

**Using Augmented Physical Objects driven by Semantic Web
technology to bridge the gap between digital and physical world**

**Het gebruik van verrijkte fysieke objecten gedreven door semantische webtechnologie
om de kloof tussen de digitale en de fysieke wereld te overbruggen**

Ruben Thys

Promotor: Prof. dr. Karin Coninx

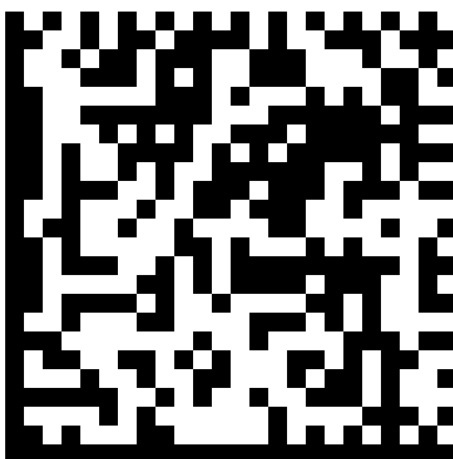
Co-Promotor: Prof. dr. Kris Luyten

Begeleider: Jo Vermeulen

Eindverhandeling voorgedragen tot het behalen van de graad
Master in de informatica afstudeervariant human-computer interaction

School voor Informatie Technologie
transnationale Universiteit Limburg

6 juni 2007



Dit document probeert de kloof tussen digitale en fysieke wereld te overbruggen.

Abstract

De digitale revolutie van het afgelopen decennium is geëxplodeerd tot een digitale wereld die op vele vlakken onafhankelijk is van onze reële wereld. De problemen die hieruit volgen zijn een gevolg van een kloof die bestaat tussen deze digitale wereld en de fysieke wereld. In deze thesis vertrekken we vanuit de studie van technologieën die kunnen bijdragen tot een overbrugging van digitale en fysieke wereld. Vervolgens wordt er een nieuwe architectuur voorgesteld die dienst kan doen als basis om deze problemen op te lossen. De voorgestelde architectuur maakt het mogelijk dat gebruikers *augmented physical objects* kunnen laten interageren met elkaar en maakt hierbij gebruik van *Semantic Web* technieken in combinatie met specifieke identificatietechnologie. Deze architectuur wordt vervolgens als basis gebruikt voor twee toegepaste scenario's (*STalkingObjects* en *SemaNews*) die zich afspelen in verschillende toepassingsdomeinen. Bovenal tracht deze thesis een andere kijk te geven op bepaalde technologieën. Het is precies de combinatie van *Semantic Web* technologie en identificatie-technologie die ons toelaat een brug te vormen tussen fysieke en digitale wereld.

Dankwoord

Iedereen die op welke manier dan ook heeft bijgedragen aan het tot stand komen van deze thesis wil ik hartelijk bedanken.

Eerst en vooral wil ik mijn thesis co-promotor Prof. Dr. Kris Luyten bedanken voor zijn inspiratie, enthousiasme, ideeën, suggesties en opvolging gedurende de volledige thesis. Tijdens de laatste academiejaren heb ik van hem de kans gekregen om onderzoek van korter bij te beleven, waarvoor ik hem ten zeerste dankbaar ben.

Daarnaast zou ik mijn begeleider Jo Vermeulen hartelijk willen danken voor zijn waardevolle feedback, kritiek en ontelbare suggesties, maar ook voor het zeer nauwgezet nalezen en verbeteren van de uiteindelijke tekst. De vele gesprekken met hem werkten zeer inspirerend en hebben geleid tot betere ideeën en implementaties.

Verder dank ik promotor Prof. Dr. Karin Coninx voor het nalezen en bijsturen van deze thesis.

Eveneens dank ik

Mijn vrienden die ik tijdens deze opleiding heb leren kennen voor de vele toffe momenten,

Mijn vriendin Sarah voor haar dagelijkse aanmoediging, motivatie, liefde en discussies die mij hebben doen inzien dat informatica en onderzoek meer is dan technologie alleen,

Mijn broer Kristof die op inhoudelijk vlak steeds bereid was zijn mening te geven en die eveneens een voorbeeld voor mij was gedurende mijn gehele opleiding,

Mijn zus Eline die zo vriendelijk was om haar acteertalent te tonen in het filmpje omtrent *SemaNews*,

Yannick Van de Weerd voor het snel ter beschikking stellen van extra webruimte,

Elias Torres voor zijn hulp en uitleg die nodig was bij het implementeren van een *RDFa parser*.

Tenslotte mogen hier ook mijn ouders niet ontbreken omdat zij mij de kans gegeven hebben om aan deze universiteit te studeren en mij hierin op alle vlakken steeds hebben gesteund.

Voorwoord

Tijdens mijn dagelijks leven word ik voortdurend geconfronteerd met praktische problemen waarbij technologie aan de basis van een oplossing zou kunnen liggen. Het uiteindelijke onderwerp van deze thesis is een voortvloeisel uit mijn interesse naar zulke problemen. Het boek *Being Digital*[29] is doorheen de volledige thesis een enorme drijfveer voor mij geweest. De manier waarop de auteur van dit boek filosofeert rond technologie was een leidraad bij het zoeken naar oplossingen.

Vanuit een selectie van praktische problemen heb ik getracht een algemeen patroon achter deze problemen te herkennen en dit vervolgens te voorzien van een werkbare (technologische) oplossing. Om deze oplossing te realiseren heb ik beroep gedaan op een gedeelte basiskennis die ik reeds had opgedaan tijdens mijn stage en bachelor eindwerk afgelopen jaar.

Het onderzoeksgebied is tijdens deze thesis vaak verschoven wat de thesis voor mezelf zeer boeiend maakte. Hetgeen ik vooral, naast technologische kennis, heb geleerd uit deze thesis is dat een onderzoeksidee nooit op bestelling komt. Het zoeken naar (en in de overgrote meerderheid van de gevallen niet vinden van) goede ideeën vond ik moeilijk en vroeg veel geduld en inspanning.

Inhoudsopgave

I	Introductie	1
1	Inleiding	2
1.1	Inleiding	2
1.2	Thesis onderwerp	3
1.3	Overzicht van deze thesis	5
II	Onderzoek	7
2	Augmented Physical Objects	8
2.1	Inleiding	8
2.2	Definities	8
2.3	Onderzoeksgebied	9
2.4	Concepten	10
2.4.1	Medium	10
2.4.2	Digitale tegenhanger	11
2.5	Literatuur	13
2.5.1	Overbrugging van fysieke en virtuele wereld	13
2.5.2	Implementatie-problemen	14
2.5.3	‘Augmented Physical Object’-systems	15
2.5.4	Internet of Things	15
3	Identificatietechnologie	17
3.1	Inleiding	17
3.2	Radio Frequency Identification	18
3.2.1	Algemeen	18
3.2.2	Componenten	18
3.2.3	Soorten tags	19
3.2.4	Barcode vs. RFID	22
3.2.5	RFID toepassingen	24
3.3	Semacode Tags	24
3.3.1	Algemeen	24
3.3.2	Identificatie	25
3.3.3	Gerelateerde technologieën	25

4	Het Semantisch Web	27
4.1	Inleiding	27
4.2	Syntax vs. Semantiek	28
4.2.1	Syntax	28
4.2.2	Semantiek	28
4.3	Wat is het Semantisch Web ?	28
4.4	Traditionele web	30
4.5	Semantic Web architectuur	31
4.5.1	XML - XML Schemas	32
4.5.2	RDF - RDF Schemas	32
4.5.3	OWL	36
4.5.4	De relatie tussen RDF, RDFS en OWL	38
4.5.5	RDF/A	38
4.6	Kritieken	39
4.7	Het Semantisch Web in deze thesis	39
5	Web Services	40
5.1	Inleiding	40
5.2	Algemeen	40
5.3	Principe van een web service	42
5.4	Web services lagenmodel	42
5.4.1	Network	43
5.4.2	XML-based Messaging - SOAP	43
5.4.3	Service Description - WSDL	43
5.5	Semantische web services	44
5.5.1	Algemeen	44
5.5.2	OWL-S	45
5.5.3	WSDL-S	45
5.5.4	SAWSDL	46
5.5.5	WSDL 2.0	47
5.5.6	WSMO	47
5.6	Literatuur	47
5.7	Toekomst	48
III	Ontwikkeling	49
6	Architectuur	50
6.1	Inleiding	50
6.2	Doelstelling	51
6.3	Architecturaal overzicht	51
6.3.1	Het creëren van een Semantisch Web	51
6.3.2	Het efficiënt gebruik van Semantische informatie	57
6.3.3	Veralgemeend concept	61
6.4	Mogelijke toepassingen	62
6.5	Aannames	62
6.6	Gerelateerd werk	65

7 Applicaties	67
7.1 Inleiding	67
7.2 ‘STalkingObjects’-scenario	68
7.2.1 Aanpak	68
7.2.2 Ontologie beschrijving	69
7.2.3 Beheer van semantische informatie	70
7.2.4 Domeingerelateerde toepassingen	71
7.2.5 3D virtuele shopping-omgeving	72
7.2.6 Communicatie server applicatie	75
7.2.7 Veronderstellingen	76
7.2.8 Bestaande shoppingconcepten	76
7.3 ‘SemaNews’-scenario	79
7.3.1 Algemeen idee	79
7.3.2 Creatie van fysieke krant op basis van digitale informatie	80
7.3.3 Gebruik van digitale informatie in de fysieke wereld	82
7.3.4 Registratie en selectie van web services	84
7.3.5 Toegepast voorbeeldscenario	85
7.3.6 Vereisten	87
IV Conclusie	88
8 Conclusie	89
V Appendices	91
A Afbeeldingen	92
B Afkortingen	96
C Extra’s	98
C.1 Radio Frequency Identification	98
C.1.1 Grondprincipes	98
C.1.2 Standaarden	99
C.2 Fysiek-Digitaal problemen in winkelomgevingen	100
C.2.1 Scenario : consumentenproblemen	101
C.2.2 Scenario : bedrijfsgerichte problemen	102
Bibliografie	104

Lijst van figuren

2.1	Weiser's visie van de afzetmarkt van <i>Ubiquitous Computing</i> in de toekomst.	9
2.2	Augmented Physical Objects vereisen het bestaan van een medium dat instaat voor de interactie tussen gebruikers.	10
2.3	De huidige situatie van voorwerpen en hun informatie	12
2.4	Ideaalbeeld van een digitale tegenhanger	12
3.1	Voorbeeld van een RFID-tag	18
3.2	De 3 basiscomponenten in de RFID technologie	19
3.3	Voorbeeld van verschillende soorten passieve RFID-tags	20
3.4	RFID-tags kunnen reeds ingeplant worden bij mensen	20
3.5	Voorbeeld van een actieve tag	22
3.6	Overzicht eigenschappen van verschillende RFID frequenties	23
3.7	Voorbeeld van een Semacode tag	25
4.1	Traditionele web vs. Semantisch Web	30
4.2	Semantisch Web architectuur	31
4.3	XML	33
4.4	XML Schemas	33
4.5	XML Schemas zijn te weinig flexibel voor de opbouw van een Semantisch Web	33
4.6	RDF Triple voorbeeld	34
4.7	De verschillende RDF notaties	35
4.8	Voorbeeld van een RDF Schema	35
4.9	Voorbeeld van een OWL	37
4.10	RDF - RDF Schemas - Ontologieën	37
4.11	RDF - RDFS - OWL	38
4.12	Voorbeeld RDFa-code	39
5.1	Algemene concept van een webservice	42
5.2	Webservices lagenmodel	43
5.3	Voorbeeld van een OWL-S composite service	46
6.1	Conceptuele architectuur - het creëren van een Semantisch Web	53
6.2	Invoeren van RDF-informatie in OpenWYSIWYG met FOAF2RDFa extensie	54
6.3	Resultaat na FOAF-parsing in OpenWYSIWYG met FOAF2RDFa extensie	55
6.4	Overzicht architectuur : OpenWYSIWYG met FOAF2RDFa extensie	55
6.5	Screenshot van de RDF/A monkey website	56
6.6	Overzicht Augmented Physical Objects concept-architectuur	58
6.7	Vereenvoudigde voorstelling van Augmented Physical Objects	58

6.8	Gedetailleerde implementatie van Augmented Physical Objects concept-architectuur . . .	60
6.9	Overzicht van de volledige concept-architectuur	63
6.10	Alternatief overzicht van de volledige concept-architectuur	64
6.11	Voorbeeld van een mogelijke toepassing van de concept-architectuur	64
6.12	Een universeel identificatiemiddel is reeds realiteit.	65
7.1	Concreet voorbeeld van een SparQL query die gebruik maakt van de shopping-ontologie .	71
7.2	Structuur van de <i>STalkingObjects</i> -implementatie ten opzichte van een uitwerking in de echte wereld	72
7.3	Screenshots 3D virtuele shoppingomgeving	72
7.4	Shoppingbuddy output : 'winkelwagen overzicht' afkomstig van semantische web service .	74
7.5	Shoppingbuddy output : 2D plattegrond van de winkel	74
7.6	Shoppingbuddy output: mashups en productinformatie	75
7.7	Fujitsu Wireless U-Scan Shopper	77
7.8	IBM customer Stop & Shop	78
7.9	Metro's Personal Shopping Assistant	79
7.10	SemaNews : creatie van fysieke krant op basis van RSS en Atom feeds.	81
7.11	SemaNews : gebruik van digitale informatie in de fysieke wereld.	83
7.12	De belangrijkste onderdelen van de SemaNews website	85
7.13	Overzicht van generatie-proces van gepersonaliseerde SemaNews-krant	86
7.14	Het vastleggen van een Semacode tag door middel van de SemaNews mobiele telefoonap- plicatie	86
7.15	Visualisatie van de mogelijk uit te voeren web services	86
A.1	Graafmodel een semantische beschrijving van een product	92
A.2	Graafmodel van een beschrijving van semantisch gekoppelde producten	93
A.3	Graafmodel van een semantisch beschreven shopping	94
A.4	Graafmodel van semantisch gekoppelde shoppings	95
C.1	Eenvoudig overzicht op de werking van passieve RFID-tags	98
C.2	Enkele alomtegenwoordige consument-gerelateerde problemen in de supermarkt.	101
C.3	Bedrijfsgerelateerde problemen in de supermarkt.	103

Deel I

Introductie

Hoofdstuk 1

Inleiding

Inhoudsopgave

1.1	Inleiding	2
1.2	Thesis onderwerp	3
1.3	Overzicht van deze thesis	5

1.1 Inleiding

‘*Computing is not about computers any more. It is about living.*’, zo schreef Nicholas Negroponte¹ reeds meer dan 10 jaar geleden in zijn befaamde boek ‘*Being Digital*’[29].

In 10 jaar tijd is de digitale wereld geëxplodeerd mede dankzij de opkomst van het internet in al zijn facetten. De hoeveelheid informatie die het internet ons biedt is onnoemelijk groot en deze enorme digitale kennisbron vervult vandaag de dag reeds een belangrijke rol in ons dagelijkse leven. Deze digitale revolutie is echter zo snel uitgebroken dat er een kloof is ontstaan tussen deze nieuwe digitale wereld en onze reële wereld. In deze thesis wordt verwezen naar deze kloof onder de benaming *fysiek-digitaal kloof*. Digitale informatie blijft namelijk doorgaans in zijn vertrouwde (digitale) wereld van computers circuleren. Zo is het vandaag de dag bijvoorbeeld nog steeds onmogelijk om aan de koelkast te informeren welke zuivelproducten over datum zijn of in de hifi-winkel aan een flatscreen TV de vraag te stellen hoe andere consumenten over hem dachten. Dit alles blijft toekomstmuziek, hoewel in veel gevallen deze informatie digitaal wel bestaande is.

Desalniettemin probeert men deze *fysiek-digitaal kloof* stap per stap te dichten. Een voorbeeld hiervan is de komst van de GPS-systemen, welke digitale informatie in een vertrouwde reële omgeving (de auto en het verkeer) bruikbaar maken. Een ander, minder voor de hand liggend, voorbeeld zijn *Wireless music systems*, deze maken het mogelijk om internetradio’s te kunnen spelen op eender welke plaats in een vertrouwde huisomgeving.

De basis van het probleem bestaat erin dat online verworven kennis en informatie moeilijk overdraagbaar is naar fysieke locaties waar deze informatie relevant is. Vaak zullen mensen informatie (gebruikersreacties, productbesprekingen, . . .) zoeken op het internet alvorens bijvoorbeeld een nieuw televisietoestel

¹Voorzitter van de ‘One Laptop per Child’-associatie, pionier op vlak van *computer-aided design* en één van de oprichters en voormalig directeur van het MIT Media Lab.

in de winkel te kopen. Deze informatie wordt meestal gezocht via een computer die aanwezig is op de gebruikers' thuislocatie en wordt vervolgens doorgaans afgedrukt of uit het hoofd geleerd. Het kopen van een artikel bestaat gewoonlijk uit drie stappen nl. een eerste bezoek aan de winkel die informatie geeft over de beschikbaarheid van artikelen, vervolgens het zoeken van informatie via het internet op de thuislocatie en de uiteindelijke aankoop van het artikel in de winkel. Het is duidelijk dat dit proces vrij omslachtig is en we naar betere oplossingen moeten zoeken.

'*e-commerce*²' verhelpt dit probleem gedeeltelijk: de fysieke verplaatsing kan bij het kopen op internet achterwege gelaten worden. Toch ontbreekt bij deze oplossing de interactie met de reële wereld, veel mensen willen namelijk een artikel met eigen ogen zien alvorens het aan te kopen. Bij het kopen van een televisietoestel willen mensen bijvoorbeeld de beeldkwaliteit eigenhandig controleren zonder blindelings te vertrouwen op gebruikersreacties of recenties.

De hierboven beschreven problemen die een rechtstreeks gevolg zijn tussen de afstand tussen werkelijkheid en digitale wereld, spelen zich allemaal af in de reële wereld. Echter brengt deze kloof ook problemen met zich mee die zich bevinden in de digitale wereld. Veel informatie die afkomstig is uit de reële wereld moet namelijk handmatig worden overgebracht door gebruikers. Een voorbeeld hiervan zijn de verschillende profielsites die persoonlijke gegevens (interesses, hobbies, verwijzingen naar vrienden,...) vereisen die telkens handmatig door de gebruiker moeten ingegeven worden. Een ander voorbeeld zijn de '*Photo Sharing sites*³', deze sites verlangen dat gebruikers via computer foto's uploaden en nadien metadata⁴ toevoegen. Ook op dit vlak vinden er dus veel overbodige haltes plaats om informatie uit de echte wereld over te brengen naar de digitale wereld.

Voor deze categorie problemen zijn er reeds initiatieven en projecten die een oplossing proberen te bieden. *Friend of a Friend* (FOAF⁵) maakt het bijvoorbeeld mogelijk voor gebruikers om persoonlijke gegevens in een uniform formaat te beschrijven dat leesbaar is voor computers. Voordeel van het gebruik van FOAF, is het feit dat de informatie uit een FOAF-profiel eenvoudig kan hergebruikt worden. *PhotoFOAF*[46] is een voorbeeld van een applicatie die aan de hand van FOAF technologie reeds een overbrugging tracht te maken tussen fysieke en digitale wereld. *PhotoFOAF* speelt zich af in de wereld van de fotografie.

1.2 Thesis onderwerp

Deze thesis tracht de manier waarop we tegen deze *fysiek-digitaal kloof*⁶ aankijken anders te benaderen. Startend vanuit een stevige basis die het *Semantisch Web* kan leveren, onderzoeken we of het *Semantisch Web* in combinatie met enkele andere moderne technologieën (uit andere onderzoeksgebieden) oplossing kan bieden voor een selectie eerder beschreven problemen. We beschrijven in deze thesis een conceptuele architectuur die de *fysiek-digitaal kloof* tracht te overbruggen op een generieke manier. *Semantische Web technologie* wordt als absolute basis van deze architectuur gebruikt. Vertrekkende vanuit de beschrijving van deze architectuur worden er twee verschillende applicaties uitgewerkt waarin we de overbrugging van de *fysiek-digitaal kloof* realiseren.

Er wordt in deze thesis veel belang gehecht aan generaliseerbaarheid, het gebruik van standaarden en de uitbreidbaarheid van de verschillende concepten. De concepten die in deze thesis worden voorgesteld

²Electronic commerce, de verzamelnaam van alle manieren waarop via digitale wegen handel bedreven kan worden.

³Flickr is een voorbeeld van een bekende *Photo Sharing*-site.

⁴Data die andere data beschrijft

⁵<http://www.foaf-project.org>

⁶Niet te verwarren met de *digitale (ICT-)kloof* waarbij mensen zonder financiële middelen niet kunnen mee profiteren van de realisaties op het vlak van digitale revolutie

zijn vaak niet afhankelijk van de technologie die hierbij gebruikt wordt. Hiermee wordt bedoeld dat bijvoorbeeld toekomstige identificatie-technologie binnen de architectuur in deze thesis bruikbaar blijven.

Een eerste applicatie die in deze thesis wordt geïmplementeerd, bevindt zich in het domein van winkelomgevingen. Vermits de *fysiek-digitaal kloof* vandaag de dag in dit toepassingsdomein groot is, trachten we hiervoor oplossingen te bieden. *STalkingObjects* (de naam van dit deelproject) probeert creatieve oplossingen te leveren voor domein-gerelateerde problemen door onder andere gebruik te maken van RFID-technologie.

Volgende paragraaf beschrijft een voorbeeldscenario van de *STalkingObjects*-applicatie die volledig gebaseerd is op de architectuur uit deze thesis :

Jo doet boodschappen in een gloednieuwe ‘next-gen’ shopping waarin alle producten geen barcodes meer bevatten maar zogenaamde ‘RFID’ tags. Op zijn winkelwagen bevindt er zich een zogenaamde ‘shoppingbuddy’, een geavanceerde mini-computer die hem informatie kan verschaffen tijdens het winkelen.

Deze ‘shoppingbuddy’ geeft Jo meer inzicht op de producten die in zijn winkelwagen liggen, zo kan Jo op een eenvoudige manier aan budgetbeheer doen. Daarnaast geeft deze ‘shoppingbuddy’ gepersonaliseerde informatie van producten uit een bepaald rek. Wanneer Jo geïnteresseerd is promoties of nieuwigheden, kan hij eenvoudig een overzicht opvragen van de promoties of nieuwe producten van een bepaald winkelrek. De ‘shoppingbuddy’ kan eveneens automatisch actie ondernemen afhankelijk van de positie waar Jo zich op een bepaald moment in de winkel bevindt. Jo kan ondanks de nieuwe technologie winkelen zoals hij dit altijd al deed, zonder producten te moeten scannen of speciale aandacht te besteden aan zijn ‘shoppingbuddy’.

Eveneens kan Jo zichzelf eenvoudig oriënteren doorheen de winkel, zijn ‘shoppingbuddy’ kan eenvoudig zijn huidige plaats in de winkel tonen op een plattegrond van de winkel.

De architectuur en technologie waarmee deze ‘next-gen’ shopping is opgebouwd, maakt het mogelijk om eenvoudig extra features te voorzien zoals automatische kassa’s, geavanceerd stockbeheer, . . . Daarnaast maakt de architectuur achter deze ‘next-gen’ shopping het zelfs mogelijk om features van verschillende ‘next-gen’ shoppings vrij eenvoudig te koppelen met elkaar. Hierdoor kan Jo nagaan of een product dat niet meer in voorraad is, in een andere shopping (van dezelfde shoppingketen) fysiek aanwezig is.

