

# Azure Marketplace DataMarket

## GOEDE RESOURCE VOOR HERBRUIKBARE DATA

Alex Arts

Er is al veel geschreven over Windows Azure, maar er zijn ook nog altijd dingen onbelicht gebleven. Een voorbeeld daarvan is Windows Azure Marketplace DataMarket, voorheen codenaam Dallas; een goede resource voor herbruikbare data. Dit artikel gaat dieper in op deze DataMarket en specifiek wat dit ontwikkelaars biedt.

DataMarket is aangekondigd tijdens de PDC eind 2009 en sinds de PDC10 is de DataMarket officieel live. Het is een cloudservice die data aanbiedt voor ontwikkelaars en information workers voor gebruik in applicaties. Je zou dit *Information as a Service* kunnen noemen ten opzichte van Windows Azure dat meer een *Platform as a Service* is. Deze service is een goed voorbeeld van het gebruik van data in de cloud aangezien het draait op Windows Azure en de opslag geregeld is via SQL Azure. Er is een groeiend aantal content provider partners, zoals de Verenigde Naties en de Wereldbank, die gegevens leveren die iedereen kan gebruiken in eigen applicaties of direct in Microsoft Excel. Naarmate meer datasets worden toegevoegd, wordt het interessanter om de DataMarket daadwerkelijk te gebruiken als informatiekanaal.

Omdat de DataMarket is opgezet door middel van open standaarden en protocollen kan een ontwikkelaar ook op andere

platformen applicaties creëren met datasets uit de DataMarket. De data kan overal vandaan komen, de User Interface zorgt voor de beleving en elke toepassing maakt dat uniek.

### Voor wie is het bedoeld?

Information Workers kunnen via Excel 2010 de DataMarket betreden en datasets benaderen. Deze kunnen in een SharePoint Server 2010 omgeving, PowerPivot, PivotChart of in een Excel 2010 rapport worden gepresenteerd in een formaat dat de gebruikers zelf kunnen aangeven. Doordat de data van een plek komen die te bereiken is met een enkele knop, zal het gebruik ervan vereenvoudigen. Er hoeft geen rekening te worden gehouden met verschillende bronnen of formaten, alles is gestandaardiseerd.

Alles wat nodig is om de dataset te tonen in bijvoorbeeld een PowerPivot-tabel is een Azure DataMarket dataset URL naar de te gebruiken dataset en een accountsleutel. Daarna kan de gebruiker zelf de data filteren, sorteren en presenteren. Om de data te vernieuwen, kan er eenvoudig op de verversknop worden gedrukt.

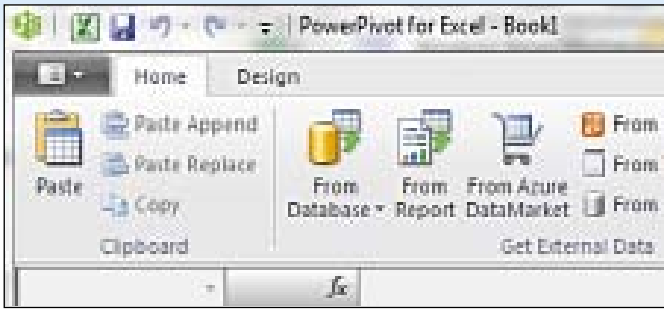
Microsoft heeft geprobeerd om dit alles zo eenvoudig mogelijk te houden voor gebruikers, zodat de leercurve laag is. Op deze manier is het platform voor iedereen te gebruiken.

Voor ontwikkelaars kan een dataset in de DataMarket als basis dienen voor de data in een applicatie. De interface kan er dan omheen gebouwd worden. Met de dataset kan een SQL database of andere opslagmethode worden opgezet om deze op te slaan en te gebruiken. Als de initiële dataset alleen wordt gebruikt om data te tonen is geen opslag nodig, omdat deze vanuit de cloud beschikbaar is.

In de DataMarket is een prijsmodel beschikbaar, vergelijkbaar met Windows Azure; je betaalt voor wat je gebruikt en de eventuele abonnementen die je afneemt. Er is een introductieaanbod waarmee je gratis gebruik kunt maken van de DataMarket service. Zodra je je hebt aangemeld, krijg je een unieke accountsleutel.



FIGUUR 1: DE STARTPAGINA VAN WINDOWS AZURE MARKETPLACE DATAMARKET.



FIGUUR 2: DATAMARKET PLUG-IN VOOR POWERPIVOT IN EXCEL.

