

XML belangrijk bij uniforme toegang contentbronnen

Over de grenzen van contentmanagement

Hoewel documentmanagement- en archieffuncties vroeger twee gescheiden werelden waren, raken de eigentijdse verschijningsvormen content- en recordsmanagement steeds vaker met elkaar verweven. Dit blijkt ook uit het productaanbod op de markt voor contentmanagement waarin meer en meer recordmanagement-faciliteiten zijn ingebouwd. Specifieke recordmanagement-applicaties bieden bovendien steeds vaker typische contentmanagementfuncties. XML zal dit integratieproces als verbindende technologie verder versnellen.

De introductie van de PC en de daarmee gepaard gaande digitalisering betekende een omslag in documentbeheer. Auteurs werden meer dan voorheen zelf eindverantwoordelijk voor het opmaken en printen van documenten. De opkomst van het internet zorgde daarna voor een nog grotere omwenteling in documentmanagement. Het vertrouwde document kreeg concurrentie van nieuwe informatiedimensies waarbij flexibele combinaties van vorm en inhoud informatieproducten op maat moesten opleveren zoals dynamische en gepersonaliseerde websites. Ook multimediale inhoud zoals audio en video zorgde voor verdere afbreking van het traditionele documentbegrip. Content- en webcontentmanagement (CM en WCM) werd geboren zodat naast documenten ook andere content goed kon worden beheerd.

De groeiende aandacht voor intranet en met name het fenomeen van de 'information portal' leidde opnieuw tot uitdagingen. Enterprise contentmanagement (ECM) probeerde een

antwoord te geven op de vraag hoe bedrijven alle content, zoals e-mail, webcontent, database-informatie én documenten, uit diverse systemen en toepassingen geïntegreerd en organisatiebreed kunnen aanbieden en beheren.

Recordsmanagement

Hoewel CM in belangrijke mate antwoord geeft op nieuwe beheervragen als gevolg van de verdergaande digitalisering, heeft het systeem ook zijn beperkingen. Zo ondersteunt

contentmanagement vooral de dynamische, actieve levensfasen van een informatieproduct. Op een zeker moment moet echter worden bepaald wat er met het informatieproduct gebeurt na zijn actieve leven. Dat moment dient zo ver mogelijk stroomopwaarts in het proces te liggen zodat al vroeg bekend is of het informatieproduct een archiefstuk is. Zeker voor overheidsinstellingen maar ook voor veel bedrijven geldt immers dat zij hun handelen altijd moeten kunnen verantwoorden. Daarom is het van belang dat de neerslag van dat handelen tijdig wordt geborgd in documenten en informatiebestanden. Hiervoor kan een wettelijke grondslag zijn zoals de Archiefwet of de Belastingwet, maar ook een juridische grondslag. Als er onverhoopt een geschil over dat handelen ontstaat, kunnen de op de transactie betrekking hebbende informatiebronnen immers als bewijs dienen. Hieruit blijkt dat de zogeheten procesgebonden informatie een zwaarder bewaarregime vereist dan andere content. Omdat contentmanagement voor een dergelijk

Ontwikkelingen	1e	2e	3e
Contentmanagement	Van papier naar digitaal	Van document naar informatieproduct	Geïntegreerd aanbieden van informatieproducten in portal
Recordsmanagement	Van separaat archiefproces naar integratie in bedrijfsproces	Rollback, versiebeheer en diverse conserveringstechnieken voor 'dynamische' informatieproducten	Standaardisatie opslagformaat (XML). Centraal afstemmen archiefregime

Afbeelding 1. Enkele verschillen tussen contentmanagement en recordsmanagement.

Recordsmanagement-uitdagingen

Enterprise contentmanagement kent een aantal recordsmanagement-uitdagingen, zoals:

- > Beheer van documenten die zijn voorzien van een elektronische handtekening, zeker wanneer die een in de tijd beperkte geldigheid heeft;
- > Vastleggen van een samenhangende e-mailcorrespondentie waarbij niet alleen het bericht maar ook eventuele attachments,

reacties en naschriften worden bewaard;

- > Registreren van transacties via websites die volledig papierloos worden afgewikkeld;
- > Archivering van webinhoud voor de langere termijn vooral bij aanwezigheid van dynamische formaten zoals audio- en video-content, verwijzingen naar content die door derden wordt beheerd en integratie van externe contentbronnen.

zwaar archiefregime tot op heden vaak onvoldoende houvast en faciliteiten biedt, komt hier Recordsmanagement (RM) om de hoek kijken. Beide disciplines vullen elkaar dan ook goed aan.