Een tweede (meer toegepast) scenario heeft de titel *SemaNews* en tracht de *fysiek-digitaal kloof* binnen het domein van kranten en digitale informatie te overbruggen. Voor deze applicatie wordt er gebruik gemaakt van twee-dimensionale barcodes (*Semacode tags*) en *Semantische Web services*. Ook dit scenario baseert zich volledig op de conceptuele architectuur die in deze thesis wordt voorgesteld.

Volgende paragraaf beschrijft een voorbeeldscenario van de *SemaNews*-applicatie en schetst hiermee de praktische mogelijkheden die de conceptuele architectuur uit deze thesis kan bieden :

Eline staat ’s morgens op en kan tijdens haar ontbijt rustig een gepersonaliseerde krant lezen. Haar computer heeft deze krant via de ‘SemaNews’-applicatie automatisch samengesteld op basis van haar favoriete (nieuws)websites en/of haar emails. Tevens wordt deze krant door haar computer elke dag op een tijdstip naar keuze afgedrukt. Ondanks het feit dat de informatie afkomstig is van digitale bronnen, hoeft Eline zich zo vroeg in de morgen niet bezig te houden met technologie. Ze kan dus een aangename papieren krantenversie lezen tijdens haar

ontbijt.

Hoewel deze krant er ‘normaal’ uitziet, bevat de informatie in de krant geavanceerde barcodes waardoor het originele (digitale) karakter van de informatie behouden blijft. Eline hoeft zich niet druk te maken over de manier waarop dit juist gebeurt, voor haar lijkt alles kinderspel. Zo kan ze namelijk via een ‘SemaNews’-applicatie op haar mobiele telefoon informatie uit de krant gebruiken om haar persoonlijke blog te updaten. Eveneens kan ze een bepaalde locatie die aangehaald wordt in een artikel, zeer snel en eenvoudig op een digitale kaart bekijken. De verschillende mogelijke toepassingen die ze kan gebruiken met de informatie uit de krant, kan Eline beheren op de ‘SemaNews’-website. Daarenboven zal de ‘SemaNews’-applicatie op haar mobiele telefoon steeds een selecte groep van toepassingen tonen, die nuttig zijn voor de informatie die Eline uit de krant wil gebruiken. Wanneer Eline dus informatie wil gebruiken over een artikel waar er geen locatie wordt aangehaald, zal de mobiele applicatie hiermee rekening houden.

De toepassingen (gerealiseerd door ‘web services’) die de ‘SemaNews’-website aanbiedt kunnen vrij eenvoudig worden uitgebreid. Deze ‘web services’ zijn ‘slim’ in de zin dat ze weten welke informatie interessant voor hunzelf is en welke niet.

Deze twee scenario’s (*STalkingObjects* en *SemaNews*) trachten naast de implementeerbaarheid van de architectuur die in deze thesis wordt voorgesteld, ook te bewijzen dat deze architectuur gebruikt kan worden met verschillende gelijkaardige technologieën. Hierdoor is de architectuur flexibel en kunnen gerelateerde toekomstige technologieën nog steeds gebruikt worden. Tevens proberen deze scenario’s de voordelen en mogelijkheden naar de toekomst van *Semantische Web* technologie aan te tonen.

De materie die in deze thesis wordt bestudeerd en toegepast is verspreid over enkele onderzoeksgebieden. De architectuur in deze thesis steunt op twee grote peilers nl. :

1. De studie van ‘*augmented objects*’ (ook wel ‘*smart objects*’). Dit zijn objecten die een meerwaarde bieden ten opzichte van ‘gewone objecten’. In deze thesis zijn dit fysieke objecten die zijn uitgerust met *Radio Frequency Identification* (RFID tags) of Semacodes. Deze identificatietechnologieën worden uitgebreid besproken in hoofdstuk 3. Vaak verwijst men in de literatuur hiernaar onder de noemer *Internet Of Things*.
2. *Semantic Web*-technologie. Deze nog vrij jonge technologie geeft een betekenis aan objecten⁷ in de digitale wereld, waardoor deze objecten interpreteerbaar worden voor computers. Wat *Semantic Web*-technologie juist inhoudt bekijken we in hoofdstuk 4. In hoofdstuk 5 gaan we dieper in op web services die uitgebreid worden met *Semantic Web* technologie.

Het is onmogelijk om binnen het bestek van deze thesis alle raakvlakken van het probleem domein in detail te bestuderen. Er is getracht de meest relevante en interessante ideeën en toepassingen van naderbij te bekijken. Belangrijk om op te merken is het feit dat de uiteindelijke scenario’s meer dienst doen als een soort leidraad dan dat ze het uiteindelijke doel op zich zijn.

1.3 Overzicht van deze thesis

Deel II van de thesis bestaat uit een uitgebreide studie naar verschillende technologieën en technieken die later in de conceptuele architectuur van deze thesis worden gebruikt. Meer bepaald gaat hoofdstuk

⁷documenten, foto’s, films, websites,...

2 dieper in op de studie rond ‘*Augmented Physical Objects*’, terwijl hoofdstuk 3 identificatietechnologie (‘*Radio Frequency Identification*’ en ‘*Semacode tags*’) van dichterbij bekijkt. Hoofdstuk 4 geeft een beter beeld op het begrip *Semantisch Web* en beschrijft enkele gerelateerde technologieën. Hoofdstuk 5 gaat dieper in op web service technologie en bekijkt het idee en de implementatie van zogenaamde *semantische web services* in meer detail.

Deel III bestaat uit een beschrijving van de ontwikkeling die in tijdens deze thesis is gebeurd. Hoofdstuk 6 kan gezien worden als het kloppend hart van deze thesis en stelt de conceptuele architectuur voor waarmee de overbrugging tussen digitaal en fysieke wereld kan vergemakkelijkt worden. Hoofdstuk 7 beschrijft de implementatie van twee uitgebreide scenario’s, nl. ‘*STalkingObjects*’ en ‘*SemaNews*’ die gebruik maken van deze architectuur.

Deel IV geeft een slotconclusie en een overzicht op het mogelijk toekomstig werk.

Deel II

Onderzoek

Hoofdstuk 2

Augmented Physical Objects

Inhoudsopgave

2.1	Inleiding	8
2.2	Definities	8
2.3	Onderzoeksgebied	9
2.4	Concepten	10
2.4.1	Medium	10
2.4.2	Digitale tegenhanger	11
2.5	Literatuur	13
2.5.1	Overbrugging van fysieke en virtuele wereld	13
2.5.2	Implementatie-problemen	14
2.5.3	‘Augmented Physical Object’-systems	15
2.5.4	Internet of Things	15

2.1 Inleiding

Om de eerder beschreven scenario’s (in sectie 1.2) te realiseren is het van belang te zoeken naar een algemeen patroon dat in zulke scenario’s steeds wederkeert. De studie omtrent *Augmented Physical Objects* is reeds een onderzoeksdomein dat dit patroon definieert en telkens tracht te verbeteren. Dit hoofdstuk geeft een inleiding tot begrippen binnen het onderzoeksgebied van *Augmented Physical Objects*. Tevens wordt er een overzicht gegeven van de reeds bestaande literatuur hieromtrent. Het inzicht dat wordt verworven binnen dit hoofdstuk, draagt bij tot een beter begrip van de uiteindelijke ontwikkeling (in deel III) binnen deze thesis.

2.2 Definities

De term ‘*Augmented Physical Objects*’ (vrij vertaald ‘*verrijkte fysieke voorwerpen*’) staat voor fysieke voorwerpen in onze reële wereld die *meer* mogelijkheden bieden dan normale objecten. Met meer mogelijkheden wordt verwezen naar het feit dat deze objecten in staat zijn om menselijke gebruikers hulp te bieden bij alledaagse activiteiten door gebruik te maken van digitale technologie in al zijn facetten.

In deze thesis zal de term ‘*Augmented Physical Objects*’ (vaak verkort tot ‘*augmented objects*’ of ‘*smart objects*’) fysieke objecten voorstellen die worden gebruikt om de digitale en fysieke wereld korter bij elkaar

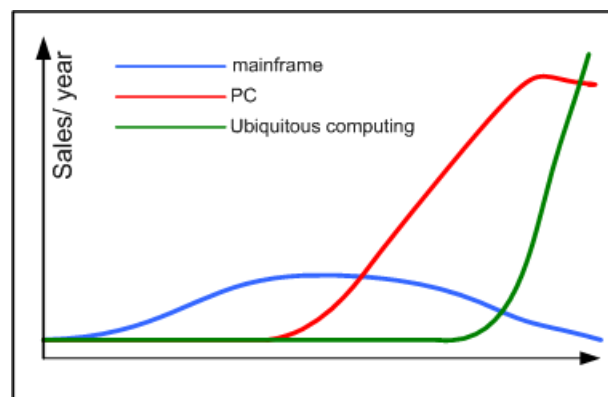
te brengen. Hierbij maken deze *augmented objects* gebruik van zowel informatie afkomstig uit de digitale wereld als uit de fysieke wereld. Menselijke gebruikers kunnen interactie hebben met deze *augmented physical objects*, al dan niet zonder dit zelf te beseffen. Deze voorwerpen proberen de digitale wereld op een zo doorzichtig mogelijke manier, een plaats te geven in onze reële wereld. Doch proberen deze objecten geen onnodige informatie op te dringen maar staan ze eerder ter hulp.

Doordat deze *augmented objects* vaak technologie bevatten die geïntegreerd wordt in de voorwerpen zelf, wordt de benaming ‘*embedded devices*’ soms ook gebruikt. Toch is er een verschil tussen *embedded devices* en *augmented physical objects*. Een voorbeeld van een *embedded device* is een GPS-systeem dat geïntegreerd is in een auto. Een *augmented physical object* is bijvoorbeeld een museum-artefact dat meer informatie over zichzelf verschaft wanneer een bezoeker zich kortbij het museumstuk begeeft.

2.3 Onderzoeksgebied

Het onderzoeksgebied omtrent *augmented physical objects* is deel van een groter geheel met de naam ‘*Ubiquitous Computing*’. Om inzicht te hebben in het onderliggend doel van deze thesis, is het van belang om te begrijpen waarvoor ‘*Ubiquitous Computing*’ staat en waar het naar streeft.

‘*Ubiquitous Computing*’ is het onderzoeksgebied dat vaak als de derde grote golf in de wereld van computers wordt gezien¹. Na de mainframe computers (veel personen, 1 computer) en personal computers (1 persoon, 1 computer) wordt verwacht dat het onderzoeksgebied van *Ubiquitous Computing* de volgende jaren een enorme groei zal kennen. De term *Ubiquitous Computing* werd als eerste gebruikt in 1988 door de helaas al overleden Mark Weiser², *Ex-Chief Technology* van het al even bekende laboratorium Xerox PARC. Weiser en John Brown³ waren tevens een van de eersten die onderzoek deden naar praktische oplossingen in het gebied van *Ubiquitous Computing*[50].



Figuur 2.1: Weiser's visie van de afzetmarkt van *Ubiquitous Computing* in de toekomst.

Het onderzoeksgebied van ‘*Ubiquitous Computing*’ spitst zich toe op technologie die een plaats vindt in de achtergrond van ons leven. Hierdoor wordt het vaak omschreven als ‘*calm technology*’, technologie die er naar streeft om de gebruiker niet te overladen met informatie. ‘*Ubiquitous Computing*’ probeert dit

¹Figuur naar <http://sandbox.xerox.com/ubicomp/>

²http://whatis.techtarget.com/definition/0,,sid9_gci211737,00.html

³<http://www.johnseelybrown.com/>

te bereiken door de gebruiker op een of andere manier zelf te laten aangeven wat centraal als informatie staat en wat zich in de rand bevindt.

Weiser[49] beschrijft het uiteindelijke doel van ‘*Ubiquitous Computing*’ kort en bondig als volgt⁴ :

‘...our highest ideal is to make a computer so embedded, so fitting, so natural, that we use it without even thinking about it’

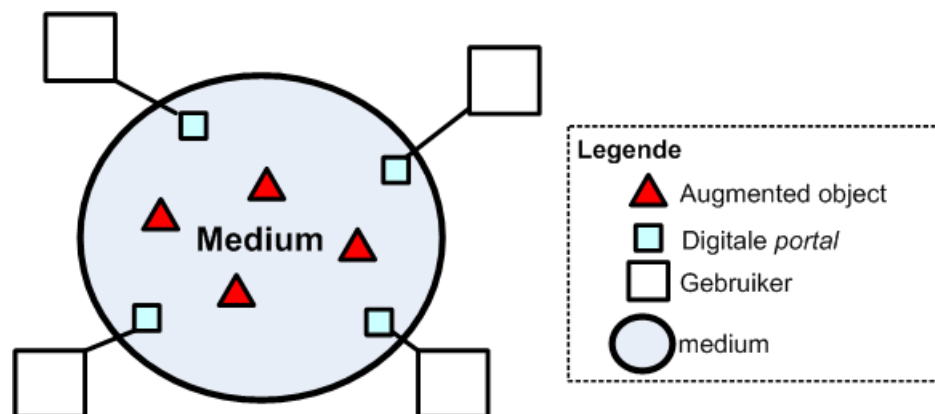
Het gebied van *Ubiquitous Computing* bestaat uit veel sub-onderzoeksgebieden, één daarvan draait rond ‘*Augmented Physical Objects*’. Hoewel er verscheidene jaren aan onderzoek voorbij zijn gevlogen sinds de opmars van ‘*Ubiquitous Computing*’, bevat elk kleiner onderzoeksgebied nog steeds de oorspronkelijke onderliggende gedachtes die *Weiser* reeds meer dan tien jaar geleden beschreef.

2.4 Concepten

De verschillende architecturen die ‘*Augmented Physical Objects*’ beschrijven, bevatten steeds enkele overeenkomstige concepten. De implementatie van deze concepten kan verschillen, doch is de achterliggende visie steeds gelijk.

2.4.1 Medium

Het hoofddoel van *Augmented Physical Objects* is de menselijke gebruiker ten dienste staan in allerlei activiteiten. Hiervoor is een interactie vereist tussen de gebruiker en het *augmented object*, dewelke enkel kan plaatsvinden binnen een interactie-ruimte. Vaak wordt de algemene benaming ‘medium’ aan een interactie-ruimte gegeven⁵. Figuur 2.2⁶ geeft een voorbeeld van een medium en zijn componenten.



Figuur 2.2: Augmented Physical Objects vereisen het bestaan van een medium dat instaat voor de interactie tussen gebruikers.

Gebruikers (meestal menselijke gebruikers) kunnen via een digitale *portal* interactie hebben binnen een bepaald medium met *augmented objects*. De digitale *portal* kan bijvoorbeeld een mobiele telefoon of draagbare computer zijn. Een medium op zich bestaat steeds uit drie vaste basiscomponenten[30, 9] :

⁴<http://www.informit.com/articles/article.asp?p=165227&rl=1>

⁵http://www.iscid.org/encyclopedia/Interaction_Space

⁶naar <http://www.acm.org/ubiquity/views/>

1. Organisatie component (sociale component)
2. Logisch medium
3. Fysiek medium

Organisatie component

Deze component definieert welke rollen (functies) de verschillende gebruikers hebben binnen een medium. Tevens worden in deze component ook de regels vastgelegd die interactie beschrijven (ook wel het protocol).

Logisch medium

Om in een medium interactie te hebben met elkaar, is het van groot belang dat men een taal spreekt die door alle componenten in het medium verstaanbaar is. Hiervoor bestaat er een logisch medium (oftewel logische ruimte) dat een formele taal beschrijft in al zijn facetten (soort symbolen, syntax, ...). Doch bevat deze logische ruimte geen semantiek, vermits die kan beschreven worden door hogere ontwikkelingen.

Fysiek medium

Het gedeelte dat instaat voor de uiteindelijke fysieke overdracht van gegevens tijdens interacties, wordt het fysiek medium (of 'carrier medium') genoemd. Op vlak van het fysieke medium bestaan er meerdere technologieën die de laatste jaren hun opgang hebben gemaakt. In deze thesis wordt er gebruik gemaakt van drie technologieën binnen het fysiek medium, nl. *RFID*, *WiFi* en *Semacode tags* technologie. Hoofdstuk 3 beschrijft identificatietechnologie in detail, vermits een deel van de implementatie in deze thesis steunt op deze technologie.

2.4.2 Digitale tegenhanger

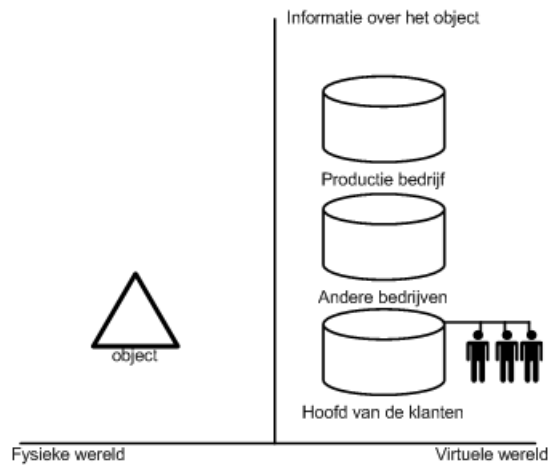
Wat *augmented physical objects* onderscheidt van normale alledaagse voorwerpen, is het feit dat ze een digitale tegenhanger (in de literatuur[32] naar verwezen als 'digital counterpart' of gewoon 'digital object') hebben.

Wanneer mensen fysieke objecten kopen, vergewissen ze zich er meestal niet van dat er veel interessante informatie bestaat omtrent deze voorwerpen. Doch bevindt deze informatie zich meestal niet bij de voorwerpen zelf maar in een digitale wereld. Deze virtuele wereld is op dit moment behoorlijk gescheiden van de fysieke wereld. Figuur 2.3⁷ geeft de huidige situatie aan omtrent fysieke objecten en de informatie over deze objecten.

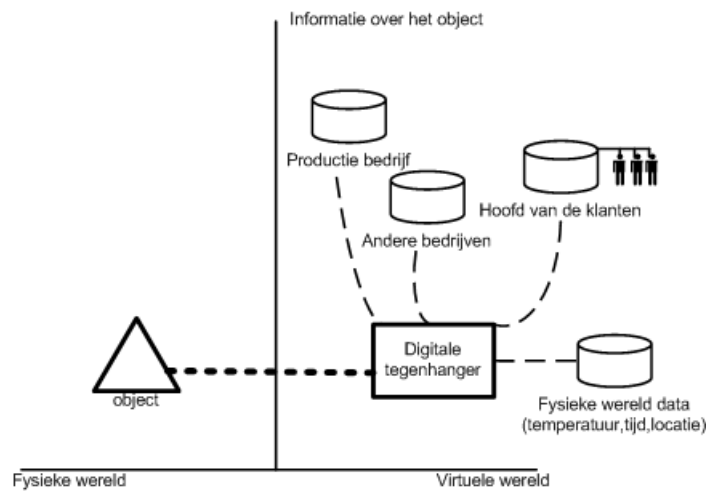
De informatie van verschillende bronnen is niet rechtstreeks gekoppeld met het fysieke voorwerp zelf. Deze verschillende bronnen zijn bijvoorbeeld het bedrijf dat een bepaald voorwerp ontwikkelt, andere bedrijven die soortgelijke voorwerpen ontwikkelen maar ook de informatie in het *geheugen* van klanten die ervaring hebben met het voorwerp.[30]

De situatie waar men in het gebied van *augmented objects* naar streeft is de realisatie van een digitale tweelingsbroer van elk fysiek object die gekoppeld is met het oorspronkelijke fysieke object. De bedoeling is dat deze tegenhanger op de een of andere manier informatie bevat die gelijkend is aan de rechterhelft uit figuur 2.3. Figuur 2.4 geeft een schematisch overzicht van het ideaalbeeld van een digitale tegenhanger. In sectie 2.5 bespreken we het gerelateerd werk dat er reeds in dit gebied gebeurd is.

⁷naar <http://www.alexandria.unisg.ch/Publications/by-projects/3180/16586>



Figuur 2.3: De huidige situatie van voorwerpen en hun informatie



Figuur 2.4: Ideaalbeeld van een digitale tegenhanger

Een digitale tegenhanger kan verschillende soorten informatie bevatten :

- Productie informatie (bv. gewicht en prijs van het voorwerp)
- Functionele informatie (bv. hoe het voorwerp gebruiken?)
- Contextuele informatie (bv. temperatuur, locatie,... van voorwerp)
- Statistische informatie (bv. aantal mensen die tevreden waren over een bepaald voorwerp)
- ...

Een digitale tegenhanger die al deze informatie bevat is en blijft vandaag de dag een ideaalbeeld, en hoogstwaarschijnlijk zal in komende jaren hier geen verandering in optreden. Toch is het van belang dat er gestreeft wordt naar een ideale digitale tegenhanger, om de kloof tussen fysiek en digitaal eenvoudiger te kunnen dichten. Toch is het een feit dat zulke tegenhangers ook veel problemen met zich mee zouden brengen, waaronder bijvoorbeeld *privacy*-problemen. In deze thesis laten we *privacy issues* links liggen, vermits dit ons te ver zou leiden.

2.5 Literatuur

In de volgende sectie bekijken we literatuur die reeds stappen heeft ondernomen in de goede richting binnen de wereld van *Augmented Physical Objects*. We bekijken literatuur die zich uitspreidt over meerdere probleemgebieden, gaande van het literatuur die zoekt naar oplossingen voor specifiekere implementatieproblemen tot uitgewerkte prototypes.

2.5.1 Overbrugging van fysieke en virtuele wereld

In 1991 stelde Pierre Wellner in een video⁸ als één van de eersten een duidelijke (en zeer indrukwekkende) visie voor die de overbrugging tussen fysieke en virtuele wereld mogelijk maakte. In zijn *Digital Desk*-scenario[52, 51, 53] slaagt erin vrijwel in om op een naadloze manier objecten in de fysieke en digitale wereld met elkaar te koppelen. Sinds deze paper is er een enorme interesse in fysieke user interfaces gegroeid, gaande van het definiëren van paradigma's[11] tot effectieve prototypes die gebruik maken van fysieke user interfaces[16, 18]. De overgrote meerderheid van dit soort onderzoek maakt gebruik van fysieke objecten die geassocieerd zijn met elektronische objecten (meestal door deze elektronische objecten in de fysieke objecten te verwerken). Typisch registreert een computer in de nabijheid de manipulaties die gebeuren op deze objecten en voorziet deze computer output via een display.

[48] werkt verder op de visie van *Wellner* en is in het bereik van deze thesis zeer interessant. Deze paper levert enkele nieuwe fysieke prototypes gebruik makend van vier hoofdtechnologieën nl. RFID-technologie (lezers en tags), RF networking, infrarood *beacons* en *portable computing*. Deze studie maakt gebruik van alledaagse voorwerpen waar RFID-tags aan worden toegevoegd.

In tegenstelling tot andere initiatieven[18, 52] (in onder andere de wereld van de *tangible interfaces*) houdt deze studie wel rekening met het kostenplaatje en reële haalbaarheid op grotere schaal. In één van hun voorgestelde prototypes is het mogelijk om de informatie van een bepaald boek te tonen wanneer deze wordt gescand via een RF-lezer. Toch is de koppeling tussen de fysieke objecten en de virtuele tegenhanger zeer primitief. Tevens wordt er geen gebruik gemaakt van de opslagruimte en herschrijfbaarheid van RFID-tags. Deze studie behoort tot één van eersten die de fysieke en digitale wereld korter bij mekaar probeerden te brengen door middel van RFID-technologie (besproken in hoofdstuk 3). [24] bekijkt de

⁸http://www.idemployee.id.tue.nl/g.w.m.rauterberg/presentations/Build-It_Story/tsld002.htm

consequenties die het gebruik van RFID op grotere schaal met zich meebrengt en trekt hier vervolgens lessen uit.

