

Od oftewel Overheidsdocumentatie is een uitgave van de SDU.

Archiveren volgens de baseline voor gemeenten (4)

De informatie voor en in het zaakdossier naar aard en vorm

Welke termen zijn nodig om zowel tekstdocumenten als databasegegevens in samenhang te benoemen en te archiveren? In deze aflevering introduceert de auteur twee nieuwe termen en bespreekt hij de OAIS-standaard, een model uit een internationale standaard voor archiefsystemen, waarin diezelfde samenhang zit.

TEKST ADRIE SPRUIT

Hoe vind je bij archiveren in een digitale omgeving en in combinatie met zaakgericht werken een balans tussen wet, praktijk en theorie? Dat is het thema van een serie artikelen van een van de auteurs van de Baseline Informatiehuishouding Gemeenten. In deze aflevering: de informatie voor en in het zaakdossier naar aard en vorm.¹

Gestructureerde gegevens uit een database zijn anders van vorm dan documenten met bijvoorbeeld teksten of tekeningen. Dat zagen we in de vorige aflevering van deze reeks over de gemeentelijke baseline. Naar dat verschil kijken we in deze aflevering nader. We doen dat in combinatie met twee

nieuwe termen waarmee we gestructureerde gegevens én gegevens in de vorm van bijvoorbeeld tekstdocumenten in samenhang benoemen.

Gestructureerde gegevens

Hoe zien gestructureerde gegevens in een database eruit?

Een database bestaat uit een of meer tabellen die vergelijkbaar zijn met een spreadsheet. Ook een Databasetabel bevat rijen en kolommen. In het vakgebied Databasebeheer wordt zo'n rij een 'record' genoemd. Omdat zo'n record iets anders is dan hetgeen we bij archivering onder een record verstaan, gebruiken we hier verder de Nederlandse term 'rij'.

