwsberichten abonneren. In Canada kan dit met updates van de websites³⁷.

In Singapore gaat men nog verder. Gebruikers van My.ecitizen³⁸ kunnen voor diverse elektronische diensten notificaties ontvangen per SMS of e-mail.

Bezoekers van Denmark.dk kunnen de zoekresultaten opslaan.

Mijn.overheid.nl is nog niet mogelijk (geen eigen content) of ongewenst (taak van de markt).

Communityvormen

Via bijvoorbeeld een forum kunnen bezoekers elkaar helpen bij het vinden van informatie of antwoorden op vragen³⁹. Op commerciële sites wordt dit middel bijvoorbeeld ingezet zodat gebruikers lijstjes van hun favoriete producten kunnen bijhouden (Amazon ListMania⁴⁰). Veel gebruikt is ook het laten stemmen op de waarde van informatie door bezoekers (bijvoorbeeld de kwaliteit van links, reviews).

De meeste van deze technieken zijn relatief goedkoop. Daartegenover staat de mogelijkheid tot misbruik en de onmogelijkheid van het controleren op de kwaliteit van informatie. Het gaat uiteindelijk wel om de overheid als betrouwbare informatiebron en om Overheid.nl als neutrale partij in het ontsluiten van informatie. Desalniettemin is het interessant om verder te onderzoeken en experimenten op te zetten.

Directgov: de inzet van content en doelgroepen

Directgov⁴¹ is/wordt de opvolger van UK Online. We noemen deze site afzonderlijk, omdat dit portal fundamenteel van andere overheidsportals verschilt. Het portal kan 90% van de zoekopdrachten direct (op de site zelf) afhandelen. De site biedt veel eigen content in plaats van externe links. De departementen vullen ieder een deel van het portal ('franchising'). Een bijkomend voordeel is de consistentie ('Common Look and Feel').

Opvallend is het grote aantal doelgroepen (automobilisten, ouders, senioren, etc.). De indeling in levensfasen is vervallen. Verder heeft het portal een context-gevoelige zoekmachine (die ook vragen in natuurlijke taal begrijpt), FAQ, een helpdesk, links naar derden, een alfabetische trefwoordenlijst en een 'subject directory'. De informatie is ook beschikbaar via digitale TV ('multi-channeling').

Uit een recente survey bleek dat het merendeel van de ondervraagden zeer tevreden is over Directgov⁴².

Advies overheid.nl kan een aantal zaken in haar visie-ontwikkeling meenemen: de indeling in doelgroepen, de hoeveelheid content, de mate van samenwerking, links naar derden en multi-channeling.

³⁶ AmsterdamMail <<http://www.amsterdammail.nl/>> en Google News <<http://www.google.com/newsalerts?hl=en>>

³⁷ Canada <<http://www.canada.gc.ca/canada/SCS?!=1&en=yes>>

³⁸ My.ecitizen <<http://www.ecitizen.gov.sg/MyeCitizen/index.htm>>

³⁹ Experts-exchange <<http://www.experts-exchange.com/>>

⁴⁰ Amazon <<http://www.amazon.com>>

⁴¹ Directgov <<http://www.direct.gov.uk/>>

⁴²

<http://europa.eu.int/ISPO/ida/jsps/index.jsp?fuseAction=showDocument&documentID=2494&parent=chapter&preChapterID=0-140-194>

Marketing/Distributie

De meest eenvoudige vorm van distributie, is het plaatsen van een klein zoekvenster op een andere site. Het aantal gemeenten⁴³ en portals dat nu al een zoekbox van Overheid.nl op hun site heeft, is klein en kan verder uitgebreid worden. Deze distributievorm kan als marketinginstrument gezien worden.

Een andere beperkte vorm van distributie is het plaatsen van RSS-feeds (bijv. met koppen van recente persberichten of kamerstukken)⁴⁴. Voor professionele gebruikers kan een browser-toolbar ontwikkeld worden. Dit zijn kleine, eenmalige investeringen.

