



Klimaatverandering in de informatiehuishouding?

Van ongestructureerde content naar gestructureerd zoeken.

SEPTEMBER, 2016



Een informatieberg is net als een ijsberg. Je ziet alleen het topje, maar driekwart of meer van de berg is onzichtbaar. De bovenkant van de ijsberg is met wat voorzichtigheid nog te bedwingen. Door goed te kijken, op te letten en soms overnieuw te beginnen komen we wel op de juiste bestemming. Voor de informatieberg geldt dat ook, de top met de nieuwste documenten kunnen we nog wel bedwingen. Daar vinden we nog wel wat we zoeken. Maar de oudere informatie die onder de top zit is ondoorzichtig, niet meer te herkennen en niet veel meer dan een enorme stapel (digitaal) papier. Is een Enterprise Search toepassing de oplossing voor het verkennen van al de verschillende informatiebergen?

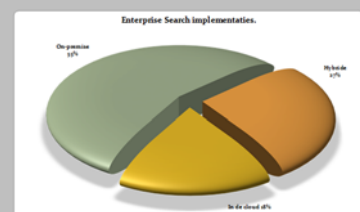
■ Enterprise Search

Het informatievolume groeit exponentieel en voor de ontsluiting daarvan krijgt Enterprise Search een hoogste prioriteit op de ICT-kalender. Dit omdat in de berg wellicht de nieuwe gouden speld kan liggen of omdat informatie uit de berg gewoon van essentieel belang is voor de bedrijfsvoering. Het aantal ES-toepassingen dat de laatste maanden (jaren) op de markt beschikbaar kwamen is aanzienlijk. Een belangrijke eerste (h)erkende toepassing was de Google Search Appliance. Deze gele doos kon als een soort van plug-in in het netwerk worden geschoven en indexeerde dan alle intern beschikbare of bereikbare digitale content. Hiermee werd als het ware een interne Google Search mogelijkheid gecreëerd. Naast de hoge aanschaf- en gebruikskosten had de toepassing ook het nadeel van een algemene Google indexmethodiek. De te indexeren informatie moest direct bereikbaar zijn. Eenvoudig voor documenten op een netwerkschijf maar lastig voor documenten die verborgen zitten in een document managementsysteem of elektronisch archief. Het zoeken van de informatie via de indexatie van alle content was een tweede nadeel. Een hitlijst van enkele honderden voor het zoeken van een vakantie is wellicht prima, maar voor een cruciale beslissing in een bedrijfsproces is dat geen optie.



Google Search Appliance

Google heeft begin 2016 het einde van de Google Search Appliance aangekondigd. In 2018 wordt de toepassing volledig beëindigd. De opvolging bij Google ligt in de cloud. Via een Google Spider functie kan alle digitale content geïndexeerd worden. De opslag en ontsluiting is via een cloud functie. Waarschijnlijk voor veel organisaties in Nederland een brug te ver. Het opslaan van de bedrijfskritische content in de cloud, nog maar even niet!



■ Relevantie

Recent onderzoek wijst uit dat een informatieberg bestaat uit 11 procent aan bedrijfskritische documenten, 44 procent aan overbodige documenten en 45 procent aan onbekende documenten. De 44 procent overbodige documenten zijn veelal dubbel, verouderde of zakelijk niet meer relevante documenten. Op zich leveren deze categorie documenten geen direct probleem op. Naast opslagruimte die ze innemen en alle kosten voor beheer en back-up, is het bedrijfsrisico gering. Het echte probleem zit hem in de documenten waarvan onbekend is wat de inhoud is en dus ook onbekend is wat de waarde of het risico is voor de organisatie. De waarde kan uitgedrukt worden in meerdere factoren;

- Waarde vanuit een nationaal belang. Het is van historisch belang dat deze documenten worden bewaard en kunnen worden ontsloten.
- Waarde vanuit een klantbelevens perspectief. Het hebben van een totaalbeeld of compleet dossier in alle klantcommunicatie helpt in de optimale klantbelevens. Ook de Customer Self Service en het voorkomen van vragen aan het callcenter wordt hiermee ondersteund. Een voorbeeld hiervan is het on-line beschikbaar stellen van de rekeningafschriften. De eerste bank begint nu met het beschikbaar stellen van de rekeningafschriften tot 6 jaar terug. Zonder extra kosten!
- Waarde vanuit een wettelijk perspectief. Wet- en regelgeving zijn vaak niet eenduidig en vaak onbekend. Zo is bekend dat facturen 7 jaar bewaard moeten worden. Maar vanaf welk moment gaat de 7 jaar in? Wet en regelgeving rondom het bewaren van identiteitsbewijzen, curriculum vitae 's, en polissen is vaak onbekend. Een belangrijk wettelijk perspectief is ook het kunnen reproduceren van bijvoorbeeld pensioenberekeningen. Het overzicht wordt wel bewaard, maar de redenatie achter de berekeningen niet.
- Waarde vanuit een strategisch belang. Het gaat dan niet zozeer om het individuele document maar om de aggregatie ervan. Welke trends kunnen we ontdekken als we naar de inhoud van de documenten kijken?

