

ICA-AtoM op speed!

Invoermodules voor ICA-AtoM in het Gemeentelijk Documentatiecentrum van Ranst



Roeland Verhaert, vrijwilliger Documentatiecentrum Ranst

In het voorjaar van 2012 besloot het Gemeentelijk Documentatiecentrum van Ranst om het opensourcearchiefbeschrijvingssysteem ICA-AtoM in gebruik te nemen. Dit artikel gaat in op deze ingebruikname en op de ontwikkeling van aanvullende tools. Die maken het beheer van grote hoeveelheden beschrijvingen van archieven, archiefvormers en foto's mogelijk¹.

Het Gemeentelijk Documentatiecentrum van Ranst² is een kleine gemeentelijke instelling, die uitsluitend op een tiental vrijwilligers³ draait. De werkingsgelden zijn afkomstig van de gemeente en voorzien op de begroting. De collectie bestaat uit documentatie (verzamelingen van affiches, boeken, foto's, krantenknipsels, kaarten, tijdschriften, personalia,...) en privé-archieven.

Twaalf jaar geleden is het documentatiecentrum gestart met de digitale inventarisatie van deze collectiestukken in een Access-databank.

Begin 2012 wordt een stand van zaken opgemaakt. Op dat moment zijn er 32.200 beschrijvingen: 15.146 krantenknipsels, 2487 affiches, 4172 foto's, 6609 beschrijvingen van personalia (doodsprentjes en -brieven), 2889 archiefbeschrijvingen, 977 beschrijvingen van audio, video en digitale kopieën van stukken. 3634 foto's zijn in 2012 gedigitaliseerd door een jobstudente.

Heel veel, minutieus geïnterpreteerde informatie, die vreemd genoeg moeilijk doorzoekbaar was. Dit had verschillende oorzaken: het ging niet meer om een databank, maar om drie Accesdatabanken met vijftien tabellen, enkele Excels en een Word-document... Ten tweede had elke tabel een ander datamodel en ten derde waren er ook veel inconsistenties, zoals het gebruik van "voorloophuizen",

datumnotaties (nu eens dd/mm/jjjj, dan weer jjjj/mm/dd, dan weer enkel een jaartal etc.) Er was dringend nood aan standaardisatie.

Een aantal andere vaststellingen: het aantal vragen via e-mail was exponentieel gestegen maar het leeszaalbezoek sterk gedaald. Er was duidelijk behoefte aan een digitale dienstverlening. Er waren veel geïnteresseerden, maar de meesten lieten het afweten omdat ze zich altijd moesten verplaatsen wanneer

“DE CONCLUSIES: ER IS EEN SCHAT AAN INFORMATIE DIE WE NIET WILLEN OPGEVEN. MAAR ER IS NOOD AAN EEN FLEXIBEL EN GEÏNTEGREERD SYSTEEM.”

ze wilden helpen. De collectie foto's was groot, maar door het gemis aan een goed overzicht begonnen er stilaan veel dubbels op te duiken,...

De conclusies: er is een schat aan informatie die we niet willen opgeven. Maar er is nood aan een flexibel en geïntegreerd systeem, waarin we alle oude inventarissen snel en efficiënt kunnen opnemen. Bij voorkeur ook een systeem met

centrale invoer en kwaliteitsbewaking. En er moet werk worden gemaakt van ontsluiting via internet.

EEN NIEUW SYSTEEM

De zoektocht naar een nieuw systeem startte met een kleine marktverkenning en enkele archiefbeheerssystemen werden onderzocht (onder meer Adlib en Brocade). De commerciële programma's waren

voor het documentatiecentrum al heel snel te duur, maakten geen gebruik van archiefstandaarden of waren niet flexibel genoeg. Met een groter aantal medewerkers, zijn er twee mogelijkheden: meer licenties kopen, of minder vlot werken. Ook het feit dat we zouden moeten betalen om eenmalig de bestaande inventarissen te laten omzetten en daarna verder handmatig



Het opbreken van de tramsproren te Broechem, rond 1958. Gemeentelijk Documentatiecentrum Ranst, inventarisnummer NB_166.