2.5.2 Implementatie-problemen

Een aanzienlijk probleem dat zich voordoet wanneer *augmented objects* in grote hoeveelheid worden gebruikt, is de zogenaamde ‘*information overload*’. In het artikel ‘*When RFID Is Everywhere, How Will You Cope*’⁹ haalt Nicholas Hoover dit probleem aan en tracht hij oplossingen te vinden. Via het wiskundig begrip ‘*topology*’ probeert men oplossingen voor dit vraagstuk te vinden¹⁰.

Ook *Mark Weiser* merkt het *information overload* probleem op en beschrijft het als volgt :

...Most important, ubiquitous computers will help overcome the problem of information overload. There is more information available at our fingertips during a walk in the woods than in any computer system, yet people find a walk among trees relaxing and computers frustrating. Machines that fit the human environment, instead of forcing humans to enter theirs, will make using a computer as refreshing as taking a walk in the woods.[27]

[13] beschrijft een systeem waarbij boeken zichzelf op basis van een user profile kunnen ‘aanraden’ aan de gebruiker. Zo wordt de ‘*information overload*’-kwestie zo goed als mogelijk omzeild doordat andere oninteressante objecten worden weggefilterd. In het voorgestelde prototype, getiteld ‘*Galatea*’, worden boeken voorzien van een *Bluetooth*¹¹-ontvanger en is de gebruiker in bezit van een mobiele telefoon. Zoals zelf wordt opgemerkt in deze studie is *Bluetooth* geen ideale identificatie-methode vermits de kostprijs vrij hoog ligt. Toch levert deze studie enkele interessante ideeën op. Zo bieden *augmented physical objects* in ‘*Galatea*’ enkel info via persoonlijke adaptatie.

Een interessante aanpak tot dit probleem wordt de laatste jaren ondernomen door zogenaamde *Web 2.0 mashups*. Volgens wikipedia¹² is een web mashup een website of applicatie die inhoud combineert van meer dan één bron in een geïntegreerde omgeving. Een bekend voorbeeld is het tonen van foto’s op de positie van een kaart overeenkomstig met de locatie waar de foto’s genomen zijn. Hierbij combineert men bijvoorbeeld de web service van Flickr met die van Google maps. [47] tracht web mashups te creëren op basis van *tangible* voorwerpen die samengebracht worden in de reële wereld.

[3] is een studie die in reikwijdte van deze thesis interessante startpunten biedt. In deze paper merkt men op dat passieve RFID-tags verspreid over een beperkt gebied, krachtige mogelijkheden kunnen bieden. Tevens merkt men op dat tags een opslagruimte hebben die in velen gevallen onbenut blijft. In deze paper wordt er gewerkt rond zogenaamde ‘*super-distributed RFID tag infrastructures*’ (SDRI), slaande op het feit dat er gewerkt wordt met een groot aantal tags die zijn uitgespreid over een beperkt gebied. Uitgaande van deze SDRI’s stelt men enkele zeer interessante ideeën voor zoals het verspreiden van geografische informatie door mobiele devices. Deze mobiele devices verspreiden GPS-coördinaten uit over een SDRI. Vervolgens kunnen mobiele devices zonder GPS-ontvanger hiervan gebruik maken, en tevens hebben devices met GPS-ontvanger een alternatief om op terug te vallen wanneer GPS zou falen.

In verschillende papers verwijst men naar problemen die zich voordoen bij het ontwikkelen van smart-object en (meer veralgemeend) ‘Ubiquitous Computing’ prototypes[25, 12]. Deze papers merken op dat

⁹<http://www.informationweek.com/news/showArticle.jhtml?articleID=193302577>

¹⁰<http://www.engr.uiuc.edu/news/?xId=068909120756>

¹¹Bluetooth is een open standaard voor draadloze verbindingen tussen apparaten op korte afstand.

¹²http://en.wikipedia.org/wiki/Mashup_%28web_application_hybrid%29

het opzetten van zulke prototypes in de fysieke wereld meestal (zeer) duur is. [25] raadt dan ook aan om devices en bestaande architecturen te hergebruiken alvorens een eigen ontwikkeling aan te vatten. Verder wordt er gewezen op het feit dat tegenwoordig PDA's en mobiele devices betaalbaar zijn geworden en uitstekende devices zijn voor 'Ubiquitous Computing'-onderzoek.

2.5.3 'Augmented Physical Object'-systems

[23] stelt enkele *smart devices* voor die kunnen interageren met andere devices. In tegenstelling tot eerdere aanpakswijzes ([21, 22]) probeert men de gebruikte *embedded devices* zoveel mogelijk interactie met elkaar te laten hebben, dit om de zwaktes van het ene *device* te verhelpen door de sterkes van andere *devices*. Het prototype dat wordt voorgesteld is weinig overtuigend, en blijft het vrij toegespitst op één bepaald scenario. Toch is het idee dat deze paper bevat waardevol, het bevat Weiser's grondgedachte achter *Ubiquitous Computing* volledig.

In het onderzoeksgebied van *Ubiquitous Computing* is er de laatste jaren enorm veel onderzoek gedaan, eveneens zijn er enorm veel plannen van systemen die onze fysieke omgeving *slim* willen maken. Van deze geplande systemen zijn er zeer weinig effectief uitgewerkt en van de uitgewerkte systemen doet maar een gering aantal interessante dingen. Een veelgemaakte fout in het domein van *Ubiquitous Computing* is de vaststelling dat veel technologie enorm interessant is maar er te weinig aandacht wordt besteed aan de uiteindelijke bruikbaarheid en uitwerking in applicaties. Rogers[36] maakt een overzicht van de fouten die binnen het domein van *Ubiquitous Computing* vaak worden gemaakt.

'*Smart-Its*'¹³ is een zeer bekend project in de wereld van de toekomstige *Ubiquitous Computing* applicaties dat interessant werk levert[38, 20]. Via Bluetooth en RFID-technologie in combinatie met *Wireless Networks* tracht men een werkbare architectuur voor *Ubiquitous Computing* te creëren. De applicaties die worden ontwikkeld met *embedded smart devices* (zogenaamde 'Smart-Its') doen in dit project eerder dienst als beschrijving van wat mogelijk *zou* zijn. Het project is meer toegespitst op de hardware technologie waardoor er nog niet zo veel software toepassingen zijn ontwikkeld op het *Smart-Its* platform. *Smart-Its Friends* is een voorbeeld van een bestaande toepassing die mensen (met *Smart-Its*) in contact met elkaar brengt[19]. Deze applicatie kan eveneens dienst doen als monitoring-service voor ouders en hun kinderen. Wanneer een *Smart-It* van een kind verder dan een bepaald bereik is, zal dit gemeld worden via een geluidssignaal. Merk op dat men ook hier de standaard user interface (zoals we die kennen van de desktop computer) naar de achtergrond verwijst en alle interactie zoveel mogelijk via *smart objects* laat plaatsvinden.

2.5.4 Internet of Things

De architectuur die later in deze thesis wordt voorgesteld toont de meeste overlappings met het onderzoeksgebied dat onder de noemer *Internet of Things* bestaat. Dit onderzoeksgebied tracht door middel van *object hyperlinking*, het huidige internet uit te breiden naar objecten en locaties in onze reële wereld, om zo te komen tot een internet bestaande uit 'dingen'. *Object hyperlinking* is het proces waarbij fysieke objecten worden voorzien van een URL door middel van tags. Deze tags kunnen voorzien zijn door verschillende technologieën, waaronder *RFID*, *Graphical tags* (Semacodes, QR Codes, ...), SMS-tags, ...

[1] beschrijft een architectuur waarin fysieke objecten worden voorzien van barcodes. Eveneens wordt er een *service broker* voorzien die functionaliteit kan bieden op basis van een barcode. Toch spitst deze paper zich vooral toe op de interpretatie van barcodes op mobiele devices, en wordt er vaag gesproken over het uiteindelijke gebruik van services. Tevens is de manier waarop deze services worden aangeboden

¹³<http://www.smart-its.org/>

en gebruikt vrij primitief. . . Neil Gershenfeld ea. beschrijven manieren waarop alledaagse devices de mogelijkheid krijgen om te connecteren tot een data netwerk. Om dit praktisch werkbaar te maken wordt er gekeken naar het data protocol van het huidige internet. Zo wordt er uiteindelijk gewerkt naar het idee van een netwerken van alle soorten types van devices, beter gekend als *interdevice networks*[14]. Na de beschrijving van de architectuur (in hoofdstuk 6.3.3), volgt een verdere studie van gerelateerde literatuur hieromtrent.

Deze thesis tracht het *internet of things* verder uit te breiden door middel van *Semantische Web* technologie (semantische beschrijvingen en web services), en hecht hierbij veel belang aan een werkbaar scenario dat de voordelen van deze aanpak praktisch kan aantonen.

Hoofdstuk 3

Identificatietechnologie

Inhoudsopgave

3.1	Inleiding	17
3.2	Radio Frequency Identification	18
3.2.1	Algemeen	18
3.2.2	Componenten	18
3.2.3	Soorten tags	19
3.2.4	Barcode vs. RFID	22
3.2.5	RFID toepassingen	24
3.3	Semacode Tags	24
3.3.1	Algemeen	24
3.3.2	Identificatie	25
3.3.3	Gerelateerde technologieën	25

3.1 Inleiding

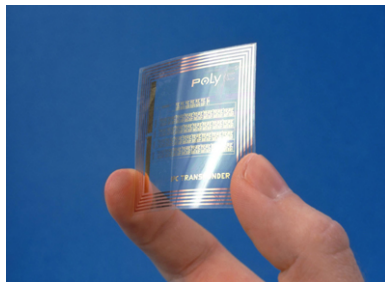
Volgend op hoofdstuk 2 over *Augmented Physical Objects* is er binnen dit concept nood aan specifieke technologie die informatie kan overdragen van fysieke naar digitale wereld. De technologie die hiertoe in staat is, beschrijven we in dit hoofdstuk onder de noemer *identificatietechnologie*.

Wanneer we *Augmented Physical Objects* willen ontwikkelen is het belangrijk kennis te hebben over de mogelijkheden van de identificatietechnologie die we gebruiken om deze objecten te ontwikkelen. In deze thesis wordt er gebruik gemaakt van twee verschillende identificatietechnologieën, die beiden voor- en nadelen hebben. Vermits *Radio Frequency Identification* technologie de meest belovende identificatietechnologie is, bekijken we deze technologie uitvoerig in sectie 3.2. Later in sectie 3.3 bekijken we *Semacode tags* (die eveneens in de uiteindelijke implementatie worden gebruikt) wat meer in detail.

3.2 Radio Frequency Identification

3.2.1 Algemeen

Radio Frequency Identification (RFID) is een technologie die toestaat om informatie op een afstand te lezen van zogenaamde *RFID-tags*¹. Deze RFID-tags zijn er in alle vormen en maten (gaande van flinterdunne RFID-tags tot tags ter grootte van een baksteen) en kunnen op fysieke objecten zitten. Hierdoor wordt het mogelijk om fysieke objecten te identificeren via een zogenaamde RF-lezer, een apparaat dat informatie van een RFID-tag kan inlezen. De RFID-technologie werkt op basis van radiogolven en kan eenvoudig uitgelegd worden als een moderne variant van de *barcode*. In sectie 3.2.4 wordt er een vergelijking gemaakt tussen de oorspronkelijke barcode en RFID-tags. Figuur 3.1² toont een voorbeeld van een flinterdunne RFID-tag die men aan drukwerk kan toevoegen.



Figuur 3.1: Voorbeeld van een RFID-tag

Er is enige onduidelijkheid rond het precieze ontstaan van RFID. Enkele bronnen melden dat RFID technologie accidenteel is uitgevonden in 1946 door *Léon Theremin* als spionage instrument. In 1948 publiceerde *Harry Stockman* ‘*Communication by Means of Reflected Power*’ en concludeerde dat er nog een heleboel werk was alvorens bruikbare applicaties te hebben. In de jaren '60 explodeert het onderzoek rond RFID technologie, wat tot gevolg had dat er een heleboel implementaties ontstonden die werkten volgens het principe van RFID. Vandaag de dag bestaan er veel toepassingen die gebruik maken van RFID. Algemeen wordt aangenomen dat RFID in verloop van tijd de barcodes zullen verdringen als standaard identificatiemiddel. Sectie 3.2.5 geeft een overzicht van enkele toepassingen die gebruik maken van RFID in de huidige wereld.

3.2.2 Componenten

De RFID technologie bestaat uit drie basis componenten. Figuur 3.2 geeft een overzicht van deze componenten.

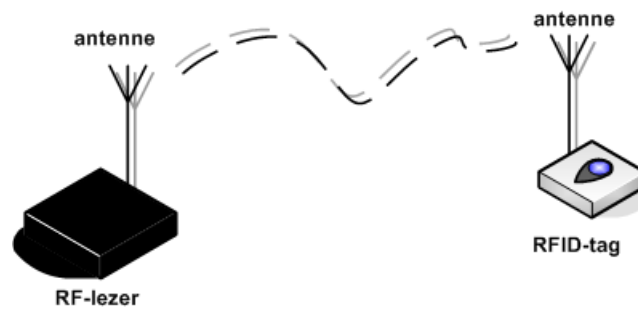
RF-lezer

De RF-lezer (RF-reader) is de meest intelligente component, deze lezer is in staat om informatie te lezen van een zogenaamde tag. De RF-lezer heeft de mogelijkheid om data te moduleren³ en te demoduleren volgens de geïmplementeerde standaard. Sectie C.1.2 geeft een overzicht van de verschillende standaarden

¹http://en.wikipedia.org/wiki/Radio_Frequency_Identification

²<http://www.rfidradiotag.info>

³De techniek om een bronsginaal (bv. een radiogolf) om te zetten in data.



Figuur 3.2: De 3 basiscomponenten in de RFID technologie

binnen RFID technologie.

Antenne

De antenne (die deel uitmaakt van de RF-lezer) kan radiogolven verzenden en ontvangen. Het vermogen waarmee dit signaal uitgezonden wordt is van cruciaal belang voor de maximale communicatieafstand. Indien dit vermogen niet gelimiteerd zou zijn door allerlei politieke wetten, zouden grotere communicatieafstanden mogelijk worden. De plaatsing van de antenne ten opzichte van de tags is ook van groot belang. Wanneer een tag bijvoorbeeld verticaal geplaatst is ten opzichte van de antenne, is het onmogelijk voor de antenne om een signaal te ontvangen.

Tag

De RFID-tag is ongetwijfeld de meest complexe component van de RFID technologie. De tag bestaat uit een microchip die enkele geheugencellen bevat. Tevens kan deze microchip signalen moduleren en demoduleren. Deze microchip is bevestigd aan een kleinere antenne die radiogolven uit de lucht kan plukken.

3.2.3 Soorten tags

Er bestaan drie grote categorieën in RFID-tags nl. *passieve*, *actieve* en *semi-passieve* tags.

Passieve tags

Passieve tags bevatten geen eigen interne energiebron. Toch zijn deze tags in staat zich te identificeren door gebruik te maken van het elektromagnetisch veld van de RF-lezer. Het inkomende radio frequentie signaal (meestal opgevangen via de antenne van de RFID-tag) levert net voldoende energie voor het *CMOS*-interne circuit⁴ om antwoorden te kunnen terugsturen. De meeste passieve tags realiseren dit door het signaal van de ontvanger te weerkaatsen.

De informatie die op de tag staat kan zeer verschillend zijn, gaande van een simpel identificatie nummer of woord tot een korte tekst. Het soort geheugen van RFID tags kan het best beschreven worden als *Non-volatile EEPROM*.

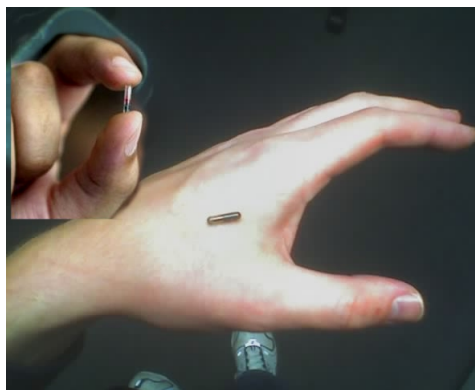
⁴Complementary Metal Oxide Semiconductor, een halfgeleidertechniek die gebruik maakt van metaal-oxide veldeffect-transistoren



Figuur 3.3: Voorbeeld van verschillende soorten passieve RFID-tags

- *Non-volatile* (Niet-vluchtig) geeft aan dat het geheugen op de RFID-tag blijft bestaan ook wanneer er geen energiebron is.
- *EEPROM* (*Electrically Erasable Programmable Read Only Memory*) oftewel ‘*elektrisch wisbaar, programmeerbaar alleen-lezen geheugen*’) geeft aan dat het om geheugen gaat dat met behulp van een elektrische spanning gewist kan worden. Het grote voordeel van dit type geheugen is dat het gewist en opnieuw beschreven kan worden terwijl het is ingebouwd in elektrische apparatuur. Meestal kan dit proces (van wissen en opnieuw beschrijven) zo’n 10 000 tot 20 000 maal herhaald worden, duurdere varianten halen zelfs 100 000 maal.

Doordat passieve tags geen interne energie toevoeren bevatten, kunnen dit soort tags zeer klein gemaakt worden. Zo is het reeds mogelijk om een RFID-tag onder de huid te laten plaatsen. Figuur 3.4⁵ toont het soort passieve tags dat bij mensen in de hand kan worden ingepland.



Figuur 3.4: RFID-tags kunnen reeds ingeplant worden bij mensen

De kleinste RFID-tag meet tegenwoordig 0.15 mm x 0.15 mm en is 7.5 micrometer dik (dunner dan een blad papier)⁶. De goedkoopste RFID-tags (die reeds in een heel aantal grootwarenhuizen worden

⁵<http://www.engadget.com>

⁶<http://www.eetimes.com/news/design/showArticle.jhtml?articleID=179100286>

gebruikt) zijn beschikbaar voor 5 Dollar cent per stuk⁷.

Het is ook mogelijk om een antenne toe te voegen aan een passieve RFID-tag, dan verschilt de grootte van een postzegel tot een postkaart. De reikwijdte van passieve tags verschilt van 10cm tot enkele meters (hiertoe behoren EPC⁸ RFID-tags), afhankelijk van de radiofrequentie en het soort en de grootte van de antenne. De levensduur van RFID-tags is niet gelimiteerd. Tegenwoordig bestaan de meeste RFID-tags uit polymeer⁹ semiconductoren die geen siliconen bevatten. Dit soort RFID-tags zijn veel goedkoper dan de tags op basis van siliconen.

Afhankelijk van het soort tag zijn sommige RFID-tags enkel leesbaar (read-only), andere zijn dan weer leesbaar en herschrijfbaar (rewritable). Voor passieve tags is het onmogelijk om de communicatie tussen een RF-lezer te starten, deze tags kunnen enkel antwoorden. Deze eigenschap wordt aangegeven door de afkorting RTF, oftewel '*Reader Talks First*'. Figuur 3.3¹⁰ toont enkele passieve RFID-tags.

Actieve tags

Actieve tags hebben een eigen energiebron (batterij) die gebruikt wordt om uitgaande signalen te kunnen genereren. Actieve tags zijn over het algemeen veel betrouwbaarder dan passieve tags omdat ze een sessie tussen de RF-lezer en zichzelf kunnen 'leiden'. Door het gebruik van een interne energiebron kunnen actieve tags gegevens overbrengen tegen hogere stroom niveaus. Hierdoor is het mogelijk om ook in meer uitdagende situaties (onder water, tussen metaal, . . .) RFID-tags te gebruiken over grotere afstanden. De meeste actieve tags hebben een bereik van enkele honderden meters en een levensduur tot 10 jaar.

Actieve tags hebben het voordeel dat ze kunnen worden uitgebreid met andere soort sensoren zoals bijvoorbeeld temperatuur sensors, lichtsensors, schok en vibratiesensoren, . . . Een ander voordeel dat actieve tags ten opzichte van passieve hebben, is het feit dat hun geheugen veel groter kan zijn en de verzender de mogelijkheid heeft om extra data toe te voegen aan een actieve tag. De kleinste actieve tags zijn ongeveer zo groot als een munt en kosten meer dan 100 Dollar.

Actieve tags kunnen in tegenstelling tot passieve tags communicatie opstarten met een RF-lezer, dit wordt ook wel '*Tag Talks First*' (TTF) genoemd. Het is dus mogelijk om een soort van peer-to-peer¹¹ communicatie te doen tussen de RF-lezer en de actieve RFID-tag[35].

Figuur 3.5¹² toont een actieve RFID tag die dienst kan doen als *locatietracker* voor verschillende objecten. Deze tag heeft een reikwijdte van 150 meter en is ongeveer zo groot als een postzegel.

Semi-passieve tags

Semi-passieve tags (soms ook wel semi-actieve tags genoemd) zijn, zoals 'semi' al verraadt, RFID-tags die eigenschappen van beide voorgaande types bevat. Semi-passieve tags bevatten een interne energiebron (in de vorm van een batterij) net zoals de actieve soort. Toch behoren ze tot het RTF-type, ze kunnen namelijk (net zoals de passieve tags) enkel antwoorden geven. Semi-passieve tags hebben als voordeel

⁷<http://www.rfidjournal.com/article/articleview/2295/1/128/>

⁸Electronic Product Code

⁹molecuul dat bestaat uit een sequentie van één of meerdere identieke of soortgelijke onderdelen die aan elkaar zijn gekoppeld.

¹⁰<http://www.phidgets.com>

¹¹Manier van communicatie waarin alle netwerkspelers als evenwaardig worden gezien.

¹²<http://mobiqitous.com>



Figuur 3.5: Voorbeeld van een actieve tag

dat ze een (in grootte) beperkte batterij bevatten en toch een vrij grote afstand kunnen overbruggen.

Indeling volgens frequentiebanden

Over het algemeen streven producten er naar om in licentievrije radiobanden te werken, dit is bij RFID niet anders het geval. RFID-technologie strekt zich uit over vier belangrijke frequenties¹³ :

Low Frequency (LF)	125kHz en 134kHz
High Frequency (HF)	13.56Mhz
Ultra high frequency (UHF)	869Mhz tot 950Mhz
Microwave (μ W)	2.45Ghz

Elke frequentieband onderscheidt zich door specifieke eigenschappen[2]. Deze bepalen onder andere :

- manier waarop energie wordt overgebracht, via magnetisch of elektrisch veld.
- reikwijdte
- doordringbaarheid met verschillende materialen (metaal, vloeistof, ...)
- beschikbare bandbreedte
- grootte en vorm van de antenne en spoel van de RFID-tag
- gevoeligheid van menselijk lichaam aan de specifiek frequentieband

Figuur 3.6 geeft een algemeen overzicht van de specifieke eigenschappen van de verschillende frequentiebanden.

In appendix C.1.1 gaan we dieper in op de achterliggende principes en werking van *Radio Frequency Identification*. Tevens bekijken we in appendix C.1.2 de standaarden die er omtrent RFID technologie bestaan.

3.2.4 Barcode vs. RFID

RFID technologie wordt vaak beschouwd als de nieuwe opvolger van de traditionele barcode. Beide aanpakswijzes hebben hun voor en nadelen. Tabel 3.1 geeft een overzicht van de voornaamste voor- en nadelen van barcodes en RFID-tags.