Uitdagingen

Traditioneel RM stond vroeger dicht tegen de archieffunctie aan. Het vaak in de kelder weggestopte archiveringsproces werd doorgaans los van de primaire processen uitgevoerd. De voortgaande digitalisering heeft echter ook het stof van recordsmanagement afgeblazen. Daarbij zijn de ontwikkelingen op het gebied van recordsmanagement te spiegelen aan die van contentmanagement. Zo veranderde door de komst van het digitale document de traditionele levenscyclus van het papieren equivalent doordat het in beheer nemen van digitale documenten naar voren moest worden gehaald. Als een digitaal stuk niet al bij binnenkomst in verzekerde bewaring wordt genomen, kan er immers van alles mee gebeuren, tot en met ongeoorloofde wijziging en vernietiging aan toe.

De nieuwe dimensie van het informatieproduct heeft tevens voor nieuwe RM-vraagstukken gezorgd. Traditionele digitale documenten hebben

nog altijd een vaste vorm- en inhoudcombinatie waardoor de authenticiteit en originaliteit relatief eenvoudig kunnen worden gegarandeerd door het digitale stuk te beveiligen tegen bewerking achteraf. Vaak wordt voor dit doel ook simpelweg een papieren bewijskopie geprint. Bij informatieproducten op maat is dit veel lastiger. Een video- of audio-

Traditionele digitale documenten hebben nog altijd een vaste vormen inhoudcombinatie

fragment is niet te printen. Het vormaspect van een informatieproduct is bovendien lastig te vatten. Door verschillen in hardware, software en configuratie kan eenzelfde informatieproduct zoals een website er op hetzelfde moment bij de een anders uitzien dan bij de ander. Door de tijd heen wordt dit 'look and feel'-probleem nog groter. Door het dynamische karakter van informatieproducten zien deze er wellicht niet meer

hetzelfde uit als eerst doordat zowel de inhoud, vorm en structuur zijn veranderd. Het is daardoor de vraag of informatieproducten later nog wel reproduceerbaar zijn en hoe bedrijven kunnen garanderen dat eerdere versies van webpagina's na een wijziging bewaard blijven. Wie weet of een gehyperlinkte informatiebron over twee maanden nog wel bestaat? Het zal niet eerste keer zijn dat klanten refereren aan informatie die enige tijd geleden op een website stond, waarbij medewerkers van de betreffende organisatie zelf niet wisten dat die informatie was gepubliceerd. Laat staan dat de website van dat moment nog opvraagbaar is.

Reproducieren

Met de opkomst van het informatieproduct ontstond de behoefte om zowel de content als bijbehorende context van een bepaald moment in bevroren toestand te kunnen reproducieren. Om in die behoefte te voorzien zijn in de loop der jaren verschillende oplossingen bedacht. De eerste categorie oplossingen poogt daarbij, getrouw aan het traditionele document, de gehele originele content en context van informatieproducten te bewaren. Dit leidde onder andere tot contentbeheer in combinatie met de originele software en hardware.

Tegenwoordig leeft steeds vaker de gedachte dat de originele vorm en structuur niet altijd bewaard hoeft te blijven, mits de relevante aspecten van de context maar ondubbelzinnig en betrouwbaar zijn af te leiden uit de bewaarde procesgerelateerde gegevens. Dit inzicht biedt interessante mogelijkheden om van elk dynamisch informatieproduct uiteindelijk slechts platte, statische procesdocumenten als archiefrecords over te houden. De RM-complexiteit kan daarmee wellicht weer worden teruggebracht tot documentproporties. De opkomst van de beschrijvingstaal XML heeft hier overigens nadrukkelijk aan bijgedragen.