Een processie te Broechem, 1946. Gemeentelijk Documentatiecentrum Ranst, inventarisnummer NB_162.

zouden moeten invoeren was een obstakel.

Uiteindelijk werd besloten om te kiezen voor het opensource-archiefbeschrijvingsstelsel ICA-AtoM.

De keuze voor ICA-AtoM lag eigenlijk voor de hand. Het is gratis open source, het ondersteunt alle internationale standaarden voor het beschrijven van archief (ISAD(G), ISAAR(CPF), ISDF, en ISDIAH)⁴, het is web based, kan ook de functie van beeldbank vervullen en het voorziet modules voor bulkimport (EAD, EAC, CSV, enz.). De beschrijvingen worden in een open systeem opgeslagen en zijn te allen tijde opnieuw exporteerbaar.

Weliswaar zijn een aantal functies nog niet voorzien, zoals depotbeheer of

leeszaalmodules, maar binnen de categorie van de gratis beschikbare systemen was het toch duidelijk een van de betere⁵.

De standaardfuncties voor het importeren van beschrijvingen in bulk boden een goede opportuniteit om het probleem van de retroactieve invoer van de reeds aanwezige digitale beschrijvingen aan te pakken. We besloten dan ook om te investeren in de ontwikkeling van een omzettingstool. Dit programma moest het mogelijk maken om zelf onbepaald onze inventarissen in Access-databanken en semi-gestructuurde lijsten op een efficiënte wijze om te zetten naar EAD, één van de importformaten van ICA-AtoM. Dit was een eenmalige investering, die zich snel zou terugverdienen. Inzake⁶ kreeg de opdracht om dit programma te ontwikkelen.

De auteur van dit artikel testte als vrijwilliger bij het documentatiecentrum actief mee bij deze ontwikkeling.

IN GEBRUIKNAME EN INSTALLATIE VAN ICA-ATOM 1.2

In eerste instantie moest uiteraard een host gezocht worden. Bij de medewerkers van het documentatiecentrum was onvoldoende kennis aanwezig om het programma zelf te beheren op een eigen server. Daarom werd gekozen voor externe hosting. Voordeel daarbij was dat we onmiddellijk online zouden kunnen werken. De prijzen liggen daarenboven relatief laag. We besloten kleinschalig van start te gaan en kozen voor een formule van shared hosting, maar wel met de mogelijkheid om de formule te wisselen of het pakket uit te breiden. Eens gekozen, verliep het opzetten

van ICA-AtoM (op dat moment versie 1.2) verder zonder problemen.

OMZETTING NAAR EAD/ISAD(G)

Parallel met de configuratie en inrichting van ICA-AtoM werd dan de opdracht gegeven om een omzettingmodule te ontwikkelen om Excel-bestanden om te zetten naar EAD. De belangrijkste stap hierbij was het vastleggen van een invoersjabloon op basis waarvan de geautomatiseerde omzetting naar EAD zou gebeuren. Vertrekpunten voor deze oefening waren enerzijds de bestaande beschrijvingen en anderzijds de velden in de EAD-norm. Dit was een relatief moeilijke oefening omdat de bestaande Access-tabellen van het documentatiecentrum geen vaste structuur hadden.