¹³<http://www.morerfid.com/index.php?do=faq&topic=Reader-3&display=RFID>

Frequentie	LF	HF	UHF	μW
Algemene info	Lage frequentie vereist langere en duurere antenne	Bevatten anti-collision techniek	Goede evenwicht tss performatie en leesafstand bij meerdere tags	Vergelijkbaar met UHF. Biedt meest richtbaar signaal bruikbaar voor specifieke doeleinden
Kostprijs	Relatief dure tags, goedkopere lezers	Goedkopere tags tov LF	Dankzij nieuw intern circuit goedkopere tags dan LF en HF	Vergelijkbaar met UHF
Maximale leesafstand (passieve tags)	< 0.5	< 1 m	< 3 m	< 1 m
Snelheid van overdracht	2 kbit/sec	10 kbit/sec	20 kbit/sec	> 100 kbit/sec
Grootte van de tags				
Lezen in omgeving van metaal of vloeistof				

Figuur 3.6: Overzicht eigenschappen van verschillende RFID frequenties

Tabel 3.1: Overzichtstabel RFID-tags vs. Barcodes

	RFID-tags	Barcodes
Leessnelheid	Meerdere tags tegelijkertijd (anti-collision)	1 barcode per leesbeurt
Leesgemak	Geen <i>line of sight</i> ¹⁴ nodig	<i>Line of sight</i> noodzakelijk
Leesafstand	5cm tot 2m afhankelijk van het soort tags en de RF-lezer	Tot 50cm afhankelijk van de barcode-lezer
Robuustheid	Kan tegen vuile en veeleisende (water, stof, ...) werkomgeving	Kan sneller beschadigd worden en is niet bestand tegen veeleisende werkomgevingen.
Data Capaciteit	Van 2Kb tot 64Kb	13 Bytes = 0.0127 Kb
Herschrijfbaar	Ja, afhankelijk van het soort reader en tags	Nee
Kost	Gaande van \$ 0.05 tot \$ 0.50	Zeer goedkoop (drukkosten)
Overig	Kan gedrukt worden maar ook in objecten geïntegreerd worden	Kan enkel gedrukt worden en moet zichtbaar zijn op de verpakking

Hoewel RFID-technologie duurder is, is RFID in trek in de logistieke industrie. Dit is zo omdat RFID toelaat menselijke inbreng voor een groot deel achterwege te laten en het proces over het algemeen betrouwbaarder wordt. Volgende sectie geeft een kort overzicht van enkele toepassingen van RFID in uiteenlopende sectoren.

3.2.5 RFID toepassingen

In volgende secties worden enkele voorbeelden gegeven van toepassingen die gebruik van RFID-technologie.

Veiligheid en Beveiliging

De *C&A*-winkels waren de eerste winkels die reeds in 1975 een anti-diefstal systeem op basis van RFID-tags in België introduceerde. De ‘diefstal-poortjes’ slaan alarm van zodra er een specifieke RFID-tag zich voorbij de lezer verplaatst wordt. Toch is het enigszins merkwaardig dat producten bij de kassa nog steeds worden geïdentificeerd via barcodes hoewel elk product een RFID-tag bevat.

Toegangsbewijzen

Tijdens het WK2006 in Duitsland bevatte elk toegangsticket een RFID-tag, hierdoor werd het vervalsen van een ticket een veel moeilijker klus. Tevens vergemakkelijkte deze tickets de controle in het stadion zelf.

Documenttracking

De *Permeke*-bibliotheek in Antwerpen rust al zijn boeken uit met RFID-tags. Hierdoor kunnen boeken snel en eenvoudig worden uitgeleend en kan er extra informatie worden toegevoegd aan de tags zoals de laatste uitleendatum.

Distributie

De grootste winkelketen ter wereld Walmart eist dat hun leveranciers productpakketten voorziet van RFID-tags. Hierdoor wordt de logistiek van Walmart voor een groot deel vergemakkelijkt en geraken producten sneller in de winkels.

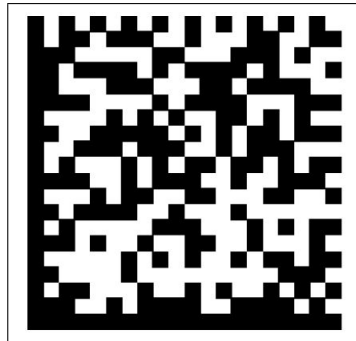
3.3 Semacode Tags

3.3.1 Algemeen

Een *Semacode tag* is een speciaal type van *Data Matrix code*, die op zijn beurt een twee dimensionale matrix barcode is bestaande uit verschillende zwart/wit vierkante regio's. Informatie kan worden geïntegreerd in zulke *Semacode tags* door de kleur van de verschillende regio's aan te passen. De hoeveelheid informatie die een *Semacode tags* kan bevatten wordt bepaald door de grootte van de dimensie van de matrix. In praktijk wordt er meestal gewerkt met *Semacode tags* gaande van enkele bytes tot twee kilobyte. *Semacode tags* kunnen tot een fysiek identificatiemiddel gemaakt worden door ze simpelweg af te drukken. Figuur 3.7 geeft een voorbeeld van een *Semacode tag*.

Semacode tags bevatten eveneens *Error correction codes* waardoor *Semacode tags* die gedeeltelijk beschadigd zijn nog steeds leesbaar zijn. Daardoor zijn deze data matrix codes decodeerbaar door goedkopere beeld-

opname apparatuur, zoals mobiele telefoons en smartphones met ingebouwde camera. *Semacode tags* zijn is volledig vrij voor niet-commercieel gebruik zonder enige restricties.¹⁵



Figuur 3.7: Voorbeeld van een Semacode tag

3.3.2 Identificatie

Doordat zulke *Semacode tags* ook data kunnen bevatten in de vorm van URL's, worden ze interessant om de overbrugging te maken tussen fysieke en digitale wereld. Een voorbeeld hiervan is het inlezen en decoderen van een Semacode tag die een URL bevat, om vervolgens met een webbrowser naar de gedecodeerde URL te surfen. Informatie die in de fysieke wereld als gedrukte 'inkt' aanwezig is, wordt door het nemen van een foto (met bijvoorbeeld een mobiele telefoon) gebruikt in de digitale wereld om een website te bezoeken. Dit is een primitief (maar duidelijk) voorbeeld van een overbrugging die plaatsvindt tussen fysieke en digital wereld.

3.3.3 Gerelateerde technologieën

Er bestaan verschillende alternatieven voor *Semacode tags*, een korte opsomming :

- *QR code*¹⁶ is een variant op *Semacode tags* die zeer veel bijval kent in Japan. *QR code* is geoptimaliseerd voor Japanse karakters doch beweert de organisatie achter Semacode dat *QR codes* minder efficiënt zijn dan *Semacode tags*¹⁷.
- *PDF417*¹⁸ (Portable Data File) is een 2D barcode technologie die tot 2710 karakters kan bevatten.
- *MaxiCode*¹⁹ is een twee dimensionele barcode technologie die gebruikt wordt door *United Parcel Service*²⁰ (UPS) bij de *tracking* en het beheer van postverzendingen. Maxicode tags kunnen snel gelezen worden maar kunnen minder informatie bevatten dan *QR codes* en *Semacode tags*.
- *Atom Tags*²¹ wordt voorgesteld als de *next generation 2D code*, en werkt via *graphical pattern matching*. Hierdoor wordt het mogelijk om (merk)logo's te koppelen aan bepaalde digitale infor-

¹⁵<http://jemmille.com/2007/03/12/semacode-tags-are-cool/>

¹⁶<http://www.denso-wave.com/qrcode/index-e.html>

¹⁷http://semacode.org/about/technical/whitepaper/best_2d_code.pdf

¹⁸<http://www.pdf417.com/>

¹⁹<https://www.aimglobal.org/estore/ProductDetails.aspx?productID=25>

²⁰<http://www.ups.com>

²¹<http://www.gotomobile.com/archives/next-generation-2d-code-atom-tag>

matie. Voordeel hiervan is het feit dat zo'n logo op die manier zowel 'leesbaar' is voor mensen als computers. Helaas is dit project nog in vroeg stadium waardoor het nog niet vrij beschikbaar is.

- *Shot Code*²² is een circulaire barcode ontwikkeld door *High Energy Magic of Cambridge University*. Deze technologie leest data van de datacirkels door de afstand en hoek tussen het middelpunt van de cirkel en de zwarte oppervlakken te berekenen. Het grootste verschil met *data matrix barcodes* is het feit dat *Shot codes* enkel een nummer van 49bits kan bevatten. Dit nummer kan vervolgens opgeslagen worden in een database en gelinkt worden aan URLs.

Semacode tags worden (naast RFID) later in de implementatie van de thesis gebruikt, vermits ze goedkoop en gemakkelijk te creëren zijn. Omdat deze tags via mobiele apparaten kunnen worden ingelezen, zijn ze extra interessant.

²²<http://www.shotcode.com/>

Hoofdstuk 4

Het Semantisch Web

Inhoudsopgave

4.1	Inleiding	27
4.2	Syntax vs. Semantiek	28
4.2.1	Syntax	28
4.2.2	Semantiek	28
4.3	Wat is het Semantisch Web ?	28
4.4	Traditionele web	30
4.5	Semantic Web architectuur	31
4.5.1	XML - XML Schemas	32
4.5.2	RDF - RDF Schemas	32
4.5.3	OWL	36
4.5.4	De relatie tussen RDF, RDFS en OWL	38
4.5.5	RDF/A	38
4.6	Kritieken	39
4.7	Het Semantisch Web in deze thesis	39

4.1 Inleiding

In dit hoofdstuk bestuderen we het *Semantisch Web*. Het is van belang de gedachten en idee achter het *Semantisch Web* te begrijpen, om een beter inzicht op de uiteindelijke implementatie in deze thesis te krijgen. Dit is niet zo voor eindgebruikers van applicaties die gebruik maken van het *Semantisch Web*, het is niet nodig dat zij de technische achtergrond ervan begrijpen.

Kort samengevat biedt het *Semantisch Web* ons de mogelijkheid om data op een gestructureerde manier te beschrijven waardoor deze data begrijpbaar is voor zowel computers als menselijke gebruikers. Daarenboven wordt zulke *semantische* data ook makkelijker *herbruikbaar* voor allerlei doeleinden. In de volgende secties wordt het *Semantisch Web* meer vanuit een algemeen standpunt bekeken vermits het gezien moet worden als een technologie die domeinoverschrijdend werkt. In functie van de thesis wordt het gebruikt als een technologie die ons toestaat *augmented physical objects* in de digitale wereld op een transparante manier te definiëren.

4.2 Syntax vs. Semantiek

Om te doorgronden waarrond het *Semantic Web* juist draait, is het van belang enkele termen in dit domein beter te beschrijven.

4.2.1 Syntax

In de computerwereld wordt het Engelse woord *syntax* gebruikt als de term voor de grammatica van een programmeertaal. In een meer complete beschrijving zou men *syntax* eerder kunnen beschrijven als

‘De presentatiewijze (de vorm van willekeurige objecten, ook wel symbolen genoemd), waarin een formeel notatiesysteem wordt aangenomen volgens een aanvaarde en gedeelde conventie.’

Computerberekeningen bestaan uit een manipulatie gebaseerd op welbepaalde regels van deze symbolen. Deze regels en manipulaties (ook wel *algoritmes* genoemd) zijn uitsluitend gebaseerd op de machinale vorm van deze symbolen¹. Deze berekeningen zijn niet gebaseerd op de betekenis van de symbolen, hoewel de combinatie van deze symbolen interpreteerbaar is door het menselijke verstand tot een bepaalde betekenis.

Met andere woorden mensen kunnen bij een opeenvolging van welbepaalde symbolen (bv. ‘*BOEK*’), een welbepaalde betekenis geven aan deze combinatie van symbolen (nl. ‘*het woord dat een boek voorstelt*’).

4.2.2 Semantiek

Integenstelling tot *syntax*, is bij *semantiek* de betekenis van symbolen van belang en niet de vorm of de manipulaties van hun vormen. Volgens Vandale² is ‘semantiek’ *de leer van de betekenis van de woorden en woordgroepen*. Waar *syntax* de presentatiewijze vastlegt, legt *semantiek* de betekenis achter deze presentatie vast. Een nog steeds onopgelost gerelateerd probleem binnen de wereld van taalkunde is het ‘*symbol grounding problem*’^{3,4}.

4.3 Wat is het Semantisch Web ?

Om het met de woorden van de ontwikkelaars te definiëren : ‘Het Semantisch Web is een uitbreiding op het huidige World Wide Web (WWW), waar informatie een goed gedefinieerde betekenis krijgt en op die manier een betere samenwerking tussen computers en mensen toelaat’⁵.

Het *Semantisch Web* is dus een project dat de intentie heeft om een universeel medium te creëren voor het uitwisselen van informatie. Dit wordt bereikt door een semantische betekenis, in de zin dat het verstaanbaar is voor machines, te geven aan de inhoud van documenten aanwezig op het web. Het is ontwikkeld door het W3C⁶ (World Wide Web Consortium) met als grote drijvende kracht Professor Tim Berners-Lee. Om deze uitbreiding op het traditionele web te realiseren wordt er gebruik gemaakt van

¹<http://www.wisageek.com/what-is-syntax.htm>

²<http://www.vandale.be>

³<http://www.ecs.soton.ac.uk/~harnad/Papers/Harnad/Harnad90.sgproblem.html>

⁴Bij dit probleem probeert men te achterhalen hoe symbolen hun betekenis krijgen. Redenering van symbolen gebeurt in de menselijke hersenen en het is dan ook aannemelijk dat dit gerelateerd is aan onze sensori-motor capaciteit. Er bestaat echter nog steeds veel discussie over de manier waarop symbolen met elkaar worden geconnecteerd in het menselijke brein, volgens velen komt hier geen ‘betekenis’ van symbolen aan te pas

⁵<http://infomesh.net/2001/swintro/>

⁶<http://www.w3c.org>

verschillende standaarden, markup languages en gerelateerde *processing tools*. Kort en bondig zou men kunnen concluderen dat het *Semantisch Web* een ‘web van data’ is⁷.

Dagelijks gaat de gemiddelde computergebruiker met een heleboel data om. Tegenwoordig kunnen computergebruikers ondermeer bankgegevens, persoonlijke foto’s en kalenderafspraken online bekijken. Waarom is het vrijwel onmogelijk om foto’s of bankafschriften geïntegreerd in een persoonlijke kalender te bekijken ?

Het antwoord is verrassend eenvoudig: we hebben namelijk geen ‘web van data’. Alle data wordt gecontroleerd door programma’s en elk programma houdt de data voor zich. Er is dus geen sprake van informatie-uitwisseling tussen verschillende (web)applicaties en bijgevolg zijn de hierboven aangehaalde voorbeelden in de overgrote meerderheid van de gevallen onmogelijk. Figuur 4.1⁸ toont de oorspronkelijke illustraties die Tim Berner-Lee in 1994 gebruikte om het verschil tussen het traditionele web (dat wordt verder wordt beschreven in 4.4) en het (toen gloednieuwe) concept van het *Semantic Web* duidelijk te maken.

We verduidelijken de illustratie uit fig 4.1 wordt getoond aan de hand van een kort voorbeeld. Stel dat data A en data B zich op een netwerk bevinden. In het traditionele web kunnen we deze data in een afzonderlijke HTML documenten gebruiken en verwijzingen naar elkaar voorzien. Het traditionele web biedt ons de mogelijkheid om ‘een’ link te leggen tussen 2 delen data, vereenvoudigd opgeschreven als (A,B). In het *Semantisch Web* wordt het mogelijk om eveneens de relatie tussen de 2 delen data te beschrijven, vereenvoudigd opgeschreven als (A, relatie, B). Zo zouden we kunnen beschrijven dat data A en B beiden een boek voorstellen of beiden een beschrijving van een persoon zijn.

Het *Semantisch Web* steunt op twee belangrijke pijlers :

1. Gemeenschappelijke formaten voor de uitwisseling van *data*. In het oorspronkelijke web is er enkel sprake van uitwisseling van *documenten*.
2. De taal voor het vastleggen van de relaties tussen data en reële objecten. Hierdoor kan een persoon of een machine starten in een database en zich digitaal verplaatsen doorheen een enorme set van andere databases. Deze databases zijn gekenmerkt doordat ze niet via bedrading in relatie staan, maar een semantische relatie met elkaar hebben (bijvoorbeeld over hetzelfde onderwerp gaan).

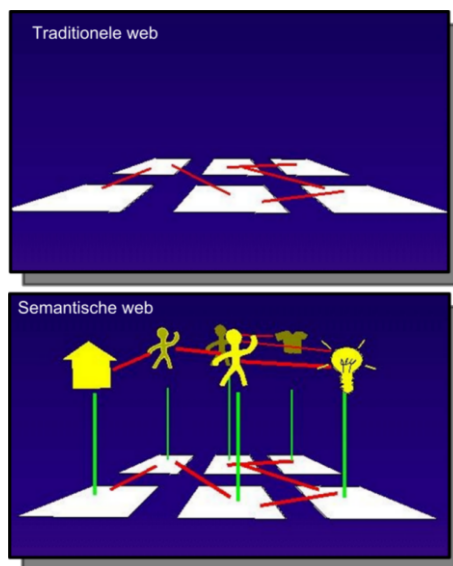
Simpel uitgelegd draait het *Semantisch Web* hoofdzakelijk rond *slimme* goed gestructureerde data. De machines die deze data uiteindelijk interpreteren hoeven geen geavanceerde artificiële intelligentie te bevatten. Hoewel het *Semantisch Web* een groot aantal hulpmiddelen biedt om semantische relaties te leggen tussen documenten, is het in de grond nog steeds een ‘*syntactisch* web’. Hiermee wordt bedoeld dat het *Semantisch Web* er in slaagt om *syntax* op het World Wide Web toegankelijker te maken op veel vlakken, maar er niet in slaagt ook computer interpreteerbare *semantiek* aan deze *syntax* te geven.

De interpretatie van syntax van het *Semantisch Web* dient nog steeds te gebeuren door kennisoverdracht van menselijke ontwikkelaars. Met andere woorden ontwikkelaars moeten nog steeds syntax interpretern en acties aan deze interpretatie koppelen. Het *Semantisch Web* is dus eerder een uitblinker in de categorie van de syntactische webs.

Volgende secties trachten het *Semantisch Web* beter uit te leggen aan de hand van praktische voorbeelden.

⁷<http://en.wikipedia.org/wiki/Semantic.Web>

⁸<http://www.w3.org/Talks/WWW94Tim/authlevel>



Figuur 4.1: Traditionele web vs. Semantisch Web

4.4 Traditionele web

Het huidige World Wide Web is tegenwoordig in de eerste plaats gebaseerd op documenten die geschreven zijn in HyperText Markup Language (HTML). HTML wordt gebruikt om documenten te beschrijven die tekst bevatten afwisselend gecombineerd met multimedia objecten (video's, afbeeldingen, ...) en invulformulieren. HTML heeft beperkte mogelijkheden om stukken tekst te kunnen classificeren op een pagina afhankelijk van de rol die ze spelen in het document.

Hieronder volgt een voorbeeld dat meteen de tekortkomingen van HTML duidelijk maakt. PC-shop 'Powertools' plaats zijn prijslijst online met volgende items :

- NR10235 - WD Mybook 320GB - 155 euro
- NR10536 - Maxtor 500GB - 240 euro

Vermits 'Powertools' nog nooit van het *Semantisch Web* gehoord heeft, plaatst het zijn prijslijsten in een zelfgeschreven HTML-formaat online. Dit achterliggende HTML-code van deze items ziet er dan als volgt uit :

```
<table>
<tr><td>NR10235</td><td>WD Mybook 320GB</td><td>155 euro</td></tr>
<tr><td>NR10536</td><td>Maxtor 500GB</td><td>240 euro</td></tr>
</table>
```

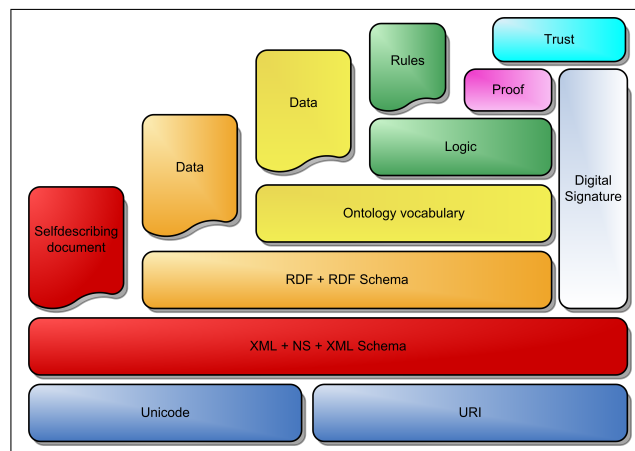
In dit voorbeeld is het voor een HTML-parser (bijvoorbeeld een webbrowser) onmogelijk om beweringen rond de eigenlijke data te doen. Uit de HTML-code kan bijvoorbeeld niet afgeleid worden dat artikel met nummer 'NR10536' een Maxtor harde schijf is met een prijs van 240 euro. HTML kan enkel beschrijven dat de tekst 'NR10536' moet gepositioneerd worden in een cel (<td>) en de tekst 'WD Mybook 320GB' en '155 euro' eveneens in dezelfde rij (<tr>) geplaatst moeten worden. Er is ook geen mogelijkheid

in HTML om te beschrijven dat deze beide artikelen behoren tot de hogere categorie van ‘*harde schijven*’.

Het *Semantisch Web* richt zich op zulke verkortkomingen van het traditionele web en gebruikt hiervoor meerdere standaarden (zie sectie 4.5). Doordat de syntax in het *Semantisch Web* relateerbaar is aan een bepaalde betekenis, wordt het voor machines mogelijk om elders informatie op te halen die overeenkomsten vertoont op het vlak van betekenis! In dit voorbeeld zou men door gebruik te maken van het *Semantisch Web*, bijvoorbeeld de webpagina vrij eenvoudig automatisch kunnen laten aanvullen met vergelijkende prijzen van identieke producten in andere PC-shops.

4.5 Semantic Web architectuur

Figuur 4.2⁹ geeft de architectuur van het *Semantisch Web* volgens Sir Prof. Tim Berners-Lee weer.



Figuur 4.2: Semantisch Web architectuur

Hieronder volgt een opsomming van de belangrijkste onderdelen op dit moment uit de oorspronkelijke architectuur van het *Semantisch Web* :

- RDF : eenvoudig data model (geen taal) dat toelaat om referenties te leggen naar objecten (ook wel ‘resources’) en een relatie te beschrijven ten opzichte van deze objecten. RDF maakt onder andere gebruik van een XML-syntax om de semantiek van documenten vast te leggen.
- RDF Schema (RDFS) : woordenschat die toelaat om de structuur van RDF-resources te beschrijven. Met RDFs (dat RDF als beschrijvende kracht gebruikt) kan men dus RDF-vocabulaires (ook wel ‘ontologieën’) vastleggen. Een voorbeeld van een RDF Schema is de ‘*Dublin Core element set*’.
- OWL : uitbreiding op RDFS, nl. een woordenschat die als extensie kan gebruikt worden voor het (meer in detail) beschrijven van eigenschappen en klassen. Zo is het bijvoorbeeld mogelijk om met OWL de relaties tussen classes aan te geven. OWL zegt niets over het uiteindelijke gebruik van relaties in een uiteindelijke applicatie.

Vermits deze termen nogal vaag kunnen overkomen, is het van belang dat we stap voor stap de functie van de verschillende onderdelen uit deze architectuur (zie figuur 4.2) verduidelijken.

⁹naar <http://rhizomik.net/roberto/thesis/figures/sw-stack-2002.png>

In de volgende secties zal duidelijk worden waarom er nood is aan een *Semantisch Web*. We trachten in volgende secties te werken naar een oplossing waarbij we documenten kunnen voorzien van semantiek, waardoor zowel mensen als computers de inhoud van een document kunnen gebruiken of herbruiken.

4.5.1 XML - XML Schemas

eXtensible Markup Language¹⁰ (XML) is een standaard voor het definiëren van formele markup-talen voor de representatie van gestructureerde gegevens in de vorm van tekst. Deze representatie is zowel leesbaar voor machines als leesbaar voor de mens. XML biedt een gestandariseerde syntax die kan gebruikt worden voor allerlei doeleinden.