Bij een verdergaande vorm kan de zoekmachine als Shared Service worden opgezet. Hierdoor hoeven overheidsorganisaties geen eigen zoekmachine meer te hebben, maar gebruiken zij de zoekmachine van Overheid.nl. Dan kan ingesteld worden dat de zoekresultaten van de eigen site als eerste worden getoond. Op de interface kan de gebruiker aangeven of hij/zij op de site wil zoeken of "in de hele overheid". Overheid.nl is dan een soort application service provider. De zoekmachine van FirstGov (die gebruik maakt van FAST) is bijvoorbeeld ook op de sites van het ministeries van Landbouw en Buitenlandse Zaken te vinden⁴⁵. Verdere integratie is mogelijk door bijvoorbeeld binnen dossiers van ministeries rechtstreeks naar zoekresultaten van de zoekmachine te verwijzen, die continu geïndexeerd worden. Een deel van het dossier wordt hiermee automatisch samengesteld.

Indien de resultatenpagina goed in de site geïntegreerd kan worden en de site goed geïndexeerd kan worden, willen andere overheidsorganisaties misschien van zo'n dienst gebruik maken. De behoefte en wenselijkheid dienen te worden onderzocht.

Metadatastructuren

Veel landen en overheidsorganisaties werken aan structuren gebaseerd op metadata zoals thesauri, taxonomieën en topic maps. We gaan hier kort in op een aantal vormen.

Soorten metadatastructuren

Gartner onderscheidt vijf soorten metadatastructuren geordend op toenemende complexiteit en mate van semantiek.

1. woordenlijsten: indices, glossaries, dictionaries
2. classificatiesystemen: taxonomieën, subject heading, hiërarchieën
3. semantische representaties: knowledge maps, semantische netwerken, thesauri, topic maps
4. kennisgebaseerde representaties: case-based reasoning, causale modellen, KRL
5. ontologieën: SHOE, RDF, DAML+OIL, Ontology Exchange Language

Taxonomie

Een taxonomie is een concept waarbij op basis van een hiërarchische classificatie (met gebruikmaking van onderlinge relaties) en applicatie (ICT-tool) de mogelijkheid wordt gecreëerd om interne en externe informatie vanuit het gebruikersperspectief (of corporate gezichtspunt) te ontsluiten en waarbij hulpmiddelen voor navigatie zorgdragen voor een snelle retrieval (bron: Reekx) of simpel gezegd een hiërarchische structuur van termen en categorieën.

⁴³ Gemeente Roermond <<http://www.roermond.nl/index8653.htm>>, MSN <<http://www.msn.nl/>>

⁴⁴ NU.nl <<http://www.nu.nl/rss/>>

⁴⁵ FirstGov <<http://www.firstgov.gov/>>, USDA <<http://www.usda.gov/>>, US Dept. of State <<http://www.state.gov/>>, FAST <<http://www.fast.no/>>

Met behulp van de taxonomie en op basis van karakteristieke termen in documenten, kunnen documenten handmatig of automatisch geclassificeerd of geclusterd worden ⁴⁶. Dossiers kunnen automatisch worden gevormd, wat veel meerwaarde kan bieden (bijv. van actuele thema's)⁴⁷.

Een taxonomie in de vorm van bijvoorbeeld een trefwoordenboom, biedt burgers een gebruiksvriendelijke navigatiemethode⁴⁸.

Domeinexperts dienen bij de ontwikkeling van de taxonomie betrokken te worden.

Thesaurus

Een thesaurus is een alfabetisch geordende gestructureerde en gecontroleerde lijst van termen met hun onderlinge (hiërarchische, associatieve en/of semantische) relaties, die wordt gebruikt bij het toegankelijk maken en raadplegen van (geautomatiseerde) informatie (bron: WODC). De gebruiker kan zoeken op nauwere (meer specifieke), bredere (meer algemene) of verwante termen ⁴⁹. Advies.overheid.nl beschikt wellicht niet over de middelen om zelf een thesaurus te ontwikkelen of te beheren en zal dus op zoek moeten gaan naar bestaande thesauri (bijvoorbeeld van de Tweede Kamer). De zoektermen in de logbestanden kunnen mogelijk gebruikt worden om deze thesaurus te verfijnen.

Topic Map

Een topic map kan dienen als navigatielaag bovenop een informatieverzameling. Topic maps zijn geschikt voor portals zoals Overheid.nl, omdat ze informatie uit een grote hoeveelheid heterogene bronnen kunnen combineren. Een topic map identificeert de concepten die in de informatie liggen opgeslagen, omschrijft de tussenliggende relaties (context) en legt een verbinding tussen de concepten en onderliggende informatie. Topic maps vormen een conceptuele laag waarin kennis uit en over documenten onafhankelijk van die documenten kan worden vastgelegd.