Uiteindelijk werd een invoer-

sjabloon vastgelegd, met een aantal verplichte velden: Bestandsidentificatie - Titel - BeginDatum - BeschrijvingsNiveau - OmvangMedium. Deze velden zijn ook volgens ISAD(G) verplicht. De meeste andere ISAD(G)-velden werden optioneel beschikbaar gesteld. Op basis van dit sjabloon⁷ lieten we de omzettingstool ontwikkelen. Het genereren van XML-documenten die voldoen aan de EAD-norm bleek (voor Inzake) op zich niet moeilijk, maar de ware test was de import van de EAD-documenten in ICA-AtoM. En dit laatste bleek uiteindelijk een intensief proces van trial-and-error. Er was en is weinig of geen documentatie over de EAD-implementatie in ICA-AtoM beschikbaar en de mapping van enkele EAD-velden naar ISAD(G) in ICA-AtoM leverde enkele verrassende resultaten op: enkele EAD-velden worden genegeerd, andere EAD-velden werden dan weer naar het verkeerde ISAD(G)-veld weggeschreven. Ook de EAD-export vanuit ICA-AtoM bracht geen hulp want deze ISAD(G)-mapping bleek dan weer te verschillen van de EAD-import. Na enkele dagen uitgebreid testen en afstemmen hadden we uiteindelijk een goed werkende omzettingmodule. Het resultaat was een programma, dat Atomis gedoopt werd. Je opent een Excel met beschrijvingen en het programma voert een aantal kwaliteitscontroles uit. Zo wordt gecontroleerd of alle verplichte kolommen aanwezig en ingevuld zijn. Het controleert ook of alle kolomnamen conform het sjabloon zijn.

Alle mogelijke voorkomende

datumnotaties worden automatisch omgezet naar een correct formaat dd/mm/jjjj of jjjj. Dit was nodig omdat datumnotaties nogal eens verschilden in de gebruikte tabellen. Niet geldige datumnotaties of onmogelijke data (zoals 31 februari) krijgen een foutmelding.

Bijkomende controles zijn er op geldige invoerwaarden voor bijvoorbeeld het veld 'Trefwoorden'. Dit mag enkel de vastgelegde termen bevatten. Vervolgens wordt de tabel omgezet naar EAD en tot slot wordt het gegenereerde EAD-document in ICA-AtoM geïmporteerd⁸.

HET GEBRUIK VAN HET SJABLOON IN DE PRAKTIJK

Eens de basismodule ontwikkeld was kon het inhoudelijke werk beginnen: het corrigeren van alle vroeger ingevulde tabellen. Dit omzetten, uniformiseren en opschonen van de bestaande beschrijvingen kostte veruit het meeste tijd (bijvoorbeeld het omzetten van voornaam, naam naar naam, voornaam). Gelukkig werd er van bij het begin van de werking van het documentatiecentrum gekozen voor een invoering in tabellen. De gegevens uit Word-inventarissen omzetten naar gestructureerde velden zou immers nog veel meer tijd kosten. Eens de omzetting gebeurd is, is het invoeren nog slechts een kleine moeite.

STRUCTUREREN EN VERPLAATSEN

De vrijwilligers van het documentatiecentrum hadden de gewoonte om archieven uit elkaar te halen en te beschrijven volgens de structuur die

hen het beste past. Zo kwamen affiches, foto's en "grote stukken" uit een archief in verschillende tabellen terecht, omdat men ze anders beschreef. De samenhang van de stukken werd gedocumenteerd in het veld opmerkingen. Door het werken met een vast sjabloon en met de hiërarchische structuur van ISAD(G) was dit niet meer nodig. De archiefstukken kunnen nu beschreven worden in een context die veel nauwer aansluit bij de structuur van de archiefvorming. De nieuwe hoofdstructuur bevat nu volgende nieuwe indelingen: bedrijfsarchieven - verenigingsarchieven - familie- en persoonsarchieven - archieven van publieke organisaties (scholen, parochies,...).

Andere vaststelling was dat grote lijsten van bijvoorbeeld affiches, losse foto's of plannen, moeilijk zo in het systeem konden worden ingevoerd. Technisch was dit snel opgelost, maar te veel beschrijvingen onder elkaar in een hiërarchische structuur presenteren is niet overzichtelijk en komt de raadpleging niet ten goede. Daarom deelden we deze lijsten verder onder, zodat er niet meer dan ca. 150 beschrijvingen onder elkaar terecht kwamen.

