

# Hoe semantiek kan bijdragen aan een veiliger Nederland

Is 'kennisbewust opereren' de volgende stap na datagedreven werken en informatie-gestuurd optreden?



\*\*\*\_



## Highlights

- Het gebruik van gemodelleerde kennis in RDF kan leiden tot betere acties en conclusies.
- Het 'graafmodel' van RDF sluit aan op het menselijk denken.
- RDF helpt bij het trainen en uitlegbaar houden van kunstmatige intelligentie.
- RDF kan de kans op fouten in ketens verkleinen en efficiëntie verhogen.
- RDF sluit aan bij een opkomende modelleringstandaard voor publieke organisaties.

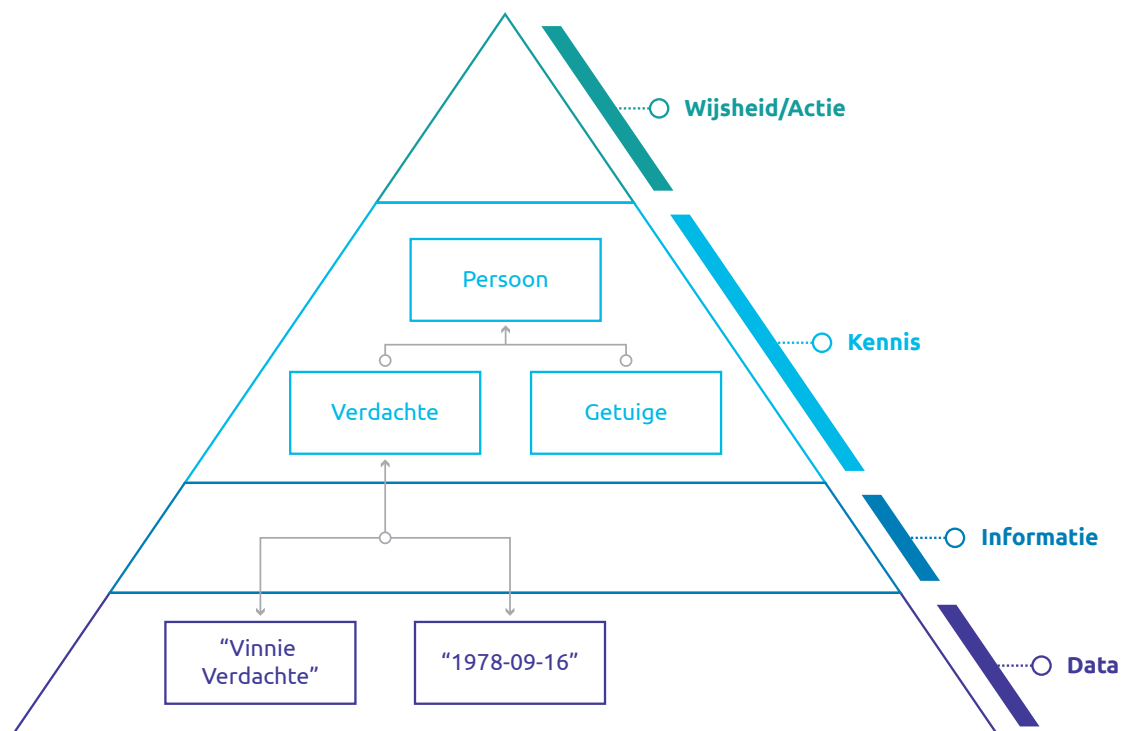


In dit artikel schetsen we hoe veiligheidsprofessionals in staat worden gesteld om effectiever met informatie om te gaan, door het gebruik van semantische modellen op basis van RDF.