Er zijn drie verschillende manieren om datasets aan te schaffen: gratis versies, transactie gebaseerd en abonnementen. Transactie gebaseerd betekent dat betaald wordt per query naar de dataset of een ander antwoord van de DataMarket. Je krijgt een bepaald aantal transacties tot je beschikking en mocht je over je aantal heen gaan, kan er geen gebruik meer van de DataMarket gemaakt worden.

Een oplossing is dan over te gaan op een abonnement op de betreffende dataset. Bij zo'n abonnement wordt een vast bedrag betaald en kan er ongelimiteerd gebruik gemaakt worden van de dataset. Bij een abonnement kan de initiële dataset ook aangesproken worden via de eigen server, zodat niet telkens een verbinding naar de Marketplace moet worden gelegd. Dit betekent dat er vanuit de DataMarket de dataset zelf wordt vernieuwd op je eigen server.

Er is ook een handige manier om een dataset uit te proberen. Bij elke dataset in de DataMarket is een pagina waarop de informatie over de dataset wordt getoond die relevant is voor ontwikkelaars, zoals interactieve selecties op de data, screenshots van hoe de dataset kan worden gebruikt en een complete lijst van parameters die uit de betreffende dataset kan worden gehaald, zodat je weet welke parameters je kunt gebruiken in je applicatie. Zo kan op de pagina van de dataset zelf al worden gekeken of de dataset geschikt is voor het beoogde doel zonder dat het aangeschaft hoeft te worden. Dit geeft ook een realtime voorstelling van hoe de data is opgebouwd of opgebouwd kan worden.

In de DataMarket zie je op elke pagina van een aanbieder welke vormen ze aanbieden, bijvoorbeeld een gratis variant. Als je een transactiegebaseerd of ongelimiteerd abonnement kiest, wordt de eerste maand meteen verrekend.

## Een dataset gebruiken

De DataMarket is een service die alle datasets beschikbaar stelt via RESTful OData feeds met HTTP/HTTPS URIs en die aan te spreken is via REST based API sets. De REST API is OData wat betekent dat de accountsleutel als een wachtwoord mee gestuurd wordt in de authenticatieheader om zo basisbeveiliging te leveren. OData is een standaardprotocol voor het delen van data.

Met OData kunnen onder andere relationele databases, het filesystem of websites worden aangesproken. Er wordt gebruik gemaakt van simpele HTTP-berichten in een uniforme gateway zodat iedere ontwikkelaar op zijn/haar vertrouwde manier kan programmeren. Door middel van strongly-typed data access en IntelliSense support kan iedereen eenvoudig datasets gebruiken. De API geeft altijd een XML-formaat terug door gebruik van een

C# proxy class. Deze proxy class is beschikbaar op de Details-pagina van de aangeschafte dataset en wordt toegevoegd in Visual Studio door Add Service Reference te gebruiken. Deze class gebruikt de WCF Data Service cliënt bibliotheek om een strongly-typed.NET class te creëren die dan de OData service kan aanroepen. Op deze manier is het eenvoudig om de class te gebruiken in een applicatie.

Er zijn drie dingen nodig om van een dataset gebruik te kunnen maken:

- Een Windows Azure Marketplace DataMarket account (je Windows Live ID);
- een account sleutel; te vinden in de Marketplace DataMarket account op de accountsleutelpagina;
- een referentie naar de API van de te gebruiken dataset. Deze referentie wordt gebruikt nadat een Service Reference in het Visual Studio project is opgezet naar de dataset.

Om de API aan te spreken kunnen we de volgende code gebruiken:

```
private const string SERVICE_ROOT_URI = "https://api.datamarket.azure.com/Data.ashx/UnitedNations/FAO/";
```

Natuurlijk moet je wel een abonnement nemen op een dataset, anders kun je de dataset niet gebruiken.

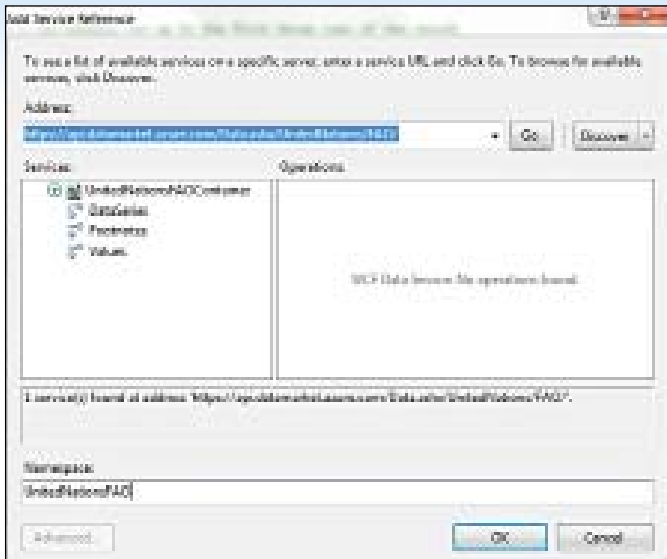
In code wordt dan een ServiceUri opgezet naar de dataset door middel van de DataMarket account en sleutel:

```
public Food()
{
    Uri serviceUri = new Uri(SERVICE_ROOT_URI);
    context = new UnitedNationsFAOContainer(serviceUri);
    context.Credentials = new NetworkCredential(USER_ID, SECURE_ACCOUNT_KEY);
}
```