Gevolg was wel dat er erg veel verplaatst en verschoven moest worden. Dat is het eenvoudigste in Excel, voor de invoer in ICA-AtoM, maar in de praktijk moesten we heel veel verplaatsen en verschuiven na de invoer. Helaas bleek dit niet eenvoudig in de webinterface van ICA-AtoM.

Daarom werd ook een module toegevoegd aan ons programma waarbij we

“DE EERSTE EN BELANGRIJKSTE LES VOOR HET DOCUMENTATIE-CENTRUM WAS DAT, WELK SYSTEEM JE OOK GEBRUIKT, JE EROP MOET LETTEN DAT DE BESCHRIJVINGEN IN EEN GESTANDAARDISEERD SJABLOON ZITTEN.”

rechtstreeks in de databank beschrijvingen konden verplaatsen.

Eens dit verplaatsen mogelijk was, is er ook voor gezorgd dat we de ingevoerde records beter konden beheeren. Sorteren, aanpassen van typfouten, zoek- en vervangoperaties... kunnen we nu ook met het programma uitvoeren.

AUTHORITY RECORDS

Een van de grote lessen uit de eerste oplaadacties was het belang van het vastleggen van goede authority records en geldige invoerwaarden. Na het verwerken van de inventaris van de affiches stelden we vast dat er heel wat archiefvormers meerdere malen in de databank voorkwamen, door een gebrek aan uniformiteit in de schrijfwijze. Zo bevatte de databank onder meer: BGJG – B.G.J.G. – BGJG Ranst – Bond van Grote en Jonge Gezinnen – Gezinsbond. Daarom werd er bij de verdere ontwikkeling van Atomis voor gekozen om ook archiefvormers (en persoonsnamen) met een sjabloon te beschrijven⁹. Deze beschrijvingen kunnen dan via EAC naar ISAAR(CPF)-fiches omgezet worden en gelden als vaste invoerwaarden, waaraan nieuwe archiefvormers worden getoetst.

Deze kwaliteitscontroles zijn belangrijk omdat dit belangrijke zoekfunctionaliteiten zijn. Een zoekopdracht naar archiefvormer "KSA", moet alle daaraan gekoppelde beschrijvingen opleveren. Er mag geen informatie ontbreken omdat er een alternatieve spelling, zoals "K.S.A." voorkomt.

WAT KON NIET IN ICA-ATOM?

ICA-AtoM is een archiefbeschrijvingssysteem, wat uiteraard maakt dat het niet 100 procent geschikt is voor alle documentatie die moet beschreven worden. Zo kozen we om de boeken en tijdschriften niet in het systeem op te nemen, omdat de openbare bibliotheek van Ranst reeds over een bibliotheekstelsysteem beschikte waarbij we konden aansluiten. Ook voor het beschrijven van artikels uit boeken is het systeem niet zo geschikt. Voor alle andere onderdelen uit onze collectie (met inbegrip van de voorwerpen) bleken de velden van de beschrijvingsstandaarden ISAD(G) of ISAAR(CPF)¹⁰ voldoende flexibel om ze te integreren.

PRAKTISCHE WERKING

Hoewel de tool vrij eenvoudig werkt, is er in het documentatiecentrum voor gekozen om slechts een persoon bestanden te laten nazien, kwaliteitscontroles te laten uitvoeren en de bulkinvoer te laten doen. Andere medewerkers blijven verder beschrijven in een Access-tabel. Voor hen is de enige verandering het gebruik van het nieuwe sjabloon. Alle medewerkers doen wel kleine aanpassingen rechtstreeks in de webinterface van ICA-AtoM. In de praktijk wordt die webinterface enkel nog daarvoor gebruikt. Deze vorm van flexibiliteit was ook een van de redenen om een eigen invoermodule te laten ontwikkelen. Wie eenmalig inventarissen laat omzetten, is nadien vaak verplicht om iedereen met een

bepaald programma te laten werken. Dat kan moeilijk zijn als je werkt met vrijwilligers met een beperkte ICT-kennis. Met een eigen invoermodule die Excel bestanden omzet kan iedereen met standaardofficeprogramma's blijven werken. Daarbij gaat werken in een office-tabel sowieso sneller dan handmatige invoer in een programma.