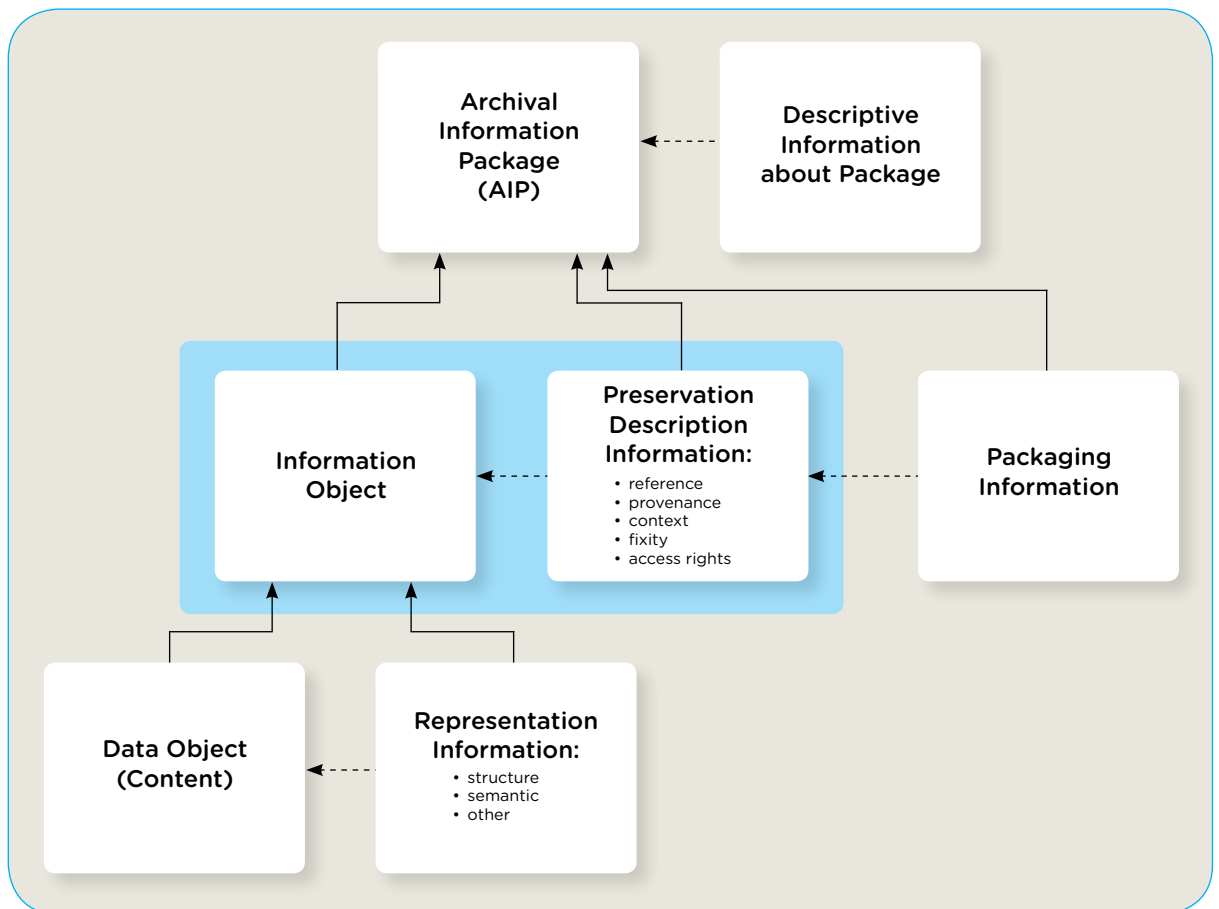
Een rij in een databasetabel bestaat uit een reeks velden. Elk veld hoort bij een kolom en in elk veld staat een gegeven. Dat kan van alles zijn, van een meetwaarde van een snuffelpaal voor luchtverontreiniging tot een huisnummer, een straatnaam, de naam van een persoon of het bedrag van een factuur. Staat er in een veld een huisnummer, dan staat in alle andere velden in dezelfde kolom ook een huisnummer. Maar kijken we horizontaal dan bevat één rij de gegevens over één adres of bijvoorbeeld de gegevens over één persoon. Welk soort gegeven in welke kolom en dus welk veld is vastgelegd – bijvoorbeeld een straatnaam in veld 1, een huisnummer in veld 2 en een plaatsnaam in veld 3 – staat in de structuurinformatie van de database. Die structuurinformatie is nodig om de inhoud van alle velden te kunnen lezen en interpreteren.

Gegevensobjecten en informatieobjecten

In een zaakdossier wil je meestal geen compleet databasebestand opnemen en archiveren, maar soms wel een voor die zaak relevante set gegevens uit zo'n bestand. Dat kunnen de gegevens zijn van enkele velden, van een rij of van meerdere rijen. Stel dat vier rijen gegevens mede de besluitvorming in een zaak bepalen, dan ziet een voor het dossier bestemde selectie er uit als het met rood gemarkeerde deel in figuur 1. Die selectie bevat gestructureerde gegevens en een set met dergelijke gegevens vormt een 'gegevensobject', in het Engels *data object*. Wil je de gegevens van een gegevensobject kunnen lezen, dan is ook de informatie no-

	kolom	kolom	kolom	kolom	kolom	kolom	kolom
rij	veld	veld 2	veld	veld 4	veld	veld	veld
rij	veld	veld 2	veld	veld 4	veld	etc.	
rij							
rij							
rij							
rij							
rij							
rij							

Figuur 1. Databasetabel met gestructureerde gegevens



Figuur 2 OAIS-model voor de structuur van de informatie in een archiefsysteem

dig die de structuur van de database beschrijft. Pas als we de selectie oftewel het gegevensobject combineren met die structuurinformatie elders uit de database en dat als een geheel opslaan in een dossier, dan hebben we iets dat we kunnen archiveren en later ook nog kunnen lezen. De combinatie van een gegevensobject en structuurinformatie is een zelfstandig én leesbaar object. Zo'n object noemen we een 'informatieobject'.