De data kan met een LINQ-expressie en een eigen keuze van tools worden opgehaald en gepresenteerd. Bijvoorbeeld om een lijst met data uit de United Nations Food Supply te halen, gebruik je een LINQ expressie zoals onderstaand.



FIGUUR 3: VOORBEELD VAN EEN DATASET PAGINA.



FIGUUR 4: TOEVOEGEN VAN EEN SERVICE REFERENTIE NAAR JE DATASET.

```
public List<UNSummary> getFoodData()
{
    IEnumerable<UNSummary> query;

    query = from c in context.Values
            select new UNSummary(c.CountryName,
                                c.ElementName,
                                c.Year,
                                c.Unit,
                                c.YearValue);

    return query.ToList();
}
```

Een class kan gemaakt worden om de data uit de dataset te halen en te formatteren.

```
public class UNSummary
{
    public string CountryName { get; set; }
    public string ElementName { get; set; }
    public Int32 Year { get; set; }
    public string Unit { get; set; }
    public float? YearValue { get; set; }

    public UNSummary(string countryName,
                    string elementName,
                    Int32 year,
                    string unit,
                    float? yearValue)
    {
        this.CountryName = countryName;
        this.ElementName = elementName;
        this.Year = year;
        this.Unit = unit;
        this.YearValue = yearValue;
    }
}
```

In de presentatiepagina wordt dan een referentie naar de class gelegd en kan de lijst die hier uit voorkomt worden gekoppeld aan een GridView door middel van een DataBind().

```
grdFood.DataSource = UNFAO.getFoodData();
grdFood.DataBind();
```

Onderstaande afbeelding is een basisrepresentatie van deze data en kun je natuurlijk zelf formatteren. Door parameters op te geven in combinatie met de Service root URL is alleen die data op

te halen waar je mee gaat werken. Deze wordt gefilterd aan de server kant, zodat je alleen download wat je gebruikt. Elke keer als de pagina wordt aangesproken wordt de data opnieuw opgehaald en blijft dus altijd actueel. Dit is heel handig als je zelf geen database nodig hebt om deze gegevens op te slaan. En doordat sommige datasets gratis zijn, bespaar je ook nog geld.

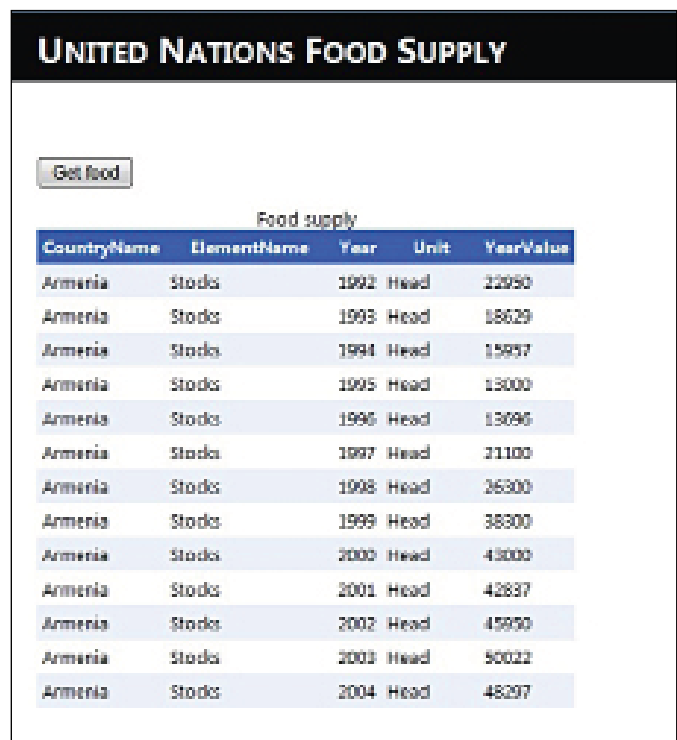
Dit is natuurlijk maar een enkele toepassing van een dataset. Je kunt een dataset ook gebruiken om een eigen applicatie uit te breiden, bijvoorbeeld met de dataset van de regering van de VS die de criminaliteitscijfers laat zien in de VS. Als je een demografische applicatie bouwt en je laat zien hoe veilig het is in de wereld, zou je deze dataset kunnen gebruiken om gefilterde resultaten te geven over de criminaliteit in de VS.