XML

In tegenstelling tot de opmaaktaal HTML beschrijft XML vooral de structuur en betekenis van informatieproducten of documenten. De Document Type Definition (DTD) speelt hierin een belangrijke rol. Door de informatieproducten in XML te coderen en de betekenis van het coderingsstelsel vervolgens vast te leggen in een zogeheten DTD kan tevens de semantische waarde van de informatieproducten worden vastgelegd. Omdat verschillende informatieproducten ook een verschillende structuur en functie hebben, worden aan die verschillende typen informatieproducten ook verschillende DTD's toegekend.

XML is een belangrijke stap in de richting van uniforme toegang tot alle contentbronnen zoals de eerder genoemde portal-gedachte binnen contentmanagement uitdraagt. Door alle informatieproducten standaard in XML te coderen, zijn ze eenduidig te ontsluiten. De juiste interpretatie van de XML-code volgt uit de bijbehorende DTD of anders gezegd blijkt de juiste betekenis van de informatieproducten uit de context waarin ze worden gebruikt.

Het is daarbij verstandig XML al vanaf de creatie van een informatieproduct als opslagformaat te hantieren. Niet voor niets werkt Microsoft aan standaardondersteuning van XML in al haar producten en ondersteunen tegenwoordig veel CM-leveranciers XML.

Naast XML heeft ook de ontwikkeling van zogeheten directory services, centrale adresboeken, geleid tot betere mogelijkheden om relevante procesgebonden gegevens uniform en authentiek, over de systeemgrenzen heen te beschrijven en te bewaren.

Conclusie

Documentmanagement en archief-functies waren vroeger twee gescheiden werelden. De eigentijdse verschijningsvormen content- en records-

management raken echter steeds vaker met elkaar verweven. Deze verwevenheid is ook nuttig vanuit het perspectief van digitale duurzaamheid en rechtsgeldigheid. Dit blijkt ook uit het productaanbod op de markt voor contentmanagement waarin meer en meer RM-faciliteiten

geheel versmelten, waarna misschien de term life cycle recordsmanagement opduikt. De integratie van recordsmanagement in de bedrijfsprocessen heeft in ieder geval het aanzien en belang van de archieffunctie in organisaties vergroot. Niettemin valt er op dit terrein in verschillende organisaties ook nog veel te winnen. De techniek lijkt er in ieder geval klaar voor.

***Het is verstandig XML al
vanaf de creatie van een
informatieproduct als
opslagformaat te hanteren***

zijn ingebouwd. Specifieke RM-applicaties bieden bovendien steeds vaker typische CM-functies. Ondersteunende en verbindende technologie zoals XML zal dit integratieproces verder versnellen. Op termijn valt zelfs te verwachten dat beide vakgebieden

Werner Zuurbier

*Werner Zuurbier (werner.zuurbier@kbenp.nl)
is adviseur Organisatie & Informatisering bij
Koenen Baak en Partners te Voorburg.*

Contentmanagement	Recordsmanagement
Staat toe dat content wordt gewijzigd en in verschillende versies beschikbaar is	Voorkomt dat records worden gewijzigd
Staat toe dat content wordt gewist door de eigenaar/auteur en heeft een soepel bewaarregime	Verbiedt dat records worden gewist behalve in enkele strikt gecontroleerde omstandigheden en heeft een stringent archiefregime
Biedt weinig voorzieningen voor het beheer van content in de laatste fase van de levenscyclus (archivering)	Kan vereisen dat sommige content wordt gewist/vernietigd op een bepaald punt in hun levenscyclus
Heeft een opslagstructuur tot gevolg die meestal is bedacht en wordt beheerd door eindgebruikers	Heeft een opslagstructuur volgens een vaste ordening (classificatie) die wordt beheerd door de recordsmanager
Is primair bedoeld om het dagelijks gebruik van content in de lopende bedrijfsvoering te ondersteunen	Is primair bedoeld om een verzekerde en betrouwbare 'archiefbewaarplaats' te verschaffen voor procesgebonden informatie op korte en lange termijn

Afbeelding 2. Enkele parallele en samenhangende ontwikkelingen van contentmanagement en recordsmanagement.