

CASE STUDY

Automatische Metadatering Is Essentiële Operatieslag Om Kenniskloof Te Overbruggen



Probleem

Bij de geplande digitaliseringslag waren documenten minder snel vindbaar in OpenText Content Server door het ontbreken van metadata.

Oplossing

Automatische metadatering van bestaande content in OpenText Content Server.

Resultaat

Snellere toegankelijkheid van interne bibliotheek en uitbreiding van het aantal beschikbare publicaties in de online bibliotheek en zodoende meer kennis ter beschikking kunnen stellen aan de samenleving.

‘Bridging Science to Practice’ is het motto van KWR, het Nederlandse instituut voor wateronderzoek. Een van de manieren van KWR om de kennisoverdracht tussen wetenschap en praktijk te vergroten, is het publiceren van wetenschappelijke artikelen, rapporten, etc. in zijn online bibliotheek. “De hoeveelheid content en de hoge snelheid waarmee het wordt aangeboden, is niet altijd geweest zoal nu”, aldus Jonie Keessen, bibliothecaris bij KWR. De bouw van het nieuwe kantoor in 2012 werd als psychologisch moment aangegrepen om een digitaliseringslag door te voeren.

“Alle 170 medewerkers zouden in het nieuwe kantoor maximaal één meter boekenplank krijgen. Vanaf dat moment zijn wij documenten die bewaard moesten worden en waarvan wij alleen een hardcopy hadden, gaan scannen en digitaal op gaan slaan in ons centrale ECM-systeem OpenText Content Server, met andere woorden de interne digitale bibliotheek.”

● Onvolledige metadata

Nadat alle documenten waren gescand en toegevoegd in de interne digitale bibliotheek was het tijd om de OpenText omgeving verder te optimaliseren. Jonie Keessen: “De wens was om een koppeling te maken tussen de interne digitale bibliotheek en onze online bibliotheek Openbare documenten die in de interne digitale bibliotheek worden opgeslagen, worden automatisch

“Na het automatisch toevoegen van de metadata kunnen onze gebruikers veel sneller de benodigde documenten vinden.”

doorgeplaatst op onze online bibliotheek. Dat proces verliep voorheen handmatig. Om documenten doorzoekbaar te maken, moeten ze over metadata beschikken.” In de praktijk bleek dat niet alle content in OpenText over voldoende of de juiste metadata beschikte en dus niet gemakkelijk toegankelijk was. Zo was het onmogelijk om te zoeken op auteur of op combinaties van meerdere woorden. Voor dit project werd de hulp ingeroepen van KBenP en Xillio's platform.

● Automatische metadatering

De nieuw gescande bestanden en oude bestanden uit de jaren '60 en '70 werden in het project meegenomen. Met behulp van automatische metadatering werden de bestanden onder de loep genomen en metadata toegekend, zoals auteur, titel, projectnummer, expertise, documentsoort en trefwoorden. “De oudere bestanden waren vaak slecht leesbare scans waren. Ook was de opmaak van de verschillende soorten documenten in de loop der tijd veranderd”, aldus Jonie Keessen. “Daardoor was het lastig om de documenttypen te herkennen. Zo'n 200 rapporten waren helemaal niet ge-OCR-d en werden nog tijdens het project ge-OCR-d en toegevoegd.”

● Efficiënter zoekproces

Het project is in juni 2016 afgerond. Later dan vooraf was ingeschat, maar dat had te maken met de diversiteit in opmaak van de publicaties door de jaren heen, de vele soorten documenten en de 200

rapporten die nog werden toegevoegd. Jonie Keessen: “Na het toevoegen van de metadata vinden onze gebruikers veel sneller de benodigde documenten. Het zoekproces is veel effectiever. We hebben ook een hele reeks aan oudere documenten via de interne digitale bibliotheek aan de online bibliotheek toegevoegd en daarmee leveren wij nog meer kennis. Ook is het achterliggende proces sterk verbeterd. Waar ik vroeger publicaties handmatig doorstuurde naar Communicatie die ze toevoegde aan de online bibliotheek, verloopt dit nu automatisch door de koppeling tussen de interne digitale bibliotheek en de online bibliotheek.”

● Continue verbetering

Om ervoor te zorgen dat ook nieuwe publicaties goed vindbaar zijn, voegen projectmanagers bij de start van een project zoveel mogelijk metadata toe in het financiële systeem (AllSolutions). Vervolgens wordt er automatisch in OpenText CS een projectmap aangemaakt. Vergaderstukken en concept-publicaties in deze map krijgen automatisch de metadata van het betreffende project. Onderzoekers sturen nu nog hun definitieve publicaties per e-mail naar Jonie, die ze vervolgens toevoegd aan de interne digitale bibliotheek. Op hetzelfde moment worden de openbare publicaties automatisch doorgezet naar de online bibliotheek. Deze stap wordt momenteel verder geautomatiseerd met een workflow. De publicaties hoeven dan niet meer rondgemaild te worden. Zo kan KWR ondanks een kleine overhead efficiënt kennis delen met de wereld.

Over Xillio

Elke organisatie wil wendbaar zijn. Maar als je content verspreid staat over legacy-systemen, lukt dat niet. Xillio helpt grote organisaties grip te krijgen op hun enterprise content. We migreren en structureren informatie uit verouderde omgevingen naar Microsoft 365 en maken die klaar voor veilig beheer, slimmere samenwerking en AI-toepassingen.

Met onze combinatie van tooling, kennis en flexibiliteit zorgen we dat zelfs de meest complexe migraties soepel verlopen.