Hetzelfde kun je dan doen voor elk ander groot land of gebied. Als je de cijfers zelf hebt, kun je ze ook beschikbaar stellen in de DataMarket door ze te publiceren. Let wel op dat je dan moet uitzoeken wie de rechten bezit van de betreffende data.

## Een dataset publiceren

Je kunt ook een dataset beschikbaar stellen in de DataMarket. Omdat deze is gebouwd op Windows Azure en gebruik maakt van een SQL Azure database, is dat eenvoudig te doen. Een dataset kan op twee manieren worden opgegeven. De connectie-string en een set van tabellen en kolommen van een SQL Azure database kunnen worden gebruikt om je data eenvoudig over te zetten.

Een andere manier is om de end points en mapping van een bestaande Webservice op te geven om zo de data te migreren naar de SQL Azure database van de DataMarket. De informatie in de dataset wordt uiteraard eerst door Microsoft gecheckt. Deze dataset moet wel van toegevoegde waarde zijn voor andere Information Workers. Zodra de dataset is toegevoegd in de DataMarket, kan iedereen met een DataMarket account er gebruik van maken.



FIGUUR 5: VOORBEELD VANEEN APPLICATIE MET GEBRUIK VAN EEN DATASET.



FIGUUR 6: DE STARTPAGINA VAN WINDOWS AZURE MARKETPLACE APPLICATIONS.

Vooralsnog is de enige manier om een dataset toe te voegen aan de DataMarket een mail te sturen naar het DataMarket team met alle informatie bijgevoegd. Het is dus niet mogelijk het proces als voorbeeld te laten zien.

Microsoft laat je zelf de prijs, hoeveelheid data en voorwaarden opstellen. Maandelijks zijn rapportages beschikbaar en worden automatisch betalingen verricht.

## Windows Azure Marketplace Applications

Naast de Marketplace DataMarket is er ook de Marketplace Applications waar applicaties en services kunnen worden gekocht om

een Windows Azure applicatie te bouwen of uit te breiden. Dit kunnen volledige applicaties zijn die je meteen kunt gebruiken of handige plug-ins om het gebruik van een eigen Windows Azure applicaties te vereenvoudigen zoals een PayPal Windows Azure integratie SDK.

Het principe is dus hetzelfde als de DataMarket, alleen betreft het hier complete applicaties of services en geen datasets. Deze Marketplace Applications is nog in bèta, dus heeft ook nog geen eigen URL.

## Conclusie

De Windows Azure Marketplace DataMarket biedt een schat aan informatie die alleen maar zal groeien en laat zien hoe Windows Azure en SQL Azure kunnen worden gebruikt. Door datasets op deze marktplaats te zetten, kan data eenvoudig gedeeld en hergebruikt worden. Wat zeker ontwikkelaars over de streep zal trekken is het gratis introductieaanbod waarbij je alleen hoeft te betalen voor gebruik. Zo kan iedereen de service uitproberen in de ontwikkeltaal van zijn/haar keuze. Deze lage drempel en het feit dat de data in Microsofts datacenters wereldwijd staat zal zeker voor een succes zorgen.



## Links

- Windows Azure Marketplace DataMarket: <https://datamarket.azure.com/>
- Excel Add-in for DataMarket: <https://datamarket.azure.com/addin>
- Windows Azure Marketplace DataMarket Blog: <http://blogs.msdn.com/b/datamarket/>

.....  
**Alex Arts**, ([alex.arts@avanade.com](mailto:alex.arts@avanade.com)) is een interne training consultant bij Avanade, een samenwerkingsverband tussen Accenture en Microsoft.



## Nieuwe boeken van Microsoft



### [Developing Windows Phone 7 Applications](#)



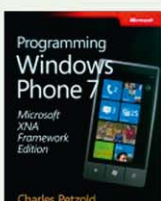
### [HTML5 Step by Step](#)



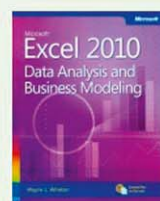
### [Managing and Implementing Microsoft SharePoint 2010 Projects](#)



### [Programming Windows Phone 7: Microsoft Silverlight Edition](#)



### [Programming Windows Phone 7: Microsoft XNA Framework Edition](#)



### [Microsoft Excel 2010: Data Analysis and Business Modeling](#)