Doelstellingen ES

Een recent onderzoek uitgevoerd door Forrester in opdracht van Mindbreeze laat ook zien dat Enterprise Search voornamelijk wordt ingezet voor de ondersteuning naar ongestructureerde content (digitale) documenten om de medewerker te ondersteunen in zijn dagelijkse werk. De belangrijkste voordelen van een Enterprise Search toepassing waren dan ook;

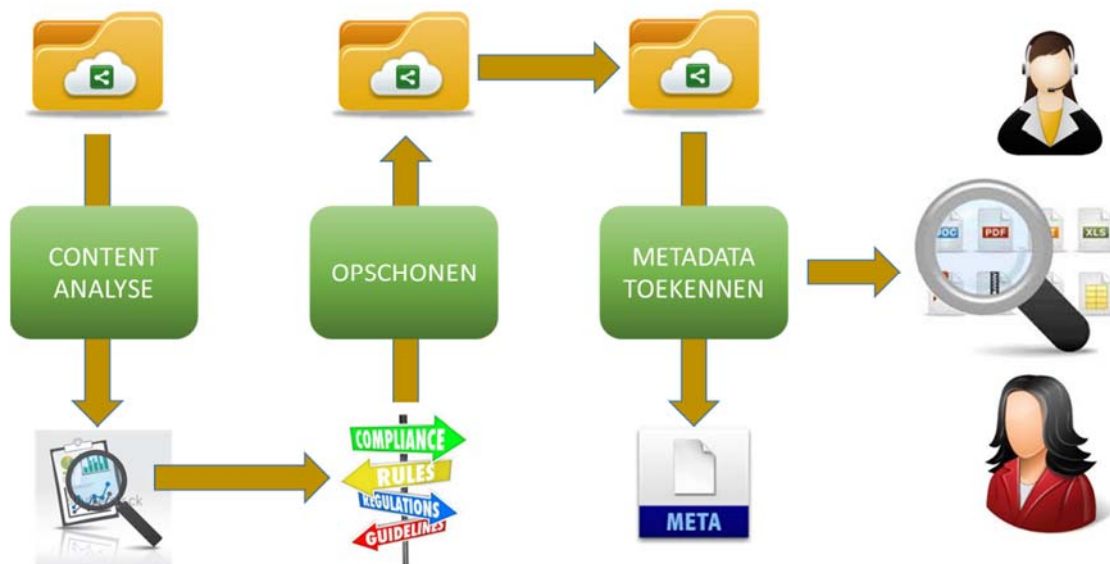
- 74 procent van de ondervraagde IT-managers gaf aan dat het bedoeld was voor de medewerker om klanten beter te kunnen bedienen.
- 62 procent gaf aan dat het de medewerker in staat stelde om sneller te kunnen werken doordat de relevante informatie sneller beschikbaar was.

■ Informatiebergen

Zoals al in het begin geconstateerd wordt het ontsluiten van ongestructureerde informatie in de vorm van documenten een steeds belangrijker aandachtspunt op de ICT-projectkalender. Het start veelal met het ontsluiten van de digitale documenten op de netwerkschijven. Via een indexeertoepassing wordt de inhoud doorlopen en ontsloten. De document management toepassing maakt daarnaast nog gebruik van een eigen (proprietary) zoektoepassing. Ook de bedrijfstoepassingen hebben weer hun eigen methodiek. Het aantal informatiebergen is de afgelopen jaren flink gegroeid en groeit nog aardig door. Een gemiddelde onderneming heeft al snel een tiental individuele verzamelingen. Elk met hun eigen methodiek in ontsluiting via autorisaties, metadatastructuren en documentformaten. Kan de Enterprise Search toepassing wel overweg met al deze diverse structuren en autorisaties? Zo ja, wat is de inspanning om dat ingeregeld te krijgen?

■ Aanpak

Een Enterprise Search toepassing loslaten op alle digitale content die op fileshares, document management toepassingen en elektronische archieven beschikbaar is niet de gewenste aanpak. Hoe geavanceerd de Enterprise Search toepassing ook is. Deze is niet instaat om de relevantie te bepalen van de gevonden documenten die aan de zoekcriteria voldoen. Lastiger wordt het zelfs wanneer de Enterprise Search toepassing resultaten toont van documenten die daar eigenlijk nooit hadden mogen staan. De documentberg zal toch eerst bekeken moeten worden, opgeschoond (ontdaan van alle overbodige documenten) en alle onbekende content geanalyseerd naar waarde, inhoud en kenmerken.