XML is niet alleen geschikt om gegevens in op te slaan, maar wordt de laatste tijd ook meer en meer gebruikt om gegevens via het internet te versturen. Een bekend voorbeeld hiervan is de Simple Object Access Protocol (SOAP)¹¹, deze technologie maakt gebruik van XML om gegevens uit te wisselen.

Figuur 4.3 geeft aan dat XML op zich onvoldoende is om een document een bepaalde semantiek te geven. De XML-tags kunnen dan wel betekenisvol zijn voor een menselijke gebruiker, voor een computer blijft het een opeenvolging van tekens. Toch kan XML al een bepaalde structuur aan een document geven, zonder dat hierbij de semantiek duidelijk is.

XML Schema¹² is een standaard waarmee de structuur van XML-documenten kan worden gedefinieerd. De formele taal van XML Schema, XML Schema Definition Language (XSD), is tevens gestandariseerd door het W3C. Het is ontwikkeld als een opvolger van het eerder ontwikkelde Document Type Definition (DTD)¹³.

Door middel van XML Schema kunnen XML-documenten formeel worden gespecificeerd, en op grond daarvan worden gevalideerd (door bijvoorbeeld XML-validators¹⁴). Een XML Schema geeft onder meer aan welke elementen mogen voorkomen in een XML document, waar deze elementen mogen voorkomen, aan welke kenmerken deze moeten voldoen, . . . Merk op dat XML Schemas nog steeds functie doen als syntax en er geen semantiek bij te kijken komt.

Een web van data (zoals het *Semantisch Web*) zouden we kunnen creëren door gebruik te maken van XML Schemas. Door de semantiek via XML Schema beschrijvingen af te spreken, zouden we gemeenschappelijke XML-termen van verschillende documenten met elkaar kunnen verbinden. En op die manier semantische relaties tussen syntax van verschillende documenten kunnen opbouwen. Dit wordt verduidelijkt in figuur 4.4. Toch bevat deze aanpak ook problemen: als een andere ontwikkelaar namelijk een ander XML schema gebruikt, wordt het onmogelijk op een eenvoudige manier deze documenten Semantisch met elkaar te verbinden (zie figuur 4.5). Het *Semantisch Web* biedt hiervoor onder andere een oplossing.

4.5.2 RDF - RDF Schemas

RDF

¹⁰<http://www.w3.org/XML/>

¹¹<http://www.w3.org/TR/soap/>

¹²<http://www.w3.org/XML/Schema>

¹³<http://www.xmlfiles.com/dtd/>

¹⁴<http://validator.aborla.net/>



Figuur 4.3: XML



Figuur 4.4: XML Schemas



Figuur 4.5: XML Schemas zijn te weinig flexibel voor de opbouw van een Semantisch Web

Resource Description Framework (RDF)¹⁵ is een eenvoudig data model (geen taal) dat toelaat om referenties te leggen naar objecten (ook wel ‘resources’) en een relatie te beschrijven ten opzichte van deze objecten. RDF kan gebruik maken van XML als syntax, het is een gestandariseerd model en wordt gebruikt door het W3C.

De ontwikkeling van RDF werd gestart door *Ramanathan V. Guha* toen hij werkte voor *Netscape Communications Corporation*. Het model is gebaseerd op de idee om kenmerken van bronnen uit te drukken in de vorm van een *onderwerp-gezegde-voorwerp* (subject-predicate-object). In RDF-termen wordt dit ook wel ‘*een triple*’ genoemd. Figuur 4.6¹⁶ geeft voorbeeld van zo een RDF-triple. Het *onderwerp* is in essentie de bron die beschreven wordt. Het *gezegde* geeft aan welk kenmerk/aspect van de bron beschreven wordt. Het *voorwerp* geeft de waarde van dat kenmerk/aspect aan. De terminologie van RDF is ontleend aan de logica en taalkunde.



Figuur 4.6: RDF Triple voorbeeld

RDF biedt een effectief mechanisme om objecten te verbinden door middel van URIs (in het geval van het web meestal URLs). Uniform Resource Identifier (URI)¹⁷, wordt gebruikt om bronnen te kunnen identificeren. Het hoofddoel van deze identificatie is het mogelijk maken om interactie te kunnen hebben met de representatie van de bron over een netwerk. Een voorbeeld van een URI is een telefoonnummer, zo'n URI maakt het mogelijk dat een bepaald telefoonnummer interactie kan hebben met een specifiek ander telefoonnummer over een netwerk. In het WWW wordt er meestal gesproken over Uniform Resource Locator (URL)¹⁸, een subklasse van een URI. Een URL identificeert een specifieke bron en zijn locatie op het internet.

Praktisch gezien zal RDF er voor zorgen dat XML syntax wordt voorzien van een identifier (URL) die de semantiek van een bepaalde element in de syntax vastlegt. Er zijn tegenwoordig verschillende soorten syntaxen voor handen om een RDF-document te beschrijven, de voornaamste zijn :

- RDF/XML¹⁹
- Notation N3 (N3)²⁰
- Turtle²¹

Figuur 4.5.2 geeft een vergelijking tussen drie van deze representaties. Voor de gemakkelijker wordt de turtle-representatie samen met de N3 notatie gezet, vermits beide representaties sterk op elkaar lijken. De N3/Turtle representatie is duidelijk veel compacter en leesbaarder (voor mensen) dan de oorspronkelijke RDF/XML notatie. Tim Berners-Lee heeft ook bij de ontwikkeling van de N3 notatie zijn steentje

¹⁵<http://www.w3.org/RDF/>

¹⁶naar <http://www.w3.org/2004/02/skos/core/guide/img/ex-triple.png>

¹⁷<http://www.w3.org/Addressing/>

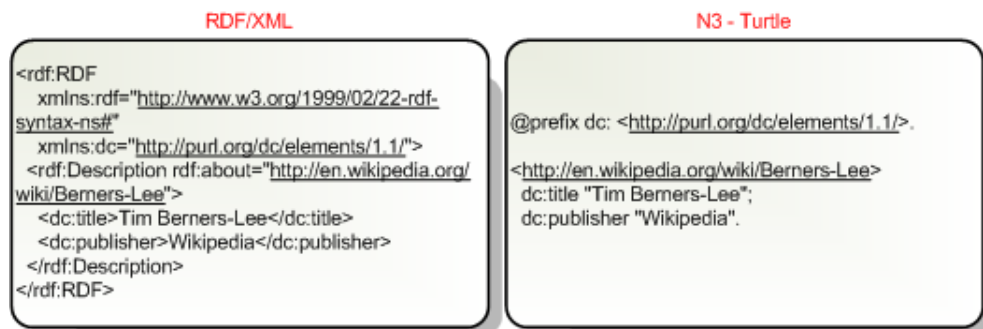
¹⁸<http://nl.wikipedia.org/wiki/URL/>

¹⁹<http://www.w3.org/TR/rdf-syntax-grammar/>

²⁰www.w3.org/DesignIssues/Notation3

²¹<http://www.dajobe.org/2004/01/turtle/>

bijgedragen.

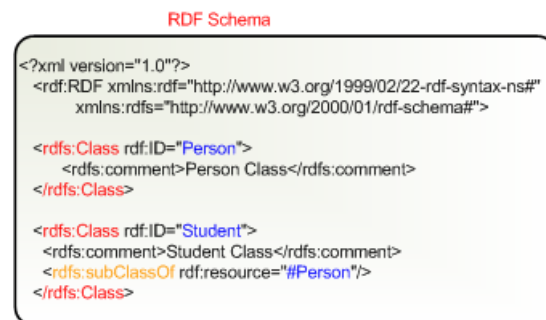


Figuur 4.7: De verschillende RDF notaties

RDF Schemas

RDF Schemas (RDFS)²² is een kennis-representatie taal die de basiselementen voor het beschrijven van ontologieën²³ (ook wel ‘RDF vocabularies’) biedt, dit om uiteindelijk RDF documenten te kunnen structureren. Het biedt dus een mechanisme om eigenschappen van objecten en klassen van objecten te definiëren.

‘RDF vocabularies’ zijn noodzakelijk in het *Semantisch Web* vermits deze de herbruikbaarheid voor een groot deel bepalen. Een voorbeeld van een RDF Schema wordt getoond in figuur 4.5.2. Aan de hand van het RDF Schema kan men een persoon beschrijven en een student. Vermits een student een specifiekere soort van persoon voorstelt, wordt het in RDFS beschreven als een *SubClass* van de *Person*-klasse. Met RDFS kunnen we dus vastleggen welke klassen een RDF-document bevat en wat de eigenschappen (properties) van deze klassen zijn.



Figuur 4.8: Voorbeeld van een RDF Schema

²²<http://www.w3.org/TR/rdf-schema/>

²³Een ontologie is het product van een poging een uitputtend en strict conceptueel schema te formuleren over een bepaald domein. Een ontologie is typisch een hiërarchische datastructuur die alle relevante entiteiten en hun onderlinge relaties en regels binnen dat domein bevat, zoals bij een domein ontologie het geval is. Het gebruik van het woord binnen de informatica is afgeleid van het veel oudere gebruik van het woord ontologie binnen de filosofie.

Aan de hand van RDFS werd er in deze thesis een (shopping)ontologie beschreven. Meer details hierover volgen in sectie 7.2 van deel III.

4.5.3 OWL

Wat kan men met RDFS niet beschrijven ?

- Het is niet mogelijk om te beschrijven dat twee klassen hetzelfde voorstellen.
- Het is onmogelijk om te definiëren dat eigenschappen van klassen meerdere waarden mogen hebben.
- Het beschrijven van relaties tussen verschillende klassen is tevens niet mogelijk. Termen als transitiviteit, symmetrie, inverse, . . . zijn onmogelijk te beschrijven in RDFS.

Om deze tekortkomingen te kunnen beschrijven, kan er gebruik gemaakt worden van Web Ontology Language (OWL)²⁴. OWL is afgeleid van een andere standaard nl. DAML+OIL Web Ontology Language²⁵. OWL representeert de semantiek van elementen in vocabulaires en de relaties tussen de elementen uit deze vocabulaires. OWL gaat nog een stapje verder dan RDFS door het mogelijk te maken verschillende aspecten van relaties tussen elementen te beschrijven. OWL wordt gezien als een belangrijke technologie in de toekomstige implementatie van het *Semantisch Web*.

Tegenwoordig bestaan er drie verschillende OWL versies :

1. OWL Lite
2. OWL DL
3. OWL Full

OWL *Lite* en OWL *DL* zijn ontwikkeld zodat elk statement binnen een eindige tijd kan afgewerkt. OWL Full daarentegen kan oneindige lussen bevatten. Verder verschillen deze versies onderling in de mate van vrijheid in beschrijvingen.

Figuur 4.5.3 geeft een voorbeeld van een OWL beschrijving die gebruik maakt van ‘owl:inverseOf’. In dit geval wordt er een blog beschreven. Door gebruik te maken van OWL kunnen we vastleggen dat ‘als een blog een auteur heeft’, we eveneens weten dat ‘de auteur de blog geschreven heeft’. Hiermee hebben we via OWL de *inverse* eigenschap op een relatie vastgelegd.

Figuur 4.10 toont het gebruik van RDF en enkele reeds bestaande OWL-ontologieën nl. FOAF²⁶ en *Resume*²⁷.

Door het gebruik van ontologieën kan data aan elkaar gelinkt worden (web) en ontstaat er data die semantiek bevat (‘slimme’ data). Hoe deze semantiek wordt geïnterpreteerd en waarvoor ze uiteindelijk gebruik wordt, wordt overgelaten aan de ontwikkelaars van *Semantic Web* applicaties.

Vermits de mogelijkheden die OWL ons biedt om ontologieën te beschrijven een *overkill* waren voor deze thesis, is er gekozen voor een eenvoudigere beschrijving in RDFS.

In de hoger gelegen gedeeltes van de *Semantisch Web* architectuur vinden er nog steeds verschuivingen plaats en het is dan ook goed mogelijk dat de architectuur er binnen enkele jaren in de hoogste lagen compleet anders uitziet.

²⁴<http://www.w3.org/TR/owl-features/>

²⁵<http://www.daml.org/>

²⁶Friend of a Friend - <http://www.foaf-project.org>

²⁷Resume : RDF schema for describing resumes - <http://captsolo.net/semweb/resume/cv.rdfs>

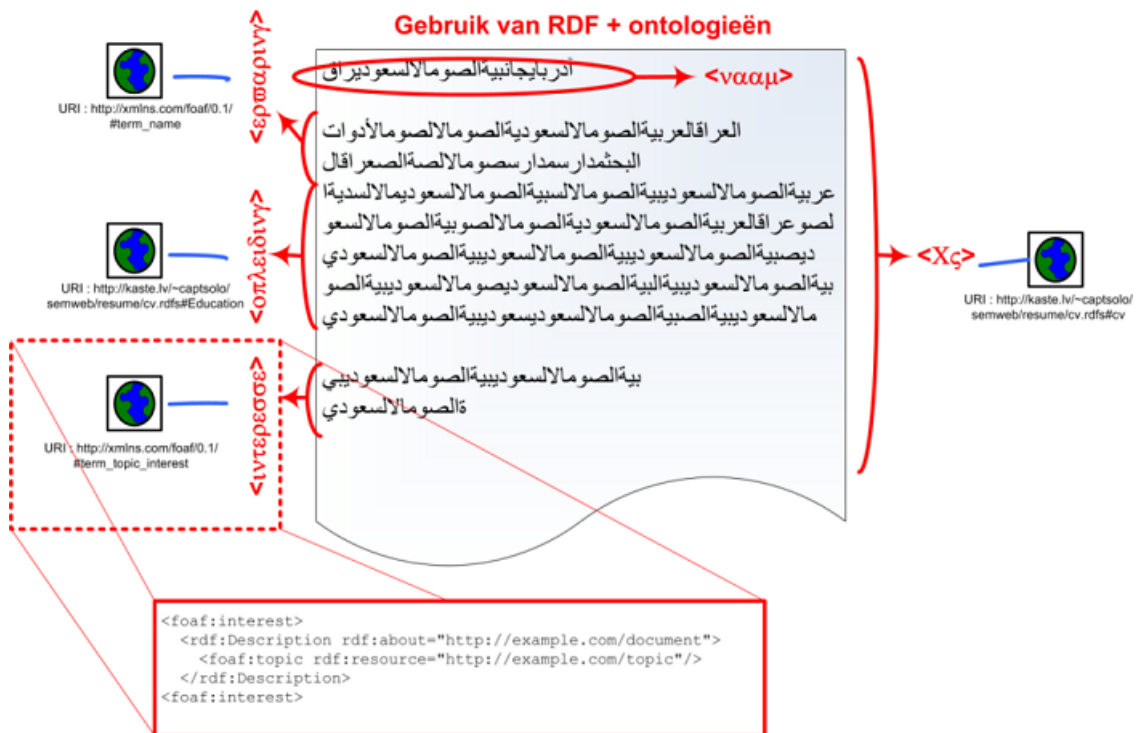
```

    OWL

    <owl:ObjectProperty rdf:ID="writes">
    <rdfs:domain rdf:resource="#Author" />
    <rdfs:range rdf:resource="#Blog" />
    <owl:inverseOf rdf:resource="#hasAuthor" />
    </owl:ObjectProperty>

    <owl:ObjectProperty rdf:ID="hasAuthor">
    <rdfs:domain rdf:resource="#Blog" />
    <rdfs:range rdf:resource="#Author" />
    <owl:inverseOf rdf:resource="#writes" />
    </owl:ObjectProperty>
    
```

Figuur 4.9: Voorbeeld van een OWL

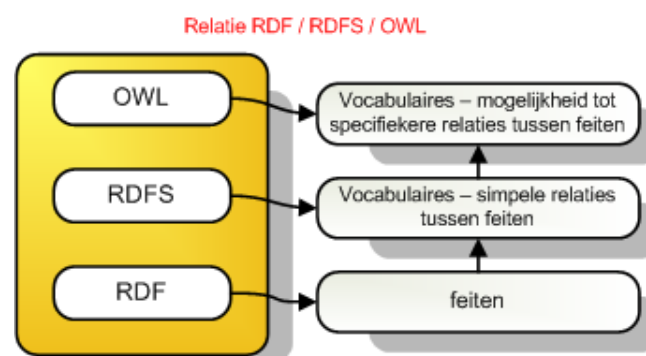


Figuur 4.10: RDF - RDF Schemas - Ontologieën

4.5.4 De relatie tussen RDF, RDFS en OWL

Figuur 4.11 toont een overzicht van het specifieke doel RDF, RDFS en OWL.

- RDF legt simpele feiten vast. Bijvoorbeeld *‘persoon X heeft als naam ‘Ruben’*.
- RDFS maakt het mogelijk een woordenschat te beschrijven zoals bijvoorbeeld *‘persoon X is een Student’*.
- OWL op zijn beurt maakt het mogelijk om relaties als *‘de studenten uit schema A zijn van dezelfde soort als de studenten uit schema B’* te beschrijven. Tevens kan men via OWL de relaties die bestaan tussen klassen in een bepaalde ontologie specifiekere beschrijven (Bijvoorbeeld *‘een student behoort toe aan juist één school’*).



Figuur 4.11: RDF - RDFS - OWL

4.5.5 RDF/A

Het traditionele web is een enorm groot web dat bestaat uit documenten. Om het traditionele web om te vormen in een *Semantisch web*, is het noodzakelijk dat men semantiek kan gebruiken in documenten op het web. In de overgrote meerderheid van de gevallen zijn deze documenten HTML-pagina's. Er zijn enkele standaarden die het mogelijk maken om semantische informatie (die wordt beschreven met de eerder aangehaalde technologie) te integreren met bestaande HTML pagina's. De meest bekende (W3C) standaarden zijn :

1. RDF/A (ook wel RDFa)²⁸
2. Gleaning Resource Descriptions from Dialects of Languages (GRDDL)²⁹

Figuur 4.12 geeft een voorbeeld van een stukje RDFa-code. Deze code beschrijft enkele eigenschappen (*auteur, titel, datum*) van een bepaalde webpagina. xHTML-elementen worden vermengd met semantische informatie die achteraf gebruikt kan worden door applicaties. In het voorbeeld wordt er gebruik gemaakt van de *Dublin Core-ontologie*³⁰ (DC) die ons toelaat bepaalde aspecten van webpagina's semantisch te beschrijven. Andere (web)applicaties kunnen door het gebruik van deze semantische informatie

²⁸<http://www.w3.org/2001/sw/BestPractices/HTML/2005-rdfa-syntax/>

²⁹<http://www.w3.org/TR/grddl/>

³⁰<http://dublincore.org/documents/dces/>

dus achterhalen wie de auteur is van de webpagina ‘<http://www.abc.be>’.

In sectie 6.3.1 in deel III wordt een RDFa parser voorgesteld. Deze parser laat toe om semantische informatie (in de vorm van RDF *N3-triples*) uit een pagina die RDFa bevat te extraheren. Eveneens maakt RDFa deel uit van de architectuur die wordt voorgesteld in hoofdstuk 6.

```
<div xmlns:dc="http://purl.org/dc/elements/1.1/" about="http://www.abc.be">
  <span property="dc:title">Titel van de webpagina</span>
  <span property="dc:author">Ruben Thys</span>
  <span property="dc:date">2007-01-01</span>
</div>
```

Figuur 4.12: Voorbeeld RDFa-code

4.6 Kritieken

- Privacy is een van de problemen waarmee het *Semantisch Web* te maken krijgt. Doordat alle data koppelbaar wordt, bestaat de vrees dat bedrijven en regeringen totale controle krijgen over persoonlijke informatie. De FOAF-ontologie (specifiek voor het beschrijven van personen) laat bijvoorbeeld reeds toe om coördinaten aan personen te koppelen, hierdoor zou anonimiteit in het niets verdwijnen.
- Het kost meer tijd om een applicatie te voorzien van alle nodige semantische informatie. Zodra de voordelen opwegen tegen deze applicatiekost zullen grotere bedrijven misschien reeds andere (semantische) technieken implementeren.
- Het *Semantisch Web* staat nog steeds in kinderschoenen en er is dan ook geen enkele zekerheid dat een applicatie die semantische technieken implementeert ook effectief bruikbaar wordt.
- Het principe rond het *Semantisch Web* is moeilijk aan te brengen, bij gebrek aan werkbare applicaties is er weinig nood om deze technologie begrijpbaar te maken.

4.7 Het Semantisch Web in deze thesis

Het *Semantisch web* speelt een belangrijke rol in deze thesis. In de uiteindelijke implementaties van verschillende applicaties in deze thesis (zie hoofdstuk 7), worden de voordelen van het *Semantisch Web* zo goed mogelijk uitgebuit. Wat zijn nu juist de belangrijkste voordelen van het *Semantisch Web* ?

- Informatie (op het internet maar ook daarbuiten) kan interpreteerbaar gemaakt worden voor computers. Deze computers kunnen dus bepaalde informatie gaan selecteren uit semantisch beschreven data.
- Vermits informatie interpreteerbaar is voor computers, kan deze informatie ook gemakkelijker worden herbruikt voor andere doeleinden. Informatie over een boek op *Amazon.com* kan bijvoorbeeld automatisch worden herbruikt op een andere website (bv. een website die boekbesprekingen aanbiedt). Deze boekbesprekingen-website, kan informatie over boeken automatisch ophalen, door een computer naar deze informatie te laten zoeken. Zo zal deze computer op *Amazon.com* deze informatie kunnen selecteren doordat ze semantisch eenduidig beschreven is. Een computer zal dus naar informatie kunnen zoeken, op een niveau dat vergelijkbaar is met mensen. Door het gebruik van *Semantic Web* technieken wordt informatie die bijvoorbeeld op *Amazon.com* te vinden is, dus ook vrij eenvoudig herbruikbaar op een andere website.

Hoofdstuk 5

Web Services

Inhoudsopgave

5.1	Inleiding	40
5.2	Algemeen	40
5.3	Principe van een web service	42
5.4	Web services lagenmodel	42
5.4.1	Network	43
5.4.2	XML-based Messaging - SOAP	43
5.4.3	Service Description - WSDL	43
5.5	Semantische web services	44
5.5.1	Algemeen	44
5.5.2	OWL-S	45
5.5.3	WSDL-S	45
5.5.4	SAWSDL	46
5.5.5	WSDL 2.0	47
5.5.6	WSMO	47
5.6	Literatuur	47
5.7	Toekomst	48

5.1 Inleiding

In dit hoofdstuk bekijken we het idee achter *web services* van korter bij. Vervolgens verdiepen we ons in een bepaald soort van *web services* nl. *semantische web services*. Deze zullen een belangrijk deel uitmaken van de architectuur die ons toelaat problemen op te lossen die te wijten zijn aan de *fysiek-digitaal kloof*. Het is van belang het principe rond het *Semantisch Web* (zie hoofdstuk 4) te begrijpen, om een beter inzicht te krijgen in *semantische web services*.

5.2 Algemeen

Voor we dit hoofdstuk aantvatten kunnen we ons eerst de meest voor de handliggende vraag stellen nl. ‘*wat is een web service juist ?*’. Om dit volkomen te begrijpen, is het van belang dat we het begrip ‘*service*’ begrijpen. Een *service* zou kort omschreven kunnen worden als *een extern zichtbaar deel van specifieke functionaliteit in een business proces* [6]. Een postkantoor biedt bijvoorbeeld de service aan om postzegels te kopen. Zo’n service wordt aangeropen (ook wel ‘*invocation*’ genoemd) simpelweg door

het postkantoor binnen te stappen, te wachten aan het loket en de bediende te vragen om een postzegel. Belangrijk hierbij is dat er geen kennis moet zijn over het ‘business proces’ van het postkantoor. Men hoeft bijvoorbeeld niet te weten waar de postzegels worden gedrukt, waar ze worden opgeslagen en in welke verpakking ze worden verkocht.