KERS OP DE TAART: AFBEELDINGEN

Zoals eerder vermeld beschikte het documentatiecentrum ook over 3634 ingescande foto's.

ICA-AtoM liet ook toe om foto's op te nemen, maar 3634 foto's handmatig importeren is te arbeidsintensief, tijdrovend en dus ook te duur. De

Kasteel Bossenstein

Gemeentelijk Documentatiecentrum Ranst
www.documentatiecentrumranst.be

Identificatie	
referentie code	Foto_1-6-NB_151
Titel	Kasteel Bossenstein
Datum(s)	Bij benadering 1945-1960 (Vervaardigd)
Beschrijvingsniveau	Stuk
Omvang en medium	Stuk

Archiefinstelling
Gemeentelijk
Documentatiecentrum Ranst

Archiefvormer(s)
Onbekend

Archief
Foto's
Broeschem
Straten en gebouwen
Abelebaan 12, vroegere diamantslijperij Van Elo...
Abelebaan 2
Abelebaan 4
kerkenpad tijdelijk afgesloten (nie...
Abelebaan 4
kerkenpad tijdelijk afgesloten (nie...
Abelebaan-
Slijkstraat zuidoost voor de serreb...
Abelebaan-

Het kasteel te Bossenstein te Broeschem, rond 1950. Gemeentelijk Documentatiecentrum Ranst, inventarisnummer NB_151.

laatste investering van het documentatiecentrum zorgde voor een module om automatisch foto's op te laden. Hierbij werden thumbnails gemaakt van de foto's en werden de foto's gekoppeld aan hun beschrijvingen, die eerder al zijn opgeladen via de omzetting naar EAD. Hierdoor is er geen aparte beeldbank nodig.

(Zie afbeelding p. 22.)

LESSEN EN CONCLUSIES

De eerste en belangrijkste les voor het documentatiecentrum was dat, welk systeem je ook gebruikt, je erop moet letten dat de beschrijvingen in een gestandaardiseerd sjabloon zitten. ISAD(G) is daarvoor bijzonder geschikt. De meeste soorten documentatie kunnen worden beschreven, en ook complexe archiefbeschrijvingen en context vormen geen probleem. Goede afspraken over de invulling van de velden (bijv. naam, voornaam) zijn nodig. Authority records en kwaliteitscontroles moeten ervoor zorgen dat de zoekfuncties optimaal zijn. Ook hier is het gebruik van een sjabloon primordiaal. Vierde fundament is een goede hiërarchische structuur voor de ISAD(G)-beschrijvingen. Verplaatsen van records die al ingevoerd zijn is in veel systemen problematisch. Indexen moeten herberekend worden, sorteren duurt langer, etc. Eens die fundamenten gelegd, is de implementatie van gelijk welk geïntegreerd systeem veel eenvoudiger.

ICA-AtoM is een geschikt geïntegreerd systeem om archiefbeschrijvingen en authority records ter beschikking te stellen. Individuele beschrijvingen

invoeren kan via de webinterface. Voor het Gemeentelijk Documentatiecentrum van Ranst was dat echter geen optie, gezien het grote aantal bestaande digitale beschrijvingen. Daarbij komt dat het gebruikersgemak van handmatige één per één invoer in ICA-AtoM heel laag is. Het gaat erg traag en het overzicht bewaren is niet zo eenvoudig. Bijgevolg is het handiger om medewerkers (zeker wanneer hun ICT-kennis beperkt is) verder te laten beschrijven in een eenvoudige tabel, en dan nadien die tabellen om te zetten naar EAD en op te laden.