## Het modelleren van kennis in RDF

Semantiek is de leer van de betekenis en houdt zich bezig met het zinvol gebruik van symbolen en taal om informatie over te brengen. In de communicatie tussen mensen is het belangrijk dat iedereen hetzelfde bedoelt met hetzelfde woord. Eveneens, als er tussen digitale systemen gegevens moeten worden uitgewisseld is een gedeelde betekenis ook essentieel. In het kader van systeemontwikkeling kan semantisch modelleren worden begrepen als het vastleggen van een werkgebied in een model waarbij elk objecttype een definitie heeft en de mogelijke relaties tussen objecttypen betekenisvol zijn gemodelleerd. Informatiearchitecten en datamodelleurs doen dit effectief al jaren, waarbij het resultaat meestal eindigt als naslagwerk. Echter, RDF (Resource Description Framework) maakt het mogelijk om modellen in dezelfde vorm op te slaan als de gegevens die door het werkende systeem stromen, waardoor data en de betekenis ervan (semantiek) beide bruikbaar zijn voor analyses en zoekvragen aan het systeem. Het potentieel van de modellen wordt zodoende optimaal benut; het systeem kan dan niet alleen zoekvragen en berekeningen uitvoeren op geregistreerde gegevens maar ook op relevante en gerelateerde begrippen. Dit wordt kennisrepresentatie genoemd.

**Figuur 1: De piramide van data, informatie, kennis en wijsheid/actie.**



Een simpel voorbeeld van kennisrepresentatie is het volgende: voor mensen is het helder dat zowel verdachten als getuigen van een incident personen zijn (in een bepaalde rol), maar een systeem heeft deze kennis in principe niet. Wanneer 'verdachte' en 'getuige' dus als specifiekere klassen van 'persoon' worden gemodelleerd, dan zijn alle personen aan de hand van de generieke klasse 'persoon' snel op te vragen, terwijl ze misschien niet als zodanig waren geregistreerd. De piramide uit figuur 1 illustreert conceptueel hoe een systeem dat gebruik maakt van kennisrepresentatie, een gebruiker kan ondersteunen: data zijn ruwe, geregistreerde symbolen, informatie is data in samenhang die voor iemand van nut is en kennis is informatie over informatie. Over de modelleringswijze van dit simpele voorbeeld valt te twisten en het voorbeeld resulteert niet direct in iets dat niet ook anders opgelost had kunnen worden. Echter, het dient om aan te geven dat wanneer bepaalde domeinkennis zinvol is gemodelleerd, dit zou kunnen leiden tot snellere inzichten of gerichtere acties. Zoals de piramide conceptueel laat zien zou deze winst dus mede tot stand komen door het toepassen van gemodelleerde kennis in RDF.

## Overeenkomst met het menselijk denken

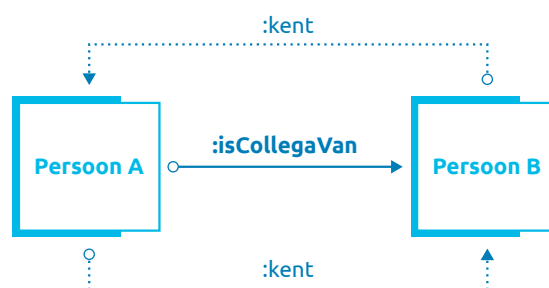
Het toepassen van RDF onderscheidt zich verder van traditionele systemen in het feit dat gegevens niet worden opgeslagen als tabellen, maar als een 'graaf'. Dit is een wiskundig concept dat zich laat visualiseren als een netwerk van bolletjes (of nodes) die zijn verbonden met lijnen die relaties voorstellen. In RDF is behalve aan de nodes, ook aan de relaties daartussen betekenis te geven. Een relatie kan worden hergebruikt en vormt zodoende een belangrijke bouwsteen in een systeem of applicatie dat met betekenisvolle relaties werkt. In het veiligheidsdomein is dit belangrijk, omdat zaken als dreiging vaak gepaard gaan met personen en objecten die op een bepaalde manier met elkaar zijn verbonden. Bovendien sluit deze manier van data verzamelen, verwerken en visualiseren beter aan op het menselijk denken, dan het werken met tabellen. Het voorbeeld van relaties tussen een groep mensen is veel makkelijker te visualiseren in bollen en lijnen (vergelijkbaar met het tekenen op een whiteboard of het maken van een mindmap) dan in één of meerdere tabellen. In grote datasets met complexe gegevens zijn in dit paradigma sneller inzichten te verkrijgen zoals welke nodes een hoge mate van verbondenheid hebben in het netwerk (en dus belangrijk zijn) of dat sommige nodes ongeveer dezelfde relaties of eigenschappen hebben (en daarmee in letterlijke zin een gemeenschap vormen). Dit soort inzichten zijn typerend voor wat ook wel graph-based intelligence wordt genoemd. Wanneer ook semantische modellen in RDF worden toegepast, zoals hieronder wordt geschetst, spreekt men van semantic knowledge graphs.