Op het eerste zicht lijkt dit doodnormaal maar waar haal je bijvoorbeeld de kennis om te weten dat een postkantoor postzegels verkoopt? Hoe weet je waar er een postkantoor te vinden is? En hoe weet je op welke manier je naar postzegels kan vragen aan de postbediende? Het antwoord is simpel: de mens leert simpelweg doorheen zijn leven het antwoord op al deze vragen. En bouwt doorheen zijn leven genoeg contextuele informatie op om de taak die aan dit proces gekoppeld is tot een goed einde te brengen.

Wanneer men interactie wil laten plaatsvinden tussen twee verschillende applicaties, moet alle contextuele informatie bekomen worden op het tijdstip van de interactie. *Web services* maken de interactie tussen applicaties mogelijk door de protocollen en de interfaces die nodig zijn voor de interactie te definiëren. *Web services* kunnen we bijgevolg beschrijven als ‘*een extern zichtbaar deel van een specifieke (business)functie die is opengesteld via het web via een goed gedefinieerde interface, en beschreven en aangeroepen via standaard webprotocollen.*’¹

Web services zijn softwaresystemen die ondersteuning bieden voor M2M² interactie over een netwerk. Het W3C definieert een *web service* als

‘a standard means of interoperating between different software applications, running on a variety of platforms and/or frameworks’ [26].

Web services zijn volgens het W3C vaak niet meer dan een soort API’s die over een netwerk (vaak het internet) toegankelijk zijn. De uiteindelijke functionaliteit wordt in zulke gevallen uitgevoerd op een *remote system* dat de web services aanbiedt.

Een voorbeeld van een web service zou kunnen zijn, een ‘*postcode conversie*’ web service, deze service zou een *stadsnaam* kunnen omzetten in een specifieke *postcode*. Net zoals bij een service (in de echte wereld) is het niet noodzakelijk iets te weten over de uiteindelijke implementatie van de web service, op welk platform deze web service draait, in welke taal deze web service geschreven is, . . .

Ook dit lijkt op het eerste zicht simpel, toch bestaan er hier ook moeilijkheden. Hoe weet een applicatie bijvoorbeeld waar een andere web service te vinden is? Welke de input en output parameters van een bepaalde web service zijn? Waarvoor een web service juist dient? Dit antwoord wordt gegeven door een aantal formaten en technieken (zoals UDDI, WSDL, Semantic Web) die door computers leesbaar zijn en tevens vrij beschikbaar zijn gesteld op een specifieke weblocatie.

Een web service kan dus gezien worden als de basis waarop toekomstige gedistribueerde systemen ontwikkeld zouden kunnen worden. Hierdoor wordt het mogelijk voor individuen en organisaties om op een relatief eenvoudige manier, hun digitale bezit deels beschikbaar te maken naar het grote publiek. De toekomstmogelijkheden van web services zijn enorm, gaande van simpele functies (bv. een valuta-conversie functie) tot volledige bedrijfsprocessen (bv. boeking van vluchten). In de context van de implementatie in deze thesis is het dus interessant om rekening te houden met web services, vermits web services een steeds grotere component zouden kunnen worden in toekomstige digitale architecturen.

¹<http://www.awprofessional.com/articles/article.asp?p=99979&rl=1>

²Machine to Machine : datacommunicatie tussen machines

Een ander belangrijk punt is het feit dat *Physical Augmented Objects* (zie hoofdstuk 2) nood hebben aan interactie met andere objecten. Deze interactie zou kunnen worden voorzien door web services. In deze thesis gaan we verder in op dit punt en trachten we *Physical Augmented Objects* te koppelen aan een speciale soort web services nl. *Semantic Web services*. Dit type van web services bekijken we in een latere sectie (5.5).

5.3 Principe van een web service

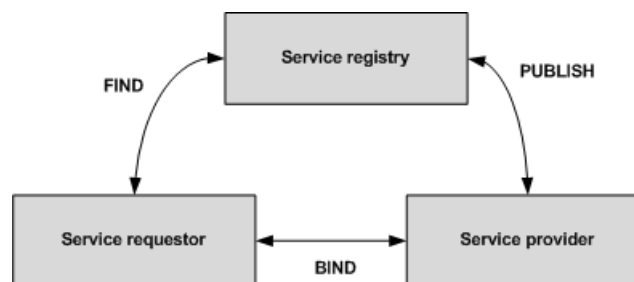
Volgens het W3C kunnen er in het algemeen in het concept van web services steeds drie verschillende rollen onderscheiden worden namelijk :

- Service provider - de aanbieder van een web service
- Service requestor - degene die de web service aanspreekt en gebruikt
- Service registry (Service Discovery Agent) - degene die de rol van ‘makelaar’ op zich neemt en vraag/aanbod bij elkaar brengt.

Eveneens zijn er drie verschillende activiteiten die plaatsvinden tussen deze rollen :

- publish (publiceren) - het publiceren van een web service door de Service provider
- find (vinden) - het vinden van een web service door de Service Requestor
- bind (koppelen) - de Service requestor maakt een koppeling met de Service provider, hierdoor kan hij de service gaan gebruiken.

Figuur 5.1³ toont de relaties tussen de rollen en activiteiten in het concept van web services.



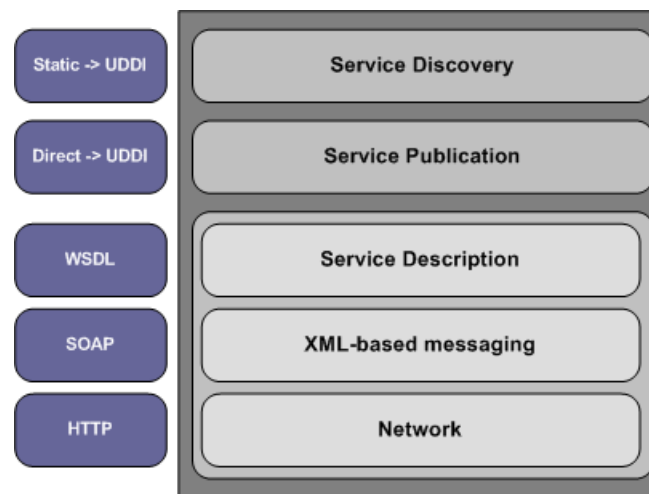
Figuur 5.1: Algemene concept van een webservice

5.4 Web services lagenmodel

Het web services lagenmodel dat in 2001 werd gepubliceerd[15], brengt enkele protocollen en interfaces bij elkaar die applicaties toestaan web services te vinden en te gebruiken. Figuur 5.2⁴ toont een gedeelte uit dit lagenmodel, deze stack is nog instabiel. Voor de elementen hoger in dit lagenmodel is er nog steeds meer onderzoek gaande dan de elementen die zich lager bevinden (vooral network, XML-based messaging). In de volgende secties overlopen we de verschillende delen uit dit lagenmodel.

³naar <http://www.novell.com/documentation/extend52/Docs/help/Director/books/UTImages/>

⁴naar <http://www.ibm.com/developerworks/kr/library/ws-featuddi/>



Figuur 5.2: Webservices lagenmodel

5.4.1 Network

In de onderste laag in het lagenmodel bevinden zich transportprotocollen die het transport van berichten tussen de applicaties regelen. Er wordt voornamelijk gebruik gemaakt van het standaard HTTP-protocol. De consensus omtrent deze laag is gelukkig vrij groot, hierdoor bevat het lagenmodel een stevige (standaard) basis.

5.4.2 XML-based Messaging - SOAP

Boven de netwerklaag bestaat er een laag die instaat voor de communicatie van web services met applicaties. Deze communicatie gebeurt via XML gebaseerde berichten. Er is gekozen voor een XML aanpak, vermits XML zich perfect leent tot het bieden van een standaard basis van gestructureerde berichten.

Simple Object Access Protocol (SOAP) is, zoals de naam al doet vermoeden, een protocol dat instaat voor de toegang tot objecten in een gedistribueerde (gedecentraliseerde) omgeving op een eenvoudige manier. SOAP definieert de structuur en de inhoud van XML-berichten die meestal met behulp van het HTTP transportprotocol worden verstuurd (TCP is eveneens mogelijk maar wordt weinig gebruikt). Daarnaast staat SOAP in voor het weergeven van functieaanvragen en antwoorden, en beschrijft het eveneens hoe applicaties hun data moeten coderen. SOAP biedt verschillende patronen om berichten uit te wisselen, doch de meest gebruikte (en bekende) is het *Remote Procedure Call* (RPC) patroon. Een precieze studie hierover valt buiten het bestek van deze thesis.

5.4.3 Service Description - WSDL

WSDL (*Web Services Description Language*) is een XML gebaseerde (en W3C gestandariseerde) taal die toelaat web services te beschrijven en eveneens aan te geven hoe deze web services toegankelijk zijn. WSDL specificeert de locatie van de web service en de mogelijke operaties (functies) die de web service aanbiedt. De operaties en inputs en outputs van een bepaalde web service worden dus abstract beschreven in WSDL. Hiernaast verschaft WSDL meer informatie over het concrete netwerk protocol dat gebruikt

kan worden. Dit alles laat toe om zogenaamde concrete ‘*web service endpoints*’ te definiëren. In principe is het mogelijk om WSDL te gebruiken in combinatie met meerdere netwerk protocols, doch in praktijk beperkt men zich tot SOAP, HTTP GET/POST of MIME.

Het algemene formaat van een WSDL-document bestaat steeds uit de volgende abstracte structuur:

```
<definitions>
<types>
  definitie van verschillende data types gebruikt door de web service
</types>

<message>
  definitie van een berichtformaat van de web service
</message>

<portType>
  definitie van een poort (operatie) van de web service
</portType>

<binding>
  definitie van communicatie protocols gebruikt door de web service
</binding>
</definitions>
```

Door het gebruik van deze simpele notatie maakt WSDL het mogelijk om web services eenvoudig te beschrijven en te gebruiken. Doch bevat WSDL ook een aantal nadelen nl.

- De data types die in WSDL beschreven kunnen worden, reiken niet verder dan de standaard XML schema types waaronder strings, integers, floats, ... Het is bijvoorbeeld niet mogelijk om te specificeren dat een bepaalde input van een web service, enkel een *persoonsnaam* kan zijn.
- WSDL biedt geen doordacht mechanisme om verschillende web services automatisch met elkaar te combineren, dit is gedeeltelijk te wijten aan de afwezigheid van specifieke input- en outputtypes.
- Met WSDL is het bijna onmogelijk om aan zogenaamde gedetailleerde ‘*service discovery*’ te doen. Ook dit is geheel te wijten aan het feit dat datatypes onvoldoende semantisch beschreven zijn.

Er is dus nood aan web services die zichzelf beter beschrijven, op een manier waardoor computers gemakkelijker (en meer automatisch) kunnen inzien waarvoor een bepaalde web service dienst doet. Met de komst van het *Semantisch web* zijn hiervoor bepaalde oplossingen ontstaan, onder de noemer *semantische web services*. Deze oplossingen zijn meestal specifieke web service ontologieën (net zoals bijvoorbeeld *FOAF* of *Dublin Core* ontologie is) die op een niveau hoger werken dan de WSDL-technologie. In de volgende sectie bekijken we enkele van deze oplossingen en gaan we tevens na welke voor- en nadelen zij hebben.

5.5 Semantische web services

5.5.1 Algemeen

In functie van deze thesis is het gebruik van *semantische web services* interessant vermits deze web services het mogelijk maken om op eigen initiatief na te gaan welke data interessant is en welke minder

interessant is. Wanneer we zulke services later koppelen aan het principe van *Augmented Physical Objects*, wordt het voor zulke objecten dus mogelijk andere bruikbare data op te sporen. In het hoofdstuk 6 gaan we hierop dieper in.

In de volgende sectie overlopen we een beperkt overzicht van bruikbare *semantische web service* technologieën.

5.5.2 OWL-S

OWL-S is een OWL-gebaseerde web service ontologie die *web service providers* een set van markup language constructies biedt voor het beschrijven van de eigenschappen en mogelijkheden van hun web services. OWL-S is gegroeid uit een initiatief dat bekend stond onder de naam ‘DAML-S’, wat op zijn beurt gebruik maakte van de voorloper van OWL namelijk ‘DAML+OIL’. De constructies die via OWL-S kunnen beschreven worden, hebben het voordeel dat ze ondubbelzinnig zijn en tevens interpreteerbaar voor computers. OWL-S maakt automatisatie taken mogelijk waaronder bijvoorbeeld ‘ontdekking van andere web services’, ‘automatische uitvoering van een web service’, ‘compositie van nieuwe web services op basis van andere web services (die in OWL-S zijn beschreven)’,...⁵

Het grote voordeel van OWL-S is het feit dat, in tegenstelling tot WSDL, OWL-S een beschrijving kan voorzien waarin de *input* en *output* types van een bepaalde web service veel meer semantiek kunnen bevatten. OWL-S biedt namelijk de mogelijkheid om semantische datatypes (datatypes uit RDF-gebaseerde ontologieën) te gebruiken. Hierdoor is het bijvoorbeeld mogelijk om voor een bepaalde service te specificeren dat deze enkel GPS-coördinaten als *input* ontvangt en een ISO-landafkorting als *output* verschaft.

In WSDL is het mogelijk om in één document meerdere functies (of ‘*ports*’) te beschrijven. OWL-S verplicht de *web service provider* om elke functie in een apart OWL-S document te beschrijven. Hierdoor ontstaan er OWL-S gebaseerde web services met exact één doelstelling waarbij op een ondubbelzinnige manier beschreven kan worden wat de *input* en *output* van deze enkelvoudige service is.

Dit alles maakt het mogelijk dat zulke web services automatisch kunnen worden ontdekt en uitgevoerd. Eveneens maakt deze aanpak het mogelijk om verschillende OWL-S gebaseerde web services te combineren met elkaar. Zoals in figuur 5.3 duidelijk wordt gemaakt, is het via OWL-S mogelijk web services te ‘ontdekken’, te combineren tot één nieuwe service (oftewel ‘*composite service*’) en vervolgens deze nieuwe services uit te voeren.

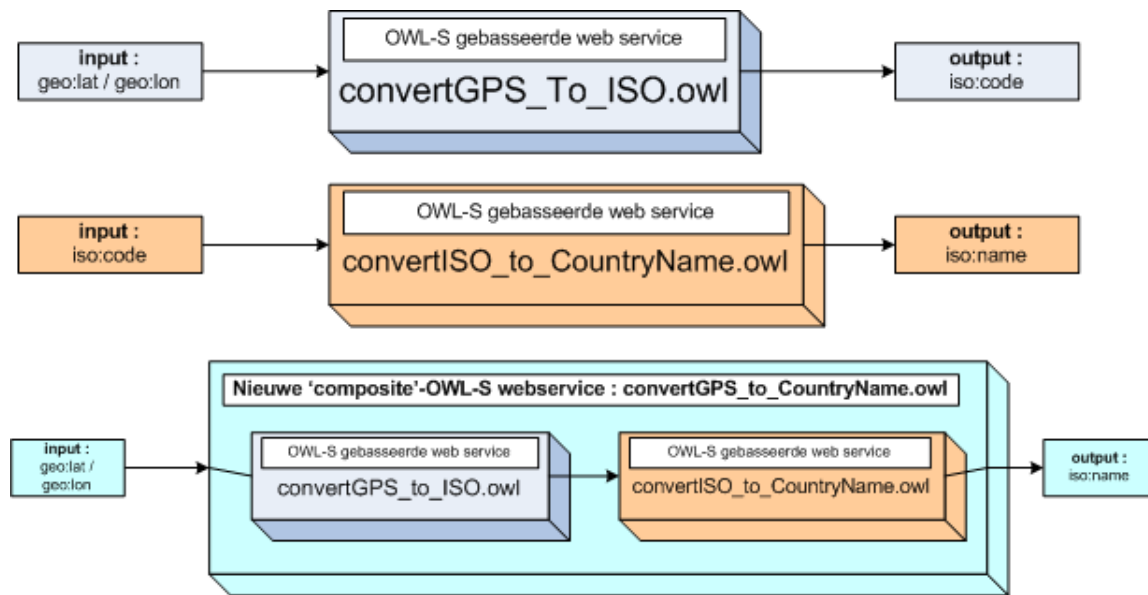
Er bestaan reeds een aantal tools die ondersteuning bieden voor het ontwikkelen van OWL-S beschrijvingen op basis van een bestaand WSDL-document. Doch vergt het helaas redelijk wat kennis omtrent meerdere technologieën (web services, semantische beschrijvingen, Semantic Web principes) om een OWL-S web service te beschrijven.

In de uiteindelijke implementaties in deze thesis (zie hoofdstuk 7) wordt er gebruik gemaakt van OWL-S met als onderliggende *grounding* WSDL services.

5.5.3 WSDL-S

In tegenstelling tot OWL-S, baseert WSDL-S zich niet op OWL. Het biedt wel de mogelijkheid om een bestaand WSDL-document uit te breiden met semantische datatypes. De bedoeling hiervan is een meer *lightweight* en praktisch gerichtere aanpak te bieden, die OWL-S niet helemaal kan leveren. WSDL-S is

⁵<http://www.daml.org/services/owl-s/1.0/>



Figuur 5.3: Voorbeeld van een OWL-S composite service

in principe in staat om dezelfde mogelijkheden te bieden dan OWL-S, en voegt hier zelfs enkele extra's aan toe.

Zo is het in WSDL-S mogelijk om verschillende soorten externe domeinmodellen (bv. UML, OWL, ...) te associëren met een *WSDL-S*-document. In OWL-S is het enkel mogelijk te verwijzen naar OWL-gebaseerde ontologieën. Een ander voordeel ten opzichte van OWL-S is het feit dat het updaten van WSDL-S web services relatief eenvoudig kan gebeuren.

Een nadeel van WSDL-S is het feit dat de uiteindelijke beschrijvingen XML-gebaseerd zijn. OWL-S-documenten zijn RDF-gebaseerd waardoor het gemakkelijker wordt om op een automatische manier deze documenten te *parsen* en te gebruiken. Hoewel de auteurs van WSDL-S voorstellen als een manier die alles vergemakkelijkt, maakt het de implementatie van bijvoorbeeld automatische ontdekking en compositie een stukje moeilijker.

5.5.4 SAWSDL

Een vrij bekend project in de wereld van het *Semantisch Web* getiteld *METEOR-S*⁶, kwam recent met SAWSDL⁷ (*Semantic Annotations for WSDL*) op de proppen. SAWSDL is sterk beïnvloed door WSDL-S (dat tevens voortbloeide uit het METEOR-S project), en biedt soortgelijke functionaliteit. Zo is het net zoals bij WSDL-S mogelijk om een WSDL-document uit te breiden met zogenaamde semantische annotaties. SAWSDL biedt twee basis types van annotaties nl. *model reference* en *schema mapping*. *Model reference* laat net zoals WSDL-S toe om verwijzingen te maken naar externe domeinmodellen. *Schema mapping* maakt het mogelijk om mappings te maken naar datatypes komende van externe XML-Schemas (ook dit is mogelijk in WSDL-S). Verder biedt SAWSDL kleinere extra features (die gebruik

⁶<http://lsdis.cs.uga.edu/projects/meteor-s/>

⁷<http://lsdis.cs.uga.edu/projects/meteor-s/SAWSDL/>

maken van WSDL 2.0 elementen) aan die buiten het bestek van deze thesis liggen. Ondanks het feit dat SAWSDL veel gelijkenissen vertoont met WSDL-S, is het in januari 2007 een zogenaamde *W3C Candidate Recommendation*⁸ geworden. Dit toont aan dat het onderzoek naar semantische annotaties voor web services zeker nog niet uitgestorven is.

5.5.5 WSDL 2.0

In 2006 kwam de opvolger van WSDL op het toneel, aanvankelijk wou men deze opvolger ‘WSDL 1.2’ dopen maar vermits de veranderingen ten opzichte van de vorige versie zo groot waren, kreeg het de titel ‘WSDL 2.0’. Ondanks WSDL veel kritiek heeft moeten slikken, heeft men met WSDL 2.0 weinig geleerd uit vroegere fouten. Ook WSDL 2.0 krijgt zware (zo niet nog grotere) kritieken⁹. Dit komt vooral door het feit dat WSDL 2.0 de zaken er niet op vergemakkelijkt (met bijvoorbeeld een gloednieuw component model, geen *backwards compatibility* met WSDL 1.x,...). WSDL 2.0 biedt wel de mogelijkheid om net zoals voorgaande initiatieven semantische datatypes te gebruiken (via nieuwe attributen zoals ‘*targetNamespace*’). WSDL 2.0 had een mooie gelegenheid kunnen zijn om de oorspronkelijke WSDL-beschrijving om te vormen tot een RDF-gebaseerd voorstel, helaas heeft men niet geopteerd voor deze verandering.

5.5.6 WSMO

Web Service Modeling Ontology (WSMO) is een ontologie die vergelijkbaar is met de vorige initiatieven. WSMO wordt ontwikkeld door de *ESSI WSMO working group*, deze werkgroep maakt deel uit van een Europees initiatief dat werkt rond *Semantic Web* systemen. Het voordeel hiervan is dat WSMO¹⁰ veel nauwer kan aansluiten bij andere initiatieven binnen de ESSI¹¹ groep zoals WSML¹² (*Web Service Modeling Language*) en WSMX¹³ (*Web Service Execution Environment*). De studie naar de precieze voor- en nadelen die WSMO ons biedt, vallen buiten het bestek van deze thesis.

5.6 Literatuur

Verschillende papers richten zich op specifieke onderdelen van *semantic web services* zoals bijvoorbeeld het annoteren van web services met semantische informatie [17] of de wijze waarop het matching proces gebeurt [34, 42]. *Semantische web services* zijn reeds enige tijd het onderwerp van veel onderzoek [28] maar worden pas vrij recentelijk gecombineerd met andere (nieuwe) technologie. Het onderzoek in de wereld van de *Semantic web services* is onderhevig aan veroudering, vermits de technologie waarmee *semantic web services* mee kunnen gebouwd worden nog veranderlijk is. In [45] stelt men DAML-S (de voorloper van OWL-S) voor, die als één van de eersten *semantische web services* beschrijven. Tevens werkt men in deze paper verder op DAML-S door de implementatie van *DAML-S/UDDI Matchmaker*. Deze applicatie breidt UDDI¹⁴ uit met semantische mogelijkheden (waarvoor DAML-S gebruikt wordt). Verder wordt er in deze paper *DAML-S Virtual Machine* voorgesteld, een applicatie die het mogelijk maakt interactie te hebben met zulke DAML-S web services. Hoewel dit onderzoek heeft bijgedragen aan de wijze waarop we tegen *Semantic web services* aankijken, blijkt de bruikbaarheid van deze tools enkele jaren na publicatie van weinig praktisch nut.

⁸<http://www.w3.org/2002/ws/sawSDL/>

⁹<http://webservices.xml.com/pub/a/ws/2004/11/17/salz.html>

¹⁰<http://www.wsmo.org>

¹¹European Semantic Systems Initiative - <http://www.essi-cluster.org/>

¹²<http://www.wsmo.org/wsml/>

¹³<http://www.wsmx.org>

¹⁴Universal Description, Discovery and Integration, XML beschrijving om webservices op het internet te presenteren en bekend te maken.