Natuurlijk hoort hier ook de vraag bij hoe je in de praktijk van de uit een database geselecteerde gegevens plus de structuurinformatie uit die database een zelfstandig en leesbaar informatieobject maakt. Het antwoord is dat daar meerdere manieren voor zijn. Een van die manieren is in deze serie al als voorbeeld langsgekomen: je maakt er een XML-document van. Net als een tekstdocument kun je dat als iets zelfstandigs én leesbaars in een dossier stoppen. Ook kun je aan zo'n XML-document metagegevens toevoegen.

Met dit voorbeeld zal ook duidelijk zijn dat digitale tekstdocumenten, zoals een MS Word- of een pdf-bestand, de structuurinformatie om de inhoud te lezen reeds bevatten. Dat zijn dus al informatieobjecten.

Ten slotte nog de vraag hoe gegevensobjecten en informatieobjecten zich verhouden tot informatie-elementen. De term 'informatie-element' staat in de gemeentelijke baseline voor elk onderscheidbaar stukje informatie. Daarmee is het een overkoepelend begrip. De informatievormen 'gegevensobject' en 'informatieobject' zijn categorieën binnen het overkoepelende begrip 'informatie-element'.

De OAIS-standaard

Het OAIS is een standaard die is vastgelegd in de ISO-norm 14721. De titel van deze ISO-norm is *Reference Model for an Open Archival Information System*. De norm bevat een model voor de informatiefuncties van een archiefsysteem én een model voor de structuur of opbouw van gearcheiverde informatie. Figuur 2 toont de hoofdlijnen van dat laatste model. Daarin zien we op het laagste niveau een data object. Tezamen met structuurinformatie – *representation information* in het model – vormt dit een *information object*. Worden aan een information object ook meta-gegevens toegevoegd over onder andere de context of de herkomst van de informatie en bijvoorbeeld de bewaartermijn, dan ontstaat een Archival Information Package (AIP). Hier vertalen we dat naar het Nederlandse 'gearcheiverd informatie-element'. Omdat een AIP of gearcheiverd informatie-element ook de structuurinformatie bevat om de inhoud te lezen heeft het qua opbouw de vorm van een informatieobject.

Gegevensobjecten in een dossier

Hoewel de regel is dat gearcheiverde informatie de vorm heeft van informatieobjecten, kunnen ook gegevensobjecten inhoud van een dossier vormen.

Als de inhoud van een dossier is verdeeld over meerdere systemen, wordt wel gesproken van een virtueel dossier. Bij een zaakdossier is dat in de dynamische fase niet ongebruikelijk, dus in de fase dat de zaak nog in behandeling is en het dossier nog in opbouw. Eerder zagen we uit welk bronssystemen de inhoud van een zaakdossier kan komen.

Die inhoud kun je, zolang de zaak nog in behandeling is, fysiek in die bronsystemen laten zitten. Het centrale digitale dossier bevat dan verwijzingen naar de inhoud in die andere systemen. Bij het voorbeeld dat eerder langskwam, de luchtverontreinigingsgegevens bij de bouwvergunningaanvraag voor een school, blijven de luchtverontreinigingscijfers dan in eerste instantie in de desbetreffende sectorale database zitten. Maar door een verwijzing zijn die gegevens in de vorm van een of meer gegevensobjecten al wel onderdeel van het virtuele dossier. Later, minimaal bij het afsluiten van de zaak, worden die gegevens alsnog fysiek overgehaald naar het centrale dossier en daarin opgeslagen als bijvoorbeeld een XML-document.

De tweede variant van gegevensobjecten in een dossier is minder duidelijk, want afhankelijk van hoe je tegen metagegevens aankijkt, en is ook lastiger uit te leggen. Dat laatste heeft te maken met waarop die uitleg zich richt. Dat is de interne structuur van een archiefsysteem.