## Toepassingen in de praktijk van het veiligheidsdomein

Een hogere mate van aansluiting bij menselijk denken is volgens ons al van aanzienlijke waarde, omdat dit tijd kan schelen in bepaalde (denk)processen, maar RDF maakt nog meer zaken mogelijk. Kunstmatige intelligentie (AI) kan semantische modellen in digitale vorm namelijk goed gebruiken om slimmer en efficiënter te worden. Sinds Garri Kasparov door een schaakcomputer werd verslagen is al veel geëxperimenteerd met het automatisch classificeren van data zodat een systeem kan leren van automatisch ontdekte patronen. Echter, dit leidt veelal tot oplossingen waarbij professionals niet meer eenvoudig kunnen uitleggen hoe er tot een bepaalde conclusie of actie wordt gekomen. Wanneer een algoritme de menselijke denkcapaciteit overstijgt, is het dus nuttig – bijvoorbeeld wanneer in de rechtszaal helder bewijs moet kunnen worden overlegd – om elementen uit een door mensen bedacht semantisch model te kunnen aanwijzen waar het algoritme rekening mee heeft gehouden. Voor processen waar dit van belang is kan een semantisch model dus essentieel zijn om AI-toepassingen uitlegbaar te houden.

Daarnaast kunnen semantische modellen een basis zijn voor het vormgeven van gebruikersinteractie met een modern systeem. In een semantisch model in RDF kan bijvoorbeeld worden uitgedrukt hoe de ene relatie zich verhoudt tot een andere. Zo is het bijvoorbeeld mogelijk om te modelleren dat zowel het bevriend zijn met iemand als het zijn van een directe collega, specifiekere vormen zijn van iemand kennen. Ook kan worden uitgedrukt dat iemand kennen een symmetrische relatie is en dat dezelfde relatie dus ook van toepassing is in de tegengestelde richting. Als dus is geregistreerd dat persoon A een collega is van persoon B, zou er achteraf ook vanuit persoon B bekeken kunnen worden wie deze persoon allemaal kent en zal in ieder geval ook persoon A naar boven komen (zie figuur 2). Op basis van dit soort modellen, ook wel ontologieën genoemd, zouden gebruikersschermen kunnen worden ontworpen waarbij het visueel leggen van relaties centraal staat. Het semantische model faciliteert dan vanwege de gemodelleerde kennis niet alleen de gegevensanalyse achteraf, maar ook direct een meer intuïtieve manier van werken voor de gebruiker.