5.7 Toekomst

Wat brengt de toekomst voor web services technologie? Naarmate de technologie vordert in de wereld van web services, stijgt ook de kritiek. Het beschrijven van *semantic web services* is een tijdrovende en relatief moeilijke procedure die technische kennis vereist. Doordat de technologie rond *semantic web services* instabiel is, weigeren veel bedrijven te investeren in zulke services. Dit heeft als gevolg dat andere (bestaande) *semantic web services* minder nuttig worden. Op het moment dat deze thesis geschreven werd, waren er nauwelijks bestaande semantisch beschreven web services te vinden. Persoonlijk zie ik hier de komende jaren weinig verandering in komen, vermits het *Semantisch Web* zelf met een identiteitscrisis zit die nog jaren kan aanslepen. De toekomst van *semantic web services* is dus vaag en moeilijk te voorspellen. Het succes zal voor een groot deel gekoppeld zijn aan het succes van het *Semantisch Web*.

Deel III

Ontwikkeling

Hoofdstuk 6

Architectuur

Inhoudsopgave

6.1	Inleiding	50
6.2	Doelstelling	51
6.3	Architecturaal overzicht	51
6.3.1	Het creëren van een Semantisch Web	51
6.3.2	Het efficiënt gebruik van Semantische informatie	57
6.3.3	Veralgemeend concept	61
6.4	Mogelijke toepassingen	62
6.5	Aannames	62
6.6	Gerelateerd werk	65

6.1 Inleiding

Het volgend hoofdstuk kan als het kloppend hart van deze thesis beschouwd worden. Er wordt in dit hoofdstuk een generieke aanpak voorgesteld die gebruik maakt van principes die in de vorige hoofdstukken beschreven werden. Het uiteindelijke doel is het beschrijven van een ‘*architectuur*’ die de *fysiek-digitaal kloof* (beschreven in hoofdstuk 1.1) verkleint op een manier waarop eenvoudig verder gebouwd kan worden in de toekomst. De uiteindelijke architectuur kan dus als basispatroon gebruikt worden voor het oplossen van problemen die te wijten zijn aan de afstand die er tussen de digitale wereld en de reële wereld bestaat. Hierbij hechten we veel belang aan de factoren generaliseerbaarheid en uitbreidbaarheid.

De rode draad doorheen deze aanpak bestaat uit *Semantische Web* technologie (zie hoofdstuk 4), vermits deze technologie ons toestaat technologieën uit andere domeinen beter met elkaar te combineren. De kracht die het *Semantisch Web* hierbij levert, schuilt in de transparante wijze waarop data kan beschreven en uitgewisseld worden.

De verschillende onderdelen in het beschreven concept kunnen afzonderlijk in andere domeinen gebruikt worden, doch ligt de kracht van het geheel in de combinatie van de verschillende technieken en technologieën. In de volgende paragrafen doen we stapsgewijs de volledige architectuur uit de doeken, om daarna in het volgende hoofdstuk deze voorgestelde architectuur toe te passen in twee verschillende applicaties.

6.2 Doelstelling

Specifiek wordt er gestreefd naar een conceptuele architectuur die toestaat dat *augmented physical objects* (zie hoofdstuk 2) in staat zijn om op een transparante manier gebruikt te worden en op die manier de digitale en reële wereld korter bij mekaar brengt. Hierbij trachten we gebruik te maken van bestaande technologie (zoals web services en *Semantic Web* technologie) die we proberen te passen binnen één groot raamwerk. De uiteindelijke architectuur zou in staat moeten zijn om als basis gebruikt te worden in tal van toepassingsdomeinen. Een overzicht van mogelijke toepassingen die gebruik zouden kunnen maken van deze architectuur beschrijven we in paragraaf 6.4. De volledige architectuur wordt stapsgewijs uit de doeken gedaan. De verschillende stappen proberen eveneens oplossingen te bieden voor deelproblemen, ook deze problemen worden nauwgezet uitgelegd.

Kort samengevat vertrekken we van problemen die zich voordoen in het *Semantisch Web* en trachten we stapsgewijs oplossingen te bieden voor deze deelproblemen. Vervolgens combineren we de resulterende oplossingen in een meer generieke architectuur die tracht de *fysiek-digitaal kloof* te dichten op een transparante manier.

6.3 Architecturaal overzicht

6.3.1 Het creëren van een Semantisch Web

Het *Semantisch Web* heeft last van het ‘*kip en het ei*’-probleem¹. Zonder het bestaan van een *Semantisch Web*, is het nut van deze technologie vrij abstract. Doch moet het *Semantisch Web* eerst ‘opgebouwd’ worden voor het nuttig wordt. Met *opbouwen* bedoelen we het systematisch voorzien van semantische informatie aan documenten. Globaal gezien komt dit neer op het annoteren van metadata. Dit proces is echter nog steeds omslachtig en moeilijk. Hoewel er veel onderzoek wordt verricht in *Semantic Web* technologie, dringt het *Semantisch Web* moeilijk door bij het grote publiek. Andere gerelateerde technieken zoals bijvoorbeeld het principe van een *folksonomy*² daarentegen (waar eveneens redelijk wat theorie achter zit), kennen wel groot succes. Folksonomy werkt volgens het principe waar metadata wordt toegevoegd door middel van *tagging* door verschillende gebruikers (*Flickr*, *Del.icio.us*, . . . zijn voorbeelden van bekende sites waar deze techniek wordt toegepast). Het verschil in succes tussen beide gerelateerde technieken is naar mijn mening deels te wijten aan het feit dat het *Semantisch Web* vrij abstract wordt voorgesteld en weinig bruikbare tools kent.

Het ‘*kip en het ei*’-probleem dat het *Semantisch Web* kent, lost men naar mijn mening geleidelijk op door technologie af te schermen en te vergemakkelijken. Tim Berners-Lee³ verwoordt de opbouw van het Semantisch Web als volgt :

‘Bit by bit, link by link, the data becomes connected, interwoven. The exciting thing is serendipitous reuse of data: one person puts data up there for one thing, and another person uses it another way.’

Wanneer we kijken naar het huidige web, zien we dat het groot geworden is door de toegankelijkheid waarop het werd voorgesteld voor de massa. Hierbij kwam weinig tot geen hergebruik van informatie aan te pas. Door de aanwezigheid van simpele WYSIWYG HTML-editors, online website tools en dergelijke (die technologische details afschermen) ontstond de mogelijkheid om eenvoudig en snel een pril web uit te breiden. Naast het feit dat het *Semantisch Web* zich perfect leent tot het hergebruik van data, is er

¹<http://nl.wikipedia.org/wiki/Kip-en-ei-probleem>

²<http://en.wikipedia.org/wiki/Folksonomy>

³<http://www.w3.org/People/Berners-Lee/>

nood aan eenvoudige tools die dit hergebruik ondersteunen. Naar mijn mening moet het *Semantisch Web* meer leren uit de geschiedenis van het huidige web en hierop inspelen. Het *Semantisch Web* wordt een uitbreiding op het oorspronkelijke web, waarom zou het op vlak van *tool-support* dan niet verder werken op bestaande concepten?

Vanuit deze filosofie is er een kleine toepassing ontwikkeld die het hergebruik van semantische informatie tracht te ondersteunen en gebaseerd is op standaarden. Daarnaast probeert deze toepassing aan te tonen dat het *Semantisch Web* doordat het transparant is, verschillende losstaande componenten vrij eenvoudig kan combineren.

Het grondidee bestaat uit een extensie voor een bestaande WYSIWYG-editor, die toelaat semantische informatie toe te voegen aan HTML-pagina's. De linkerzijde van figuur 6.1 geeft een overzicht op de architectuur die is ontwikkeld om dit proces te ondersteunen. Deze architectuur wordt in volgende paragrafen stapsgewijs uit de doeken gedaan.

In een eerste stap zou RDF-data (dit is semantische informatie (in de vorm van bijvoorbeeld een FOAF-profiel of een URL naar een website met RDF informatie) als input kunnen dienen voor een *WYSIWYG-editor*. Deze (semantic) WYSIWYG-editor zou deze RDF-data op een eenvoudige (en visueel aantrekkelijke) manier beschikbaar kunnen maken. Hierdoor voldoen we aan twee belangrijke aandachtspunten nl. het hergebruik van data (*Berners-Lee's visie*) en de ondersteuning van het *Semantisch web* op een toegankelijke manier (nl. door middel van een WYSIWYG-editor).

Vervolgens is het van cruciaal belang dat de semantische informatie niet verloren gaat in een web van markup-languages, zoals dit het geval is in het traditionele web. Vermits we RDF-data herbruikbaar willen maken, is het dus belangrijk deze data op een gestandariseerde manier toe te voegen aan webpagina's. Zoals in paragraaf 4.5.5 al werd duidelijk gemaakt, biedt RDFa een gestandariseerde basis om herbruikbaarheid van semantische informatie voor webpagina's te verwezenlijken. De aangepaste WYSIWYG-editor kan de RDF-data automatisch omzetten naar (x)HTML code die is voorzien van semantische informatie (in de vorm van RDFa-code). In de implementatie in deze thesis is er gekozen om de WYSIWYG-editor enkel RDF-data afkomstig uit de FOAF-ontologie te laten verwerken. Op die manier kunnen we de layout meer specificeren en hierdoor eenvoudiger maken.

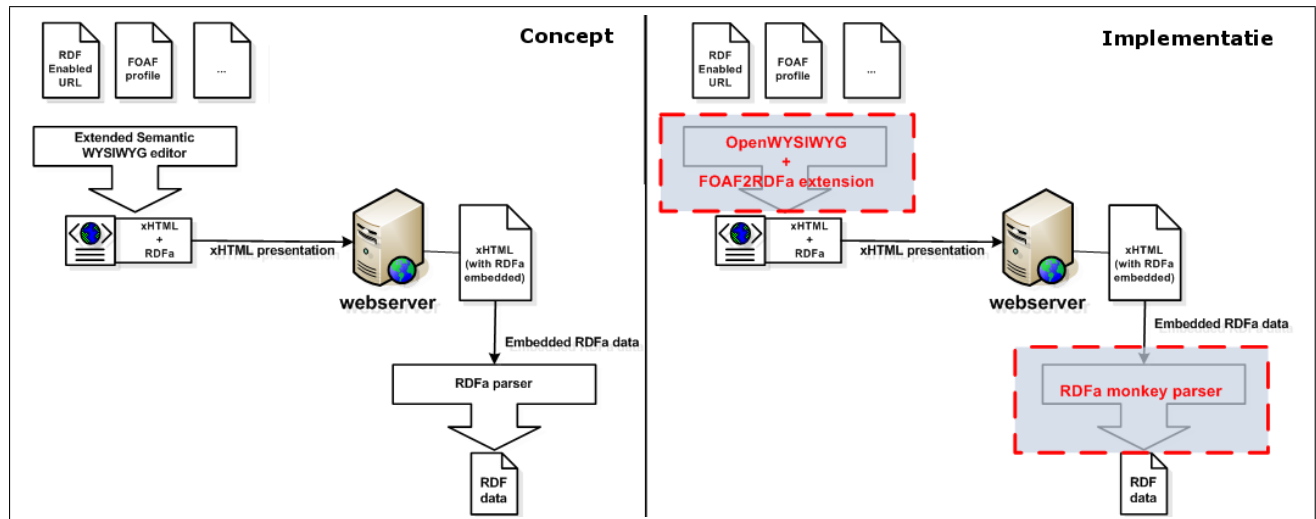
Om de uiteindelijke RDFa-code te kunnen hergebruiken elders, is er nood aan services die RDFa-code kunnen extraheren uit webpagina's en vervolgens deze code kunnen transformeren naar het originele RDF-formaat.

Het volledige proces bestaat dus uit :

1. Het eenvoudig toevoegen van semantische informatie aan normale (HTML gebaseerde) webpagina's door middel van een WYSIWYG-editor.
2. De mogelijkheid om semantische informatie in originele vorm te extraheren uit webpagina's die semantische informatie bevatten.

Dit proces ondersteunt naar mijn mening de manier waarop een *Semantisch Web* eenvoudig kan uitbreiden zonder dat hiervoor specifieke technologische kennis van de gebruikers nodig is. In deze thesis zijn er enkele implementaties ontwikkeld die trachten dit proces te ondersteunen. De rechterkant van figuur 6.1 toont een overzicht van de overlappingen met de uiteindelijke implementatie die dit proces ondersteunt. In de volgende paragrafen beschrijven we deze implementaties in meer detail.

We houden eveneens in het achterhoofd dat het *Semantisch Web* een cruciale pion is in de uiteindelijke architectuur om problemen op te lossen die te wijten zijn aan de *fysiek-digitaal kloof*. Het is dus in het belang van de uiteindelijke architectuur dat het creëren van een *Semantisch Web* zo goed mogelijk ondersteund wordt.



Figuur 6.1: Conceptuele architectuur - het creëren van een Semantisch Web

OpenWYSIWYG met 'FOAF2RDFa'-extensie

Als WYSIWYG-editor werd er geopteerd voor de xHTML-editor OpenWYSIWYG⁴ omwille van meerdere redenen. Een eerste en belangrijke reden is het feit dat OpenWYSIWYG, zoals de naam doet vermoeden, open source is en uitgegeven is onder LGPL-licentie⁵. Een tweede reden is het feit dat OpenWYSIWYG een online editor is, dit heeft als voordeel dat deze editor makkelijk elders te gebruiken is. Daarnaast zullen naar mijn mening toepassingen in de toekomst als maar meer op het web gericht worden. Vermits het *Semantisch Web* een uitbreiding wordt op het traditionele web, is het dus niet ondenkbaar dat ook de uiteindelijke creatie van het *Semantisch Web* voor een groot deel zal plaatsvinden op het traditionele web. Een online editor (zoals OpenWYSIWYG) stelt op dat vlak dus eveneens een voorbeeld, dit voorbeeld wordt stilaan ingeburgerd op vele websites⁶. Een derde reden is het feit dat OpenWYSIWYG volledig ontwikkeld is met behulp van Javascript in combinatie met xHTML en CSS, waardoor deze editor 'licht' is en eenvoudig kan gebruikt op eender welke webhosting. Als laatste reden blijkt OpenWYSIWYG volledig compatibel te zijn met de meest populaire browsers, waardoor het toegankelijk is voor een zo groot mogelijk publiek.

De uiteindelijke implementatie bestaat uit een extra feature voor OpenWYSIWYG die beschikbaar is vanuit de menubalk van de editor. Deze bevat een extra knop (zie 1 uit figuur 6.2) die resulteert in een popup-venster (zie 2 uit figuur 6.2) waar de gebruiker een url naar een FOAF-profiel⁷ kan geven. Zulke FOAF-profielen (die dus semantische informatie bevatten over een bepaalde persoon, zijn hobbies en zijn vrienden) kunnen eenvoudig aangemaakt worden op verschillende sites⁸.

⁴<http://www.openwebware.com/products/openwysiwyg>

⁵<http://www.opensource.org/licenses/lgpl-license.php>

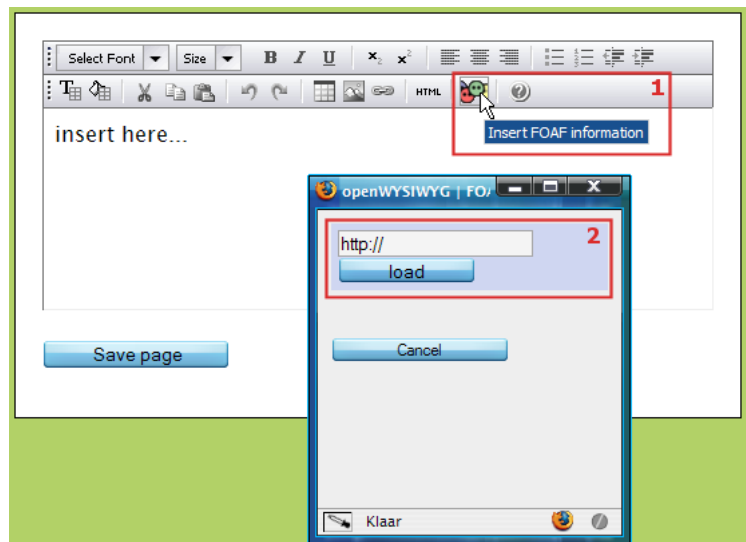
⁶bv. <http://pages.google.com>

⁷Friend of a friend (FOAF)

⁸<http://www.ldodds.com/foaf/foaf-a-matic.html>

Vervolgens zal met één druk op de knop een achterliggende service het ingegeven FOAF-profiel *parsen* en in een specifiek formaat (aangepast aan de editor) terugsturen naar de OpenWYSIWYG-editor. Deze achterliggende service is geïmplementeerd met behulp van PHP en de RDF-parser van het RAP (RDFapi-php) framework⁹. De communicatie tussen de PHP-service en de editor gebeurt via de AJAX-functionaliteit van de Dojo-bibliotheek¹⁰. Het uiteindelijk resultaat bestaat uit xHTML-code in combinatie met Javascript functieaanroepen. Dit geheel wordt getoond in het popup-venster van de editor (zie figuur 6.3). Eenmaal de verwerkte RDF-data getoond wordt, kan de gebruiker de verschillende delen informatie uit het FOAF-profiel aanklikken waarna deze worden toegevoegd aan de editor. In dit proces wordt er naast de opmaak, ook RDFa code gecreëerd en toegevoegd. Figuur 6.4 geeft een gedetailleerd overzicht op de achterliggende architectuur van dit volledige (deel)concept.

Deze extensie spitst zich toe op de specifieke FOAF-ontologie, doch is dit principe vrij eenvoudige generaliseerbaar naar andere ontologieën. Er is gekozen voor een implementatie die zich toespitst op de FOAF-ontologie vermits deze ontologie reeds vrij populair is en het principe hierdoor duidelijker wordt.



Figuur 6.2: Invoeren van RDF-informatie in OpenWYSIWYG met FOAF2RDFa extensie

RDF/A parser

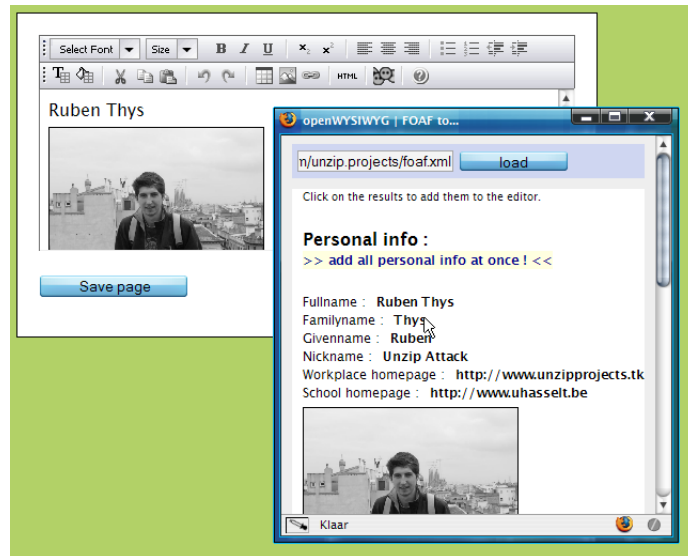
Naar aanleiding van de literatuurstudie omtrent het *Semantisch Web* (zie hoofdstuk 4), werd duidelijk dat de nood naar een parser voor documenten die RDF/A (zie sectie 4.5.5) bevatten hoog was. Vermits het schrijven van een parser het inzicht in de taal die *geparsed* wordt vereist, was dit tevens een goede manier om het *Semantisch Web* beter te begrijpen via een praktische aanpakswijze.

Het doel van een zogenaamde RDF/A parser in het hierboven beschreven concept is dat HTML-documenten die semantische informatie in de vorm van RDF/A¹¹ data bevatten (en die geleverd worden via de OpenWYSIWYG editor), kunnen omgezet worden naar *standaard* RDF-data (die veel bruikbaar is voor

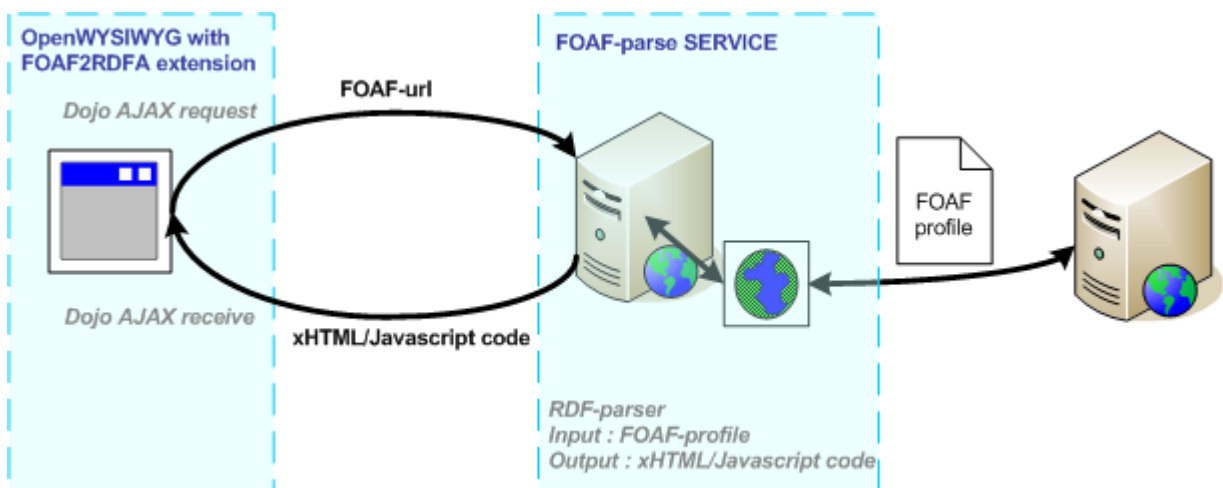
⁹<http://sites.wiwiss.fu-berlin.de/suhl/bizer/rdfapi/>

¹⁰<http://dojotoolkit.org/>

¹¹<http://en.wikipedia.org/wiki/RDFa>



Figuur 6.3: Resultaat na FOAF-parsing in OpenWYSIWYG met FOAF2RDFa extensie

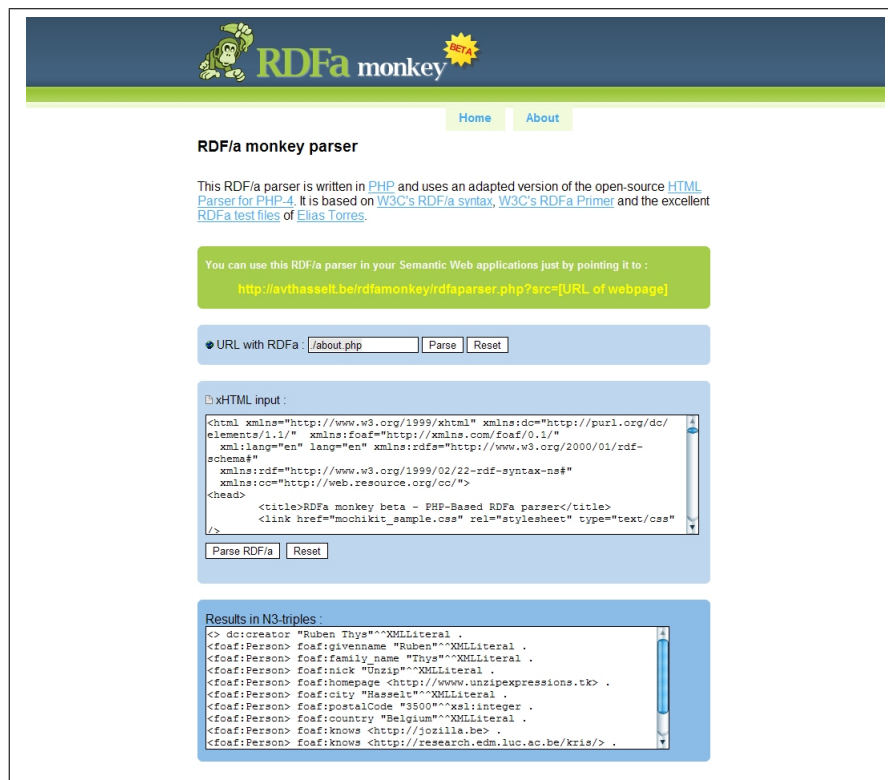


Figuur 6.4: Overzicht architectuur : OpenWYSIWYG met FOAF2RDFa extensie

Semantic Web doeleinden). Er is gekozen voor de N3-notatie (zie sectie 4.10) als voorstellingswijze voor de uiteindelijke RDF-data. Via bestaande services en bibliotheken bestaat nog steeds de mogelijkheid om de uiteindelijke N3-notatie om te zetten naar RDF/XML-notatie. Figuur 6.1 toont de plaats van deze parser in het concept rond *het creëren van een Semantisch Web*.