Met een beperkt budget is een kleine instelling zoals het Gemeentelijk Documentatiecentrum erin geslaagd om een programma te laten ontwikkelen dat invoer van grote hoeveelheden beschrijvingen mogelijk maakt. Op een tiental dagen tijd zijn met Atomis ca. 18.000 records opgeladen en ter beschikking gesteld op www.documentatiecentrumranst.be. Dankzij de kwaliteitscontroles en de functies voor het data-beheer (verplaatsen, sorteren, zoeken en vervangen,...) kan het documentatiecentrum daarna de records in ICA-AtoM veel beter beheren en bewerken. Ook het geautomatiseerd opladen van foto's is een enorm voordeel. Het zorgt ervoor dat je afbeeldingen op hun juiste plaats in het archievenoverzicht kan opladen. ICA-AtoM staat ook toe om deze heel overzichtelijk te bekijken. Een aparte beeldbank is dan ook overbodig.

Het Gemeentelijk Documentatiecentrum kan de verdere ontwikkelingen en verbeteringen van Atomis helaas niet

alleen meer dragen. We doen dan ook een oproep aan anderen om mee op de kar te springen. Door de investeringen is de basis gelegd van een tool die een eenvoudige omzetting van onbeperkte hoeveelheden beschrijvingen naar EAD en EAC mogelijk maakt. Daarbij zorgt het programma er ook voor dat ICA-AtoM veel gebruiksvriendelijker wordt en dat een aantal technische problemen opgelost worden.

ICA-AtoM kan samen met Atomis het betaalbare archiefbeheerssysteem vormen waar menig archief reeds jaren op wacht. Intussen hebben de erfgoedcellen en de stadsarchieven van de Kempen beslist om ook met ICA-AtoM en Atomis van start te gaan. ■■

> www.documentatiecentrumranst.be
> Voorbeelden van de invoer en het beheer en de gebruikte sjablonen alsook de prijzen: www.Atomis.be

-
- 1 Voor meer informatie over het gebruik en de functionaliteiten van ICA-AtoM op zich, verwijzen we naar Bart Sas, Inge Roosens, Gratis toegang tot ons geheugen? Opensourcesoftware ICA-AtoM op de testbank, in Meta 2011/3, p. 8-13.
 - 2 Ranst is een gemeente van 18.448 inwoners aan de grens van de Kempen. Ze valt buiten het werkingsgebied van de Kempense erfgoedcellen.
 - 3 De auteur van dit artikel is archivaris van opleiding en werkt ook als vrijwilliger voor het Documentatiecentrum.
 - 4 Eigenlijk was dit een van de eerste beoordelingscriteria voor een beschrijvingssysteem: je moet er op een correcte manier de hiërarchische structuur en context van archief mee kunnen beschrijven.
 - 5 Voor een uitgebreide bespreking van de voor- en nadelen van ICA-AtoM, Zie: Inge Roosens en Bart Sas, Gratis toegang tot ons geheugen? Opensourcesoftware ICA-AtoM op de testbank, in Meta 2011, 3
 - 6 Meer info: www.inzake.be
 - 7 Je kan dit sjabloon downloaden op http://www.inzake.be/Atomis/downloads/Atomis_ISADG_Sjabloon.xls
 - 8 Hoe dit in zijn werk gaat kan je zien op <http://www.inzake.be/Atomis/Functies.aspx>
 - 9 Het sjabloon vindt u hier: http://www.inzake.be/Atomis/downloads/Atomis_ISAARcpf_Sjabloon.xls
 - 10 De 6.609 personalia (persoonsnamen) werden als ISAAR-beschrijvingen ingevoerd, waaraan dan de beschrijvingen van de documenten als ISAD-fiche gekoppeld werden.