■ Eerst analyseren

De analyse is de noodzakelijke investering die het mogelijk maakt dat een Enterprise Search toepassing optimaal kan worden ingezet. Door deze investering te doen worden de medewerkers ook enthousiast over de mogelijkheden die een dergelijke toepassing biedt in hun dagelijkse werk. Het eindresultaat is dan weer dat medewerkers gaan inzien hoe belangrijk het is om hun digitale content juist te gaan opslaan. Juist is niet alleen op de juiste plaats binnen het archief, netwerkschijf of DMS. Juist in deze is ook voorzien van de juiste metadata kenmerken.

■ Migrato

Migrato levert de techniek en diensten voor het analyseren van ongestructureerde content. Migrato heeft een unieke manier ontwikkeld waarmee ongestructureerde content inhoudelijk kan worden geanalyseerd en geclassificeerd. De inhoud van het document is daarbij één van de parameters om het document te kunnen classificeren. Naast de inhoud wordt ook gekeken naar vorm, grootte, en specifieke tekstposities. Deze methodiek garandeert een herkenning van 90 procent op de digitaal beschikbare content en een maximale foutenmarge van 1 procent.

■ Aanbevelingen

Gaan zoeken in alle ongestructureerde content via een Enterprise Search toepassing levert niet de verwacht winst voor de medewerkers. Het levert waarschijnlijk meer risico's op voor de bedrijfsvoering dan nut. Beslissingen nemen op een oude versie van een document liggen dan voor de hand.

- **Verwijder de overbodige content.** Het opschonen van oude, overbodige of niet meer relevante content kan prima via een content analyse. Bestanden worden dan via diverse bedrijfsregels gecontroleerd op bijvoorbeeld ouderdom, laatst gebruik, versies en inhoud.
- **Controleer de onbekende documenten.** De inhoud van de nu nog onbekende documenten kan cruciaal zijn voor de compleetheid van een dossier. De set aan onbekenden zal moeten worden geanalyseerd en gemigreerd en/of geconverteerd.
- **Voeg de juiste metadata toe.** Zoeken op alleen een index van de inhoud levert voor een zakelijke toepassing onvoldoende resultaat. Een ondersteuning van de medewerker via een (minimale) set aan metadata is gewenst. De Enterprise Search toepassing moet instaan zijn deze aparte metadata te verwerken.

Over Migrato

Migrato is gespecialiseerd in de migratie van content uit gebruikte of oude ECM-systemen naar een nieuw ECM-systeem. Doelstelling is om de content explosie met ongestructureerde en ongecontroleerde opslag van content gereed te maken voor de volgende fase van informatievoorziening. De migratie wordt goed voorbereid met de deskundigen uit de eigen organisatie en omvat niet alleen het overzetten van content met de bijbehorende meta-data naar het nieuwe systeem, maar neemt ook functionaliteit mee als, verrijken van meta-data, opschoning, aanvullen van meta-data (voor bijvoorbeeld Records Management) en formaatconversies. Zo krijgt de content in de nieuwe omgeving meer waarde voor de medewerker. Content is makkelijker en sneller te vinden, de organisatie voldoet weer aan de compliance eisen en een klantdossier is weer compleet. Zo wordt de organisatie een echte kennisorganisatie.

Voor het uitvoeren van migraties gebruikt Migrato een specifiek ontwikkelde methodiek. Deze zorgt ervoor dat migraties via een strak geleid proces worden uitgevoerd. Medewerkers worden daarbij zo min mogelijk belast of lastiggevallen tijdens de conversie. Ze kunnen van de ene op de andere dag over naar het nieuwe systeem.

De medewerkers van Migrato hebben een jarenlange ervaring en expertise in ECM implementaties. Zo waren wij onder andere betrokken bij de eerste ECM-implementaties in Nederland binnen de farmaceutische industrie, bij de overheid en in de financiële sector. Onze implementatie experts hebben de expertise wat betreft de inrichting van de oudere ECM toepassing, en ook voldoende kennis van de nieuwe toepassingen met de functionele mogelijkheden en de wet en regelgeving.

Migrato helpt ook u bij de eerste stappen naar een betere informatiehuishouding en vernieuwde digitale archivering. Neem contact met ons op voor een vrijblijvend gesprek over de mogelijkheden.

E-Mail info@migrato.nl

Website www.migrato.nl

Afkortingen en Begrippenlijst

ECM – Enterprise Content Management

EIM – Enterprise Information Management

ES – Enterprise Search

DIV – Documentaire Informatie Verwerking

Disclaimer

Complete document. © 2016 by, "Migrato B.V.". All rights reserved. This publication may not be reproduced or transmitted in any form without prior permission from the copyright holder.

The information contained herein has been obtained from sources believed to be reliable and has been prepared with professional diligence. However, Migrato B.V. disclaim all warranties as to the accuracy and completeness of this information. Migrato B.V. its directors and employees, can accept no liability for any damages or loss occasioned to person or entity for errors or omissions in the information contained herein or for the interpretations thereof. The opinions expressed herein may change at any time without notice.

Migrato B.V.,

Nieuwe Haven 39, 3291 GA Strijen, Nederland.

Tel: (+31) 6 507.34.796;

Email: info@migrato.nl

Zie ook: www.migrato.nl