De uiteindelijke implementatie is gebeurd in PHP aangezien er nog geen parser in deze taal voor handen was (sterker nog, bij het ontwikkelen van deze parser waren er nog maar twee bestaande implementaties bekend). De parser is gebaseerd op de *W3C RDF/A syntax*¹², *W3C's RDF/A Primer*¹³ en getest met voorbeelden afkomstig van *Elias Torres*¹⁴. De parser kan via een querystring worden aangesproken en wordt hierdoor bruikbaar als externe web service in *Semantic Web* projecten. Tevens is er gebruik gemaakt van de licht aangepaste versie van de *HTML Parser for PHP-4*.

De uiteindelijke parser is verwerkt in een simpele en gebruiksvriendelijke website (zie figuur 6.3.1) die ontwikkeld is met PHP en enkele Javascript bibliotheken. Bij het schrijven van deze parser zijn er enkele tegenstrijdigheden in W3C's RDF/A syntax opgemerkt, die gerapporteerd zijn en in de toekomst verholpen zullen worden. De zogenaamde *RDFa monkey parser* is beschikbaar als open source onder de *GNU General Public License*.



Figuur 6.5: Screenshot van de RDF/A monkey website

¹²<http://www.w3.org/2001/sw/BestPractices/HTML/2005-rdfa-syntax#id69901>

¹³<http://www.w3.org/TR/xhtml-rdfa-primer/>

¹⁴<http://torrez.us/>

6.3.2 Het efficiënt gebruik van Semantische informatie

Een ander probleem dat zich voordoet in zowel de fysieke als de digitale wereld is het feit dat digitale informatie moeilijk beschikbaar is in beide werelden en er veel handmatige acties bij te pas komen om dit te realiseren. Dit probleem werd in sectie 1.1 uitgebreid uit de doeken gedaan. Het moeilijk bruikbaar maken van semantische informatie is een gerelateerd probleem dat deel uitmaakt van het hierboven aangehaalde probleem. Eenmaal documenten voorzien zijn van semantische informatie is het vaak moeilijk deze informatie op een nuttige en efficiënte manier te gebruiken voor allerlei doeleinden.

Ook hiervoor trachten we naar oplossingen te zoeken op basis van het eerder onderzoek in deze thesis. Een eerste opmerking die hierbij gemaakt kan worden is dat, indien we de digitale wereld en fysieke wereld korter bij mekaar willen brengen, we nood hebben aan technologie die ons toestaat deze overbrugging te maken. *Radio Frequency Identification* (zoals besproken in hoofdstuk 3) laat toe dat fysieke objecten digitaal kunnen worden geïdentificeerd, waardoor fysieke objecten een *digitale tegenhanger* (zie hoofdstuk 2) krijgen. Via deze digitale tegenhanger wordt het ook mogelijk te verwijzen naar zulke objecten in de digitale wereld. *Semacode tags* kunnen evenwel dienst doen als identificatietechnologie die dit principe ondersteunt. Om de architectuur beter begrijpbaar te maken, gebruiken we in de verdere uitleg steeds een concrete identificatie-technologie, nl. RFID.

Naast het probleem van een transparante overbrugging, is het vaak moeilijk om in de digitale wereld applicaties aan te sturen. Enerzijds is dit te wijten aan het feit dat deze applicaties via *zelf in elkaar geknutselde* protocollen toegankelijk zijn. Anderzijds worden applicaties op het internet toegankelijker gemaakt door het gebruik van web service beschrijvingen. Doch gaat het gebruik van een web service meestal eerst gepaard met het uitzoeken van de input- en outputparameters die noodzakelijk zijn om de web service te gebruiken. Wanneer we de kloof tussen de fysieke en digitale wereld willen verkleinen, is het van cruciaal belang deze web services nog (meer automatisch) toegankelijk te maken.

Web services die semantische input- en outputtypes ondersteunen (en dus zelf kunnen uitmaken of data geschikt is als inputparameter, zie hoofdstuk 5), kunnen gezien worden als een eerste stap naar een meer gegeneraliseerde en bruikbare oplossing.

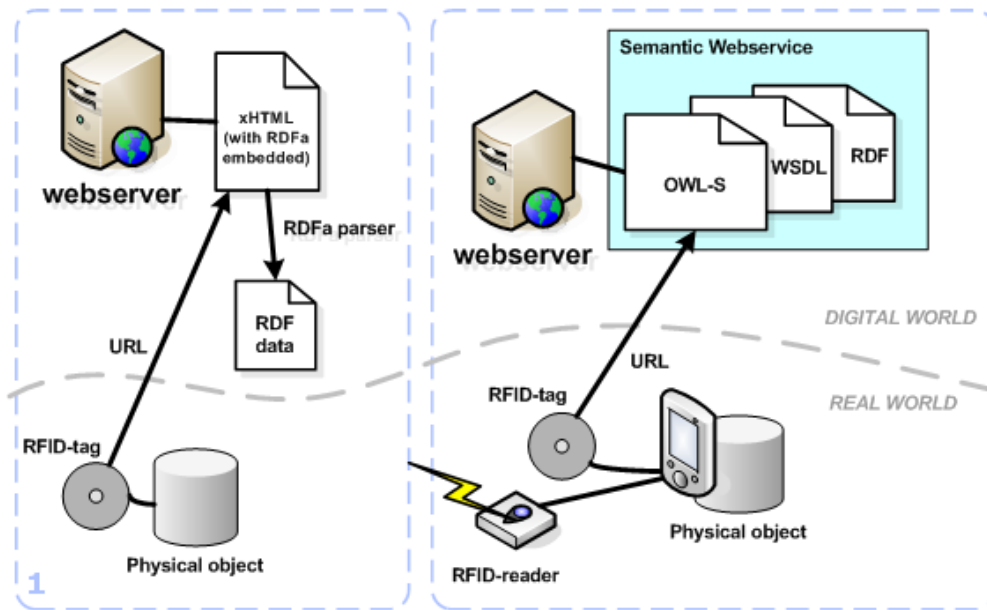
Waar leidt dit alles ons? We trachten een architectuur op te bouwen die het gebruik van semantische informatie in onze fysieke en digitale wereld vergemakkelijkt. Op basis van de bevindingen uit de vorige secties, werd er een architectuur (figuur 6.6) ontwikkeld die dit proces ondersteunt.

Wanneer het *Semantisch Web* eenmaal voldoende semantische informatie (RDF-beschrijvingen) bevat (volgend op sectie 6.3.1) wordt het vrij eenvoudig om naar deze (RDF)data te verwijzen. Dit kan namelijk door het gebruik van URL's gebeuren, vermits het *Semantisch Web* een uitbreiding is op het traditionele web. Door deze URL's op te slaan op RFID-tags (of een verwijzing te maken naar unieke URL's voor elke RFID tag, bv. 'http://abc.com/rfid0012'), ontstaan er fysieke objecten die verwijzen naar digitale semantische informatie (linkerzijde figuur 6.6). Deze fysieke objecten zijn op deze manier in zowel de digitale wereld (door middel van semantische informatie) als de fysieke wereld (door de vorm, de eigenschappen, . . . van het fysieke object) gedefinieerd.

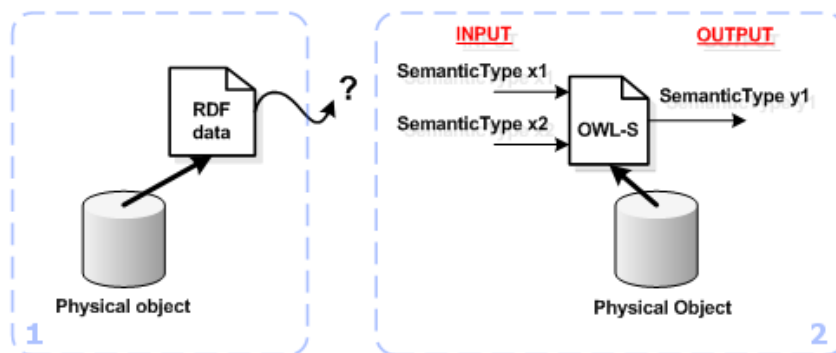
Verder bevat de architectuur in figuur 6.6 (rechterzijde) fysieke objecten in de reële wereld die enerzijds gekoppeld zijn met een RFID-reader en anderzijds een RFID-tag bevatten. Deze RFID-tag op zijn beurt verwijst via een URL naar een OWL-S¹⁵ gebaseerde web service. Op die manier ontstaat er een *augmented physical object* dat in staat is om in de digitale wereld (semantische RDF) data te 'ontvangen'

¹⁵OWL-S maakt het mogelijk om web services semantisch op een gestandariseerde manier te beschrijven. <http://www.daml.org/services/owl-s/1.0/>

en te verwerken. De interactie tussen de beide fysieke objecten (de linker en rechterkant van figuur 6.6) in de digitale wereld, kan in de fysieke wereld worden gestart. Dit kan bijvoorbeeld door de twee *augmented objects* bij elkaar te brengen. Van zodra bijvoorbeeld de RFID-lezer opmerkt dat er een ander object in de buurt is, zal een OWL-S service worden uitgevoerd met de data die afhankelijk is van het object (dat in de buurt komt). Merk op dat deze fysieke objecten nood hebben aan ‘processor’ kracht en een (draadloze) internetverbinding, hiervoor kan er bijvoorbeeld een smartphone of PDA in het fysiek object geïntegreerd zijn.



Figuur 6.6: Overzicht Augmented Physical Objects concept-architectuur



Figuur 6.7: Vereenvoudigde voorstelling van Augmented Physical Objects

Nu we beide delen uit het concept in figuur 6.6 hebben beschreven, kunnen we deze architectuur vereenvoudigd voorstellen (zie figuur 6.7). Aan de linkerzijde bestaat er een fysiek object dat op een bepaalde

manier is gekoppeld aan semantische RDF data, de rechterzijde bevat een fysiek object dat in staat is om semantische data te *ontvangen* en te verwerken.

Wat brengt dit ons op? Zoals misschien duidelijk is geworden, hebben we twee *physical augmented objects*[30] gedefinieerd, die zowel in de fysieke wereld (via identificatietechnologie, in dit geval RFID) als in de digitale wereld (via *Semantic Web* technologie) met elkaar kunnen interageren. Het is van groot belang hierbij op te merken dat dit op een transparante manier gebeurt en tevens is gebaseerd op standaarden met een zeer beperkte tussenkomst van ‘*custom-made*’ oplossingen. Dit alles heeft tot gevolg dat sommige fysieke objecten de mogelijkheid krijgen om een specifieke selectie (op basis van semantische informatie) te maken uit andere fysieke objecten en we tevens andere fysieke objecten als ‘input’ kunnen gebruiken om een bepaalde dienst of service aan te bieden.

Op het eerste zicht lijkt dit nogal vaag, daarom verduidelijken we deze situatie door middel van een kort voorbeeld dat volgens dit concept werkt :

Een identiteitskaart kan bijvoorbeeld een RFID-tag bevatten die een verwijzing naar een FOAF-profiel bevat. De identiteitskaart kunnen we in dit geval vergelijken met een fysiek object dat gekoppeld is aan bepaalde semantische informatie (overeenkomstig met de linkerzijde van figuur 6.7). Een infokiosk in een grote stad zou een RFID-reader kunnen bevatten. Deze infokiosk kan informatie tonen op basis van de interesses van personen. Deze infokiosk is dus geïnteresseerd in specifieke eigenschappen van personen (oftewel FOAF informatie). Hiervoor zou de infokiosk een verwijzing naar een web service kunnen bevatten die enkel semantische informatie over specifieke eigenschappen van personen als input ontvangt (overeenkomstig met het rechter gedeelte uit figuur 6.7).

Wanneer een persoon bijgevolg zijn identiteitskaart kort bij de RFID-reader van de infokiosk houdt, zal deze infokiosk meteen via een web service gebruik kunnen maken van de persoonlijke informatie van de persoon. Dit proces komt natuurlijk over, vermits personen kunnen werken met fysieke objecten en tegelijkertijd informatie in de digitale wereld automatisch wordt uitgewisseld.

In dit voorbeeld is er een overbrugging gebeurt tussen de fysieke en digitale wereld, zonder dat de gebruiker zich iets moet aantrekken van technische achterliggende details. De volgende paragrafen gaan dieper in op de deze achterliggende details en de realisatie van deze architectuur.

OWL-S wrapper voor PEAR::SOAP

Na de ontwikkeling van deze architectuur werd duidelijk dat er nood was aan een bibliotheek die OWL-S web services kan uitvoeren op basis van semantische input types (en bijhorende data) (zie sectie 5.5.2). Aanvankelijk werd de OWL-S API van Mindswap¹⁶ van korter bij bekeken (waarvoor men ook wil opteren in [41]). Doch werd al snel duidelijk dat deze API een aantal tekortkomingen heeft :

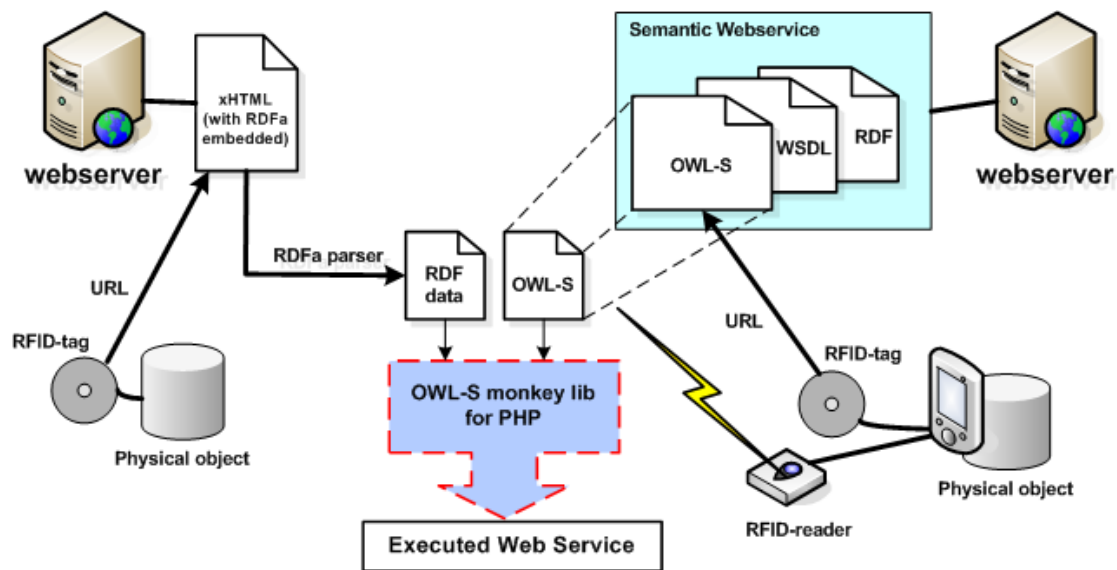
- Naast het feit dat deze API veel features bevat (zoals bijvoorbeeld het omzetten van WSDL-bestanden naar OWL-S formaat, de creatie van *composite services*,...), is de implementatie om OWL-S web services aan te spreken vrij simplistisch. Deze bibliotheek verplicht de ontwikkelaar om OWL-S services aan te spreken op een manier die weinig voordelen biedt ten opzichte van WSDL. Hiermee wordt bedoeld dat het noodzakelijk is dat ontwikkelaar op voorhand kennis heeft van de types van de input-parameters en de volgorde waarin deze parameter moeten doorgegeven worden. Hierdoor wordt de kracht van OWL-S volledig teniet gedaan en blijven er weinig redenen over om OWL-S te gebruiken.

¹⁶Maryland Information and Network Dynamics Lab Semantic Web Agents Project - <http://www.mindswap.org/2004/owl-s/index.shtml>

- De Mindswap OWL-S bibliotheek is moeilijk te installeren en draait op een zwaar (offline) Java-framework. Web services moeten licht en makkelijk te ontwikkelen zijn, het grote succes van REST¹⁷ web services zet deze uitspraak enkel kracht bij.

Op basis van deze tekortkomingen werd een nieuwe oplossing ontwikkeld. Als grondlaag werd er gekozen voor een bestaande bibliotheek genaamd PEAR¹⁸. Deze bibliotheek bevat een *SOAP Client/Server* package voor het PHP-framework, tevens is deze bibliotheek licht, makkelijk installeerbaar en webgebaseerd. Deze *PEAR::SOAP*-package bevat daarenboven een extensie¹⁹ die het mogelijk maakt WSDL-services uit te voeren. Hierdoor biedt deze bibliotheek de ideale basis om een verbeterde OWL-S implementatie te implementeren.

De uiteindelijke OWL-S *wrapper* die bovenop de *PEAR::SOAP package with WSDL extension* is geschreven, is eveneens in PHP uitgewerkt. In de uiteindelijke uitvoering zal de RDF-parser van het *RAP framework* eerst de OWL-S beschrijving parsen. OWL-S onderscheidt zich hier reeds van andere aanpakken (zoals WSDL-S, zie sectie 5.5.2) doordat het gebaseerd is op RDF en daardoor makkelijk interpreteerbaar is. Vervolgens worden de input- en outputparameters van de OWL-S beschrijving *gemapt* op deze van de onderliggende WSDL-beschrijving. Eveneens wordt de OWL-S service *gemapt* op deze van de *grounding* WSDL-service. De OWL-S wrapper bevat daarnaast enkele functies die het mogelijk maken om uit een gedeelte RDF-data (of een URL met RDF-data) de bruikbare inputparameters te extraheren (op basis van de OWL-S beschrijving). Op die manier ondersteunt deze OWL-S *wrapper* (getiteld ‘*OWL-S monkey lib*’) het concept dat in figuur 6.6 beschreven werd. Figuur 6.8 tracht de plaats van deze implementatie in het geheel concept weer te geven.



Figuur 6.8: Gedetailleerde implementatie van Augmented Physical Objects concept-architectuur

Deze OWL-S wrapper ondersteunt op dit moment de W3C-standaard van OWL-S 1.0 en 1.1. Onder de

¹⁷<http://www.xfront.com/REST-Web-Services.html>

¹⁸<http://pear.php.net/package/SOAP>

¹⁹http://www.phppatterns.com/docs/develop/extending_pear_soap_wsd

naam *OWL-S monkey lib* is deze wrapper vrij beschikbaar als open source onder de *GNU General Public License*. Figuur 6.1 toont de manier waarop de OWL-S wrapper gebruikt kan worden, figuur 6.2 toont (in analoge pseudo-code) het verschil in aanpak met de OWL-S lib van Mindswap.

Listing 6.1: OWL-S monkey lib - code

```
// Creeër een OWL-S 1.1 object
$webservice = new OWLS('http://www.abc.com/doSomething.owl', '1.1');
// RDF link (of data)
$rdf_url = 'http://www.abc.com/myRDFFile.rdf';

// check of er bruikbare inputtypes bestaan in de RDF-data
if($webservice->validRDFParams($rdf_url))
{
    // voer de OWL-S webservice uit
    echo $webservice->executeWithRDFParams();
}
```

Listing 6.2: Mindswap OWL-S lib - analoge pseudo-code

```
// Creeër een OWL-S 1.1 object
$webservice = new OWLS('http://www.abc.com/doSomething.owl', '1.1');

// array van input parameters - kennis van de inputtypes is nodig !
$p = array();
array_push($p, 'value1');
array_push($p, 'value2');

// uitvoeren van OWL-S web service met de gegeven inputparameters
echo $webservice->executeFunction($p);
```

6.3.3 Veralgemeend concept

In sectie 6.3.1 bekeken we hoe het *creëren van een Semantisch Web* vergemakkelijkt kan worden. Doch is het vaak een probleem om nieuwe domeinconcepten binnen het *Semantisch Web* te passen. Hoewel een persoonlijk profiel bijvoorbeeld reeds wordt beschreven in de vorm van een FOAF-file, bestaan er niet voor alle domeinen zulke taxonomieën. Daarom is het mogelijk om zulke nieuwe beschrijvingen (ook wel ontologieën, zie hoofdstuk 4) te creëren. Eens een ontologie is beschreven, kan deze vrij gebruikt worden in alle componenten van het *Semantisch Web*.

Nu we enkele oplossingen hebben geboden voor problemen die zich afspelen binnen de *Semantic Web* wereld, kunnen we (onder de noemer $1+1=3$) deze oplossingen combineren tot een algemene architectuur die een oplossing tracht te bieden om de digitale en reële wereld korter bij elkaar te brengen. Het generiek concept dat wordt voorgesteld in deze thesis combineert de concepten in sectie 6.3.1 en 6.3.2 met de manier waarop een (nieuw) domein gemakkelijk kan worden beheerd (in een *Semantic Web framework*). Figuur 6.9 geeft een overzicht op de volledige architectuur die toelaat dat informatie de overbrugging kan maken van fysieke naar digitale wereld (en andersom) op een gefundeerde manier. Figuur 6.10 geeft een alternatieve presentatie van deze architectuur. Het idee en de uitwerking van deze architectuur werd

grotendeels aangebracht in vorige secties (6.3.1 en 6.3.2).

In volgende hoofdstukken gebruiken we deze concept-architectuur als basis voor twee verschillende applicaties en trachten de generaliseerbaarheid van de architectuur aan te tonen.

6.4 Mogelijke toepassingen

Deze concept-architectuur kan gebruikt worden in verschillende toepassingsdomeinen. In het volgende hoofdstuk, passen we het volledige concept toe op twee specifieke domeinen, nl. een winkelomgeving en het domein van de (geschreven) media (kranten, tijdschriften, nieuws, ...). In deze sectie overlopen we enkele mogelijke toepassingen die aan de hand van dit concept eveneens zouden uitgewerkt kunnen worden.

- Foto slideshow - Stel dat foto's een verwijzing krijgen naar dezelfde foto die digitaal (semantisch beschreven) te vinden is op een website. Onder deze aanname zou een digitaal foto-frame (met ingebouwde RFID-reader) eenvoudig kunnen aangevuld worden met foto's. De digitale fotoframe spreekt een web service aan die random foto's uit een bepaalde queue teruggeeft. Door foto's kort bij de fotoframe te houden, kan dit frame op een transparante manier 'zien' dat er foto's in de buurt liggen en vervolgens de semantische informatie van deze foto's toevoegen aan de slideshow-queue. Deze toevoeging gebeurt in dat geval via de OWL-S web service.

Figuur 6.11 geeft een overzicht op dit toepassingsdomein. Op het eerste zicht lijkt dit voorbeeld nogal artificieel, toch is het enige 'pijnpoint' het feit dat RFID-tags niet ingeburgerd zijn en het moeilijk is informatie toe te voegen aan foto's. Door de explosie van digitale cameras, digitale foto-services en de komst van draadloos internet, zou het in de toekomst mogelijk moeten zijn om een RFID-tag met de nodige semantische informatie bij het afdrukken van de foto's automatisch te laten toevoegen.