Een topic map kan als een elektronische index van gecontroleerde termen (bijv. uit een thesaurus) fungeren ⁵⁰. Het indexeren kan ook zonder kennismodel, met medewerking van domeinexperts. De bouw en het onderhoud van een topic map is vrij kostbaar. Topic maps worden voornamelijk in een professionele omgeving gebruikt.

Indien voor een bepaald domein een thesaurus beschikbaar is, kan Advies.overheid.nl op den duur een experiment met de bouw en het gebruik van topic maps opzetten.

Ontologie

Een ontologie is een classificatiesysteem voor begrippen en hun onderlinge verbindingen binnen een bepaald kennisdomein. Op middellange termijn zijn ontologieën interessant om mee te experimenteren. De ontwikkelingen zijn nog in een pril stadium.

Op het gebied van semantische structuren liggen vrijwel zeker mogelijkheden om het zoekproces te verbeteren. Echter, aan de beheerkant en technologie zijn nog wel wat uitdagingen te overwinnen. In het onderzoek naar ontwikkelingen op dit gebied gaan we in eerste instantie uit van het gebruikmaken van bestaande initiatieven, in plaats van eigen ontwikkeling.

⁴⁶ BMWA <<http://www.bmwa.gv.at/>>

⁴⁷ NOVA Dossiers <<http://www.novatv.nl/>>

⁴⁸ Edulex <<http://www.ond.vlaanderen.be/edulex/database/default.asp>> en State of Connecticut <http://www.ct.gov/ctportal/taxonomy/ct_taxonomy.asp>

⁴⁹ Overheid Nieuw-Zeeland (zie zoekresultaten) <<http://www.govt.nz/>> en Canada <<http://en.thesaurus.gc.ca/>>, WODC <<http://www.wodc.nl/>>

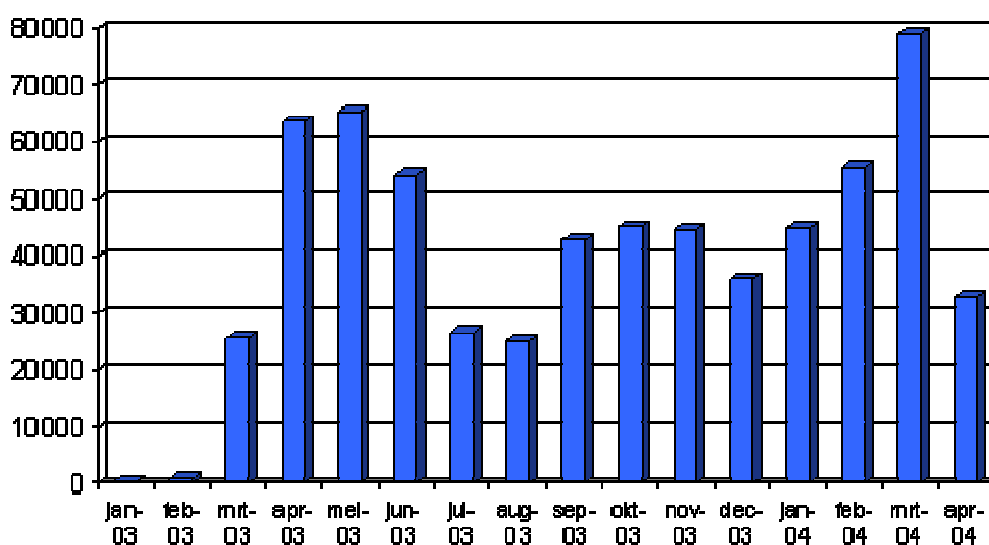
⁵⁰ Omnigator <<http://www.ontopia.net/omnigator/>>

Bijlage 3: Query-analyse en Quick wins MZD

In de aanloop naar het (deel)projectplan is een analyse gedaan van de huidige zoekmachine en van het zoekgedrag dat het afgelopen jaar heeft plaatsgevonden.