Metagegevens op het niveau van informatieobjecten, zoals een pdf of een XML-document, worden na archivering gezien als een onderdeel van het ontstane gearcheverde informatie-element. Informatieobject en metagegevens vormen dan een geheel, zoals in de AIP van het O AIS-model. Maar voor metagegevens op zaakniveau, dus op het niveau van het gehele dossier, ziet dat er meestal anders uit. We kijken daarvoor naar algemene zaakgegevens zoals de naam van de aanvrager, het zaakobject waarover de aanvraag gaat (een boom bij het aanvragen van kapvergunning), de afdeling die de aanvraag heeft behandeld en de behandelend ambtenaar. Dat zijn gestructureerde gegevens.

Als je die eenmalig in een archiefsysteem opslaat, en wel als metagegevens op zaakniveau, dan gebeurt dat over het algemeen in een daarvoor gereserveerde structuur in de database van het archiefsysteem. Dat is een structuur die wordt ingericht bij de implementatie van een archiefsysteem (nadat eerst is vastgesteld welke metagegevens de organisatie op dat niveau wil opslaan). De structuurinformatie om die metagegevens te lezen zit dan dus al in het archiefsysteem zelf. In ieder geval hoeft je dergelijke gestructureerde gegevens niet eerst om te zetten naar bijvoorbeeld een XML-document, alvorens ze – als metagegevens – in het archiefsysteem op te slaan.



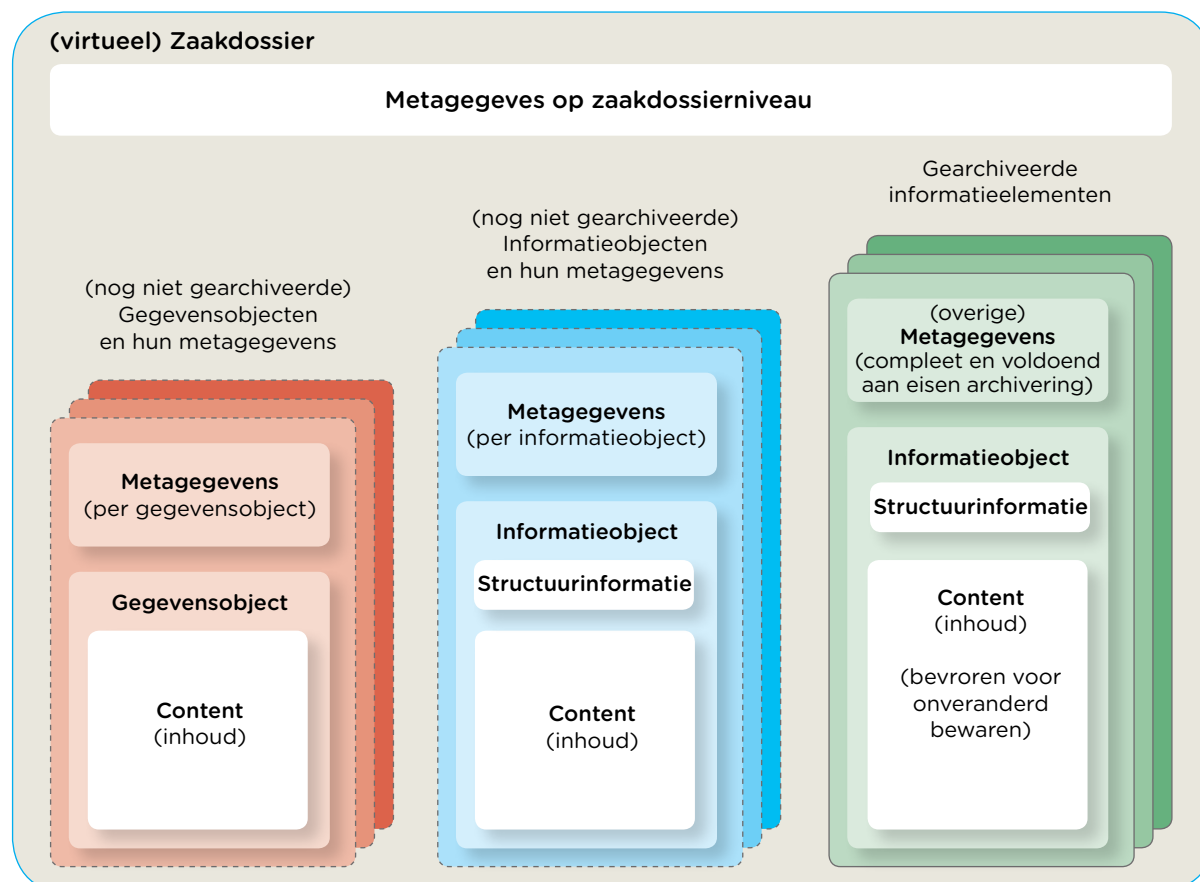
Adrie Spruit
adrie.spruit@kinggemeenten.nl