**Figuur 2: Kennen als afgeleide relatie.**

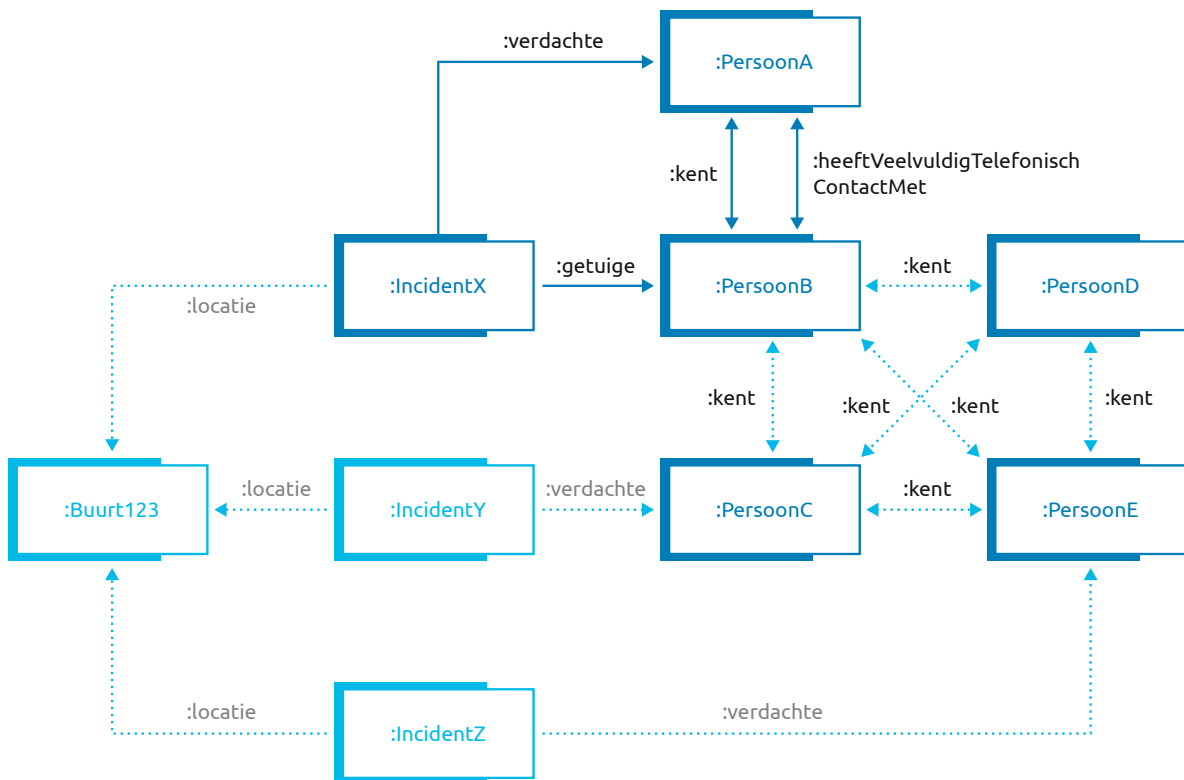




Zoals eerder aangestipt: het inzicht dat iemand een significant aantal andere mensen kent in een netwerk zou bijvoorbeeld kunnen betekenen dat iemand belangrijk is en dat verder onderzoek nuttig kan zijn. Stel dat er een aanleiding is om de afhandeling van een incident als vreemd te bestempelen en dat dit tot nader onderzoek leidt. Belgegevens van de verdachte, persoon A, worden opgevraagd en laten zien dat de verdachte veelvuldig telefonisch contact heeft gehad met persoon B, de getuige van het incident die zich vrijwillig heeft gemeld (zie Figuur 3). Omdat dit vreemd lijkt, wordt persoon B onder de loep genomen en blijkt het dat deze persoon meer mensen

kent die betrokken waren bij verschillende incidenten (Incident X, Y en Z) in de buurt. Persoon A blijkt na onderzoek als enige deze andere personen niet te kennen. Persoon B, hoewel zelf nooit verdachte van een incident en dus niet eerder bekend in het systeem, heeft mogelijk een reden om als getuige op te treden tegen persoon A en wordt op basis hiervan extra aan de tand gevoeld. Uiteraard gaat het hier om een sterk gesimplificeerd voorbeeld, maar laat het zien dat als de incidenten X, Y en Z apart werden bekeken, een stevigere aanpak ten aanzien van persoon B mogelijk achterwege was gebleven.

**Figuur 3: Visualisatie van het voorbeeld.**



## Semantische data-interoperabiliteit in ketens

Naast het feit dat RDF het toepassen van betekenisvolle relaties mogelijk maakt, heeft RDF ook de potentie om organisatie-overstijgend gebruikt en ontwikkeld te worden. Waarin RDF duidelijk verschilt van traditionele systemen is het gebruik van HTTP-URI's als identificatiekenmerk van gegevens in een systeem. Traditionele systemen en databases die niet op basis van RDF werken, gebruiken meestal een voor mensen betekenisloos identificatiekenmerk dat prima werkt binnen het gesloten systeem, maar daarbuiten – zelfs binnen dezelfde organisatie – zegt dit identificatiekenmerk niets. Een HTTP-URI (wat neerkomt op een URL, bijv: 'http://data.mijnorganisatie.nl/id/123456') heeft in de context van het internet een functie en zou dus kunnen worden gebruikt om de betekenis van data en modellen van anderen op te kunnen halen of simpelweg kunnen dienen als referentie. Zo is het bijvoorbeeld mogelijk om uit te drukken dat een gegeven of begrip van organisatie of afdeling A hetzelfde of ongeveer hetzelfde is als die van organisatie of afdeling B. Wanneer dit openlijk wordt gedaan, ontstaat een netwerk van gegevens en modellen waarbij semantiek altijd opzoekbaar is voor zowel mens als machine. Dit wordt ook wel aangeduid als 'Linked Data' en is sinds 2005 de droom van de uitvinder van RDF alsook van het wereldwijde web daarvoor, Tim Berners-Lee. Het publiceren van begrippenkaders in RDF zou verwarring over gegevens in een keten kunnen voorkomen en daarmee de kans op fouten in de keten kunnen verkleinen. Waar dit is toegestaan zouden in data-analyses zelfs gegevens van andere organisaties kunnen worden meegenomen, zonder deze te hoeven kopiëren en kan het beheer ervan worden gelaten bij de bron. Uiteraard moet goed nagedacht worden over zaken als privacy en security, al zouden dergelijke overwegingen eerder in beeld moeten zijn dan de keuze voor RDF.

De vraag rijst al gauw: als RDF al het bovenstaande mogelijk maakt, waarom is dan nog geen sprake van grootschalige adoptie? Uitdagingen die een grote rol spelen zijn het bestaande IT-landschap dat niet zomaar met RDF gemengd kan worden en dat nieuwe mogelijkheden zich het makkelijkst laten integreren met technologieën waar men al veel ervaring mee heeft. Gartner plaatste RDF in 2020 (onder de noemer Ontologies and Graphs) in de Trough of Disillusionment van hun hype-cyclus, en schetst daarmee de situatie dat steeds meer organisaties met realiteitszin de stap aandurven. In dit kader is het Metamodel Informatiemodellering (MIM) van onder andere Geonovum een belangrijke ontwikkeling om te noemen. Dit model beoogt standaardisering van informatiemodellen in het Nederlandse publieke domein en reikt RDF aan als een manier om met het metamodel te modelleren. Het is de verwachting dat Forum

Standaardisatie in 2021 de toetsingsprocedure van MIM hervat voor op de pas-toe-of-leg-uit-lijst, waardoor het MIM voor publieke organisaties een verplicht of aanbevolen karakter krijgt. Het veiligheidsdomein zou er volgens ons goed aan doen om zich voor te sorteren op deze ontwikkeling en deze standaard te volgen. Zeker wanneer dit ook wordt opgepakt door ketenpartijen en wanneer in het veiligheidsdomein steeds meer semantisch wordt gemodelleerd in RDF, zal de veiligheidsprofessional ons inziens ook steeds meer kennisbewust kunnen opereren.

### Auteurs



**Amir Westhoff**  
Senior Consultant

[amir.westhoff@capgemini.com](mailto:amir.westhoff@capgemini.com)  
<https://twitter.com/AmirWesthoff>

Amir is gespecialiseerd in Linked Data/RDF/semantisch modelleren.



**Laura van Dijk**  
Senior Consultant

[laura.van.dijk@capgemini.com](mailto:laura.van.dijk@capgemini.com)

Laura is gespecialiseerd in business- en informatieanalyse binnen het openbare orde- en veiligheidsdomein.