De logbestanden van de Metazoekdienst beslaan de periode van 23 januari 2003 t/m 17 april 2004. De eerste twee maanden zijn slechts honderden zoekopdrachten uitgevoerd. Een zoekopdracht wordt vaak enkele keren achterelkaar uitgevoerd. Een zoekactie die binnen een bereik van 20 records meerdere malen is uitgevoerd, wordt als één zoeksessie geteld.

Van de 979.100 zoekopdrachten (records) blijven dan 640.518 zoeksessies over, waarvan 259.210 uniek. De drieduizend meest voorkomende (unieke) zoekopdrachten vormen veertig procent van het totaal; 8.900 de helft. De zogenaamde 80% wordt pas behaald bij 131.219 (unieke) zoekopdrachten.

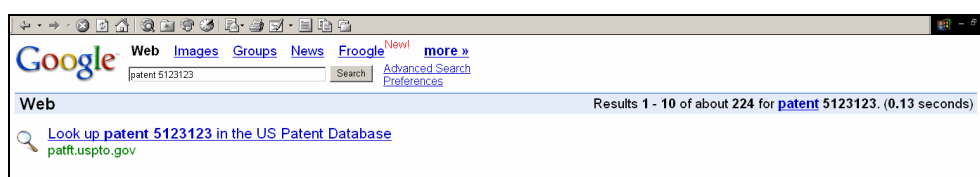


Figuur 1: Aantal zoeksessies per maand

Keyword handling

Door bepaalde trefwoorden af te vangen, kunnen bezoekers sneller worden doorverwezen.

1. **Kamerstukken:** Indien iemand op een getal tussen de 20.000 en 30.000 zoekt, wordt waarschijnlijk het publicatienummer van een Kamerstuk bedoeld. Ook bij Google is het mogelijk om op een getal te zoeken.



2. **Publicaties en publicatiesoorten:** direct verwijzen naar documenten (de Miljoenennota, het Belastingplan, de Grondwet, het Hoofdlijnenakkoord, de Troonrede, het Staatsblad, de Staatscourant, het Tractatenblad, de Rijksbegroting, het Regeerakkoord, etc.) of databanken (AMvB's, besluiten, kamerstukken, kamervragen, memories van toelichting, officiële publicaties, parlementaire stukken, verdragen, wetboeken, wet- en regelgeving, wetsvoorstellen, wetten, CAO's, etc.).
3. **Adressen:** link naar de Almanak
4. **Veel gevraagde producten uit de RPC:** algemeen (subsidies, belastingen en vergunningen) en specifiek (bijstand, identiteitsbewijs, kinderbijslag, paspoort, rijbewijs, successierecht, studiefinanciering, Zalmsnip, etc.)
5. **Afkortingen van wetten** (zie bijlage wetten.xls)
6. **(Afkortingen van) organisaties** (zie bijlage organisaties.xls)
7. **Link naar Dienst Domeinen**
8. **Veel gevraagde content** (kaart van Nederland, een kalender met vakanties en feestdagen, vlag(-instructie) en volkslied, stages en traineeships)
9. **Dossiers van ministeries** (zie bijlage dossiers.xls)
10. **Koppeling met thema-pagina's** (zie bijlage themas.xls)

Deze trefwoorden vormen 45,5% van het totale aantal zoeksessies (640.495 - 348.988 = 291.507).

In minder dan één procent van de zoeksessies wordt met (booleaanse) operatoren gezocht. Het zoeken op postcode wordt bijna nooit gebruikt.

Overige aanbevelingen

- Resultaten clusteren/splitsen: personen, organisaties, producten en diensten, (actuele) persberichten, webpagina's;
- Controleren op HTTP-statuscodes (403, 404, 500) en foute pagina's uit cache verwijderen;
- Controleren op (hexadecimale) user-id's in de URL (De Milieulijst 2003 van VROM is tienduizenden keren geïndexeerd);
- Het negeren van voorzetsels leidt (ook) tot relevante resultaten;
- Bij documenten het bestandsformaat toevoegen (bijv. [PDF]);
- Bij geen zoekopdracht <mzd:query/> in het tekstvak verwijderen;
- De breedte van een header-cel is hard gecodeerd (810 pixels) waardoor de layout niet meer klopt;
- Postbus 51 indexeren en integreren.

Een aantal aanbevelingen komt overeen met het gebruikersonderzoek dec. 2003.

Bijlage 4: Detailbegroting en -planning