Adrie Spruit is adviseur gemeentelijke informatiearchitectuur bij KING en een van de auteurs van de Baseline Informatiehuishouding Gemeenten.

De informatie in een dossier op een rij

Als een dossier in de dynamische fase nog in opbouw is, kunnen allerlei soorten informatie (of vormen van informatie) nog naast elkaar in het dossier voorkomen. Een deel van de informatie kan al gearcheveerd zijn, met bevroren inhoud en alle bij archivering horende metagegevens, terwijl andere informatie in de vorm van zowel gegevens- als informatieobjecten, nog niet is gearcheveerd. Zetten we dat op een rij, dan ontstaat de volgende lijst:

- nog niet gearcheveerde gegevensobjecten, zoals de lucht-



Figuur 3. De (mogelijke) inhoud van een zaakdossier in de dynamische fase

verontreinigingscijfers uit het voorbeeld; misschien zijn de gegevens nog in bewerking, maar ze horen al wel bij de zaak;

- nog niet gearchiveerde informatieobjecten; ook die zijn wellicht nog in bewerking, zoals een brief of een conceptbesluit waaraan nog wordt getypt, maar zodra ze bij een zaak horen, horen ze – ook al zijn ze nog niet definitief – al wel in het dossier;
- gearchiveerde informatie-elementen met informatie die niet meer verandert, zoals een pdf van een aanvraag of een besluit; en zelfs van een conceptbesluit als dat een versie is die ter visie is gelegd;
- metagegevens van gegevensobjecten (bijvoorbeeld de creatiedatum van een rij gegevens in een sectorale database), informatieobjecten en gearchiveerde informatie-elementen;
- metagegevens op het niveau van het gehele zaakdossier.

Figuur 3 is een visualisatie van een en ander.

Tot slot

In de volgende aflevering kijken we naar wat er bij het afsluiten van de dynamische fase gebeurt met de in figuur 3 afgebeelde inhoud van een zaakdossier. Met de fasen die daarop volgen ontstaat dan een beeld van de levensloop

van het dossier. Ook zullen we dan zien wat dat betekent voor de levensloop van de afzonderlijke elementen in het dossier en welke stappen of overgangen daar dan voor die elementen bijhoren. ●

¹ In eerdere afleveringen van de serie artikelen over de gemeentelijke baseline was het onderwerp:

- De Archiefwet en de praktijk van alledag
- Wat hoort er in een zaakdossier en wat zijn de bronnen?
- Archiefbescheiden, record en document: drie lastige termen

In volgende afleveringen komen achtereenvolgens aan de orde:

- De levensloop van zaakdossier en inhoud
- Processtappen voor bewaren en archiveren

Zie ook www.od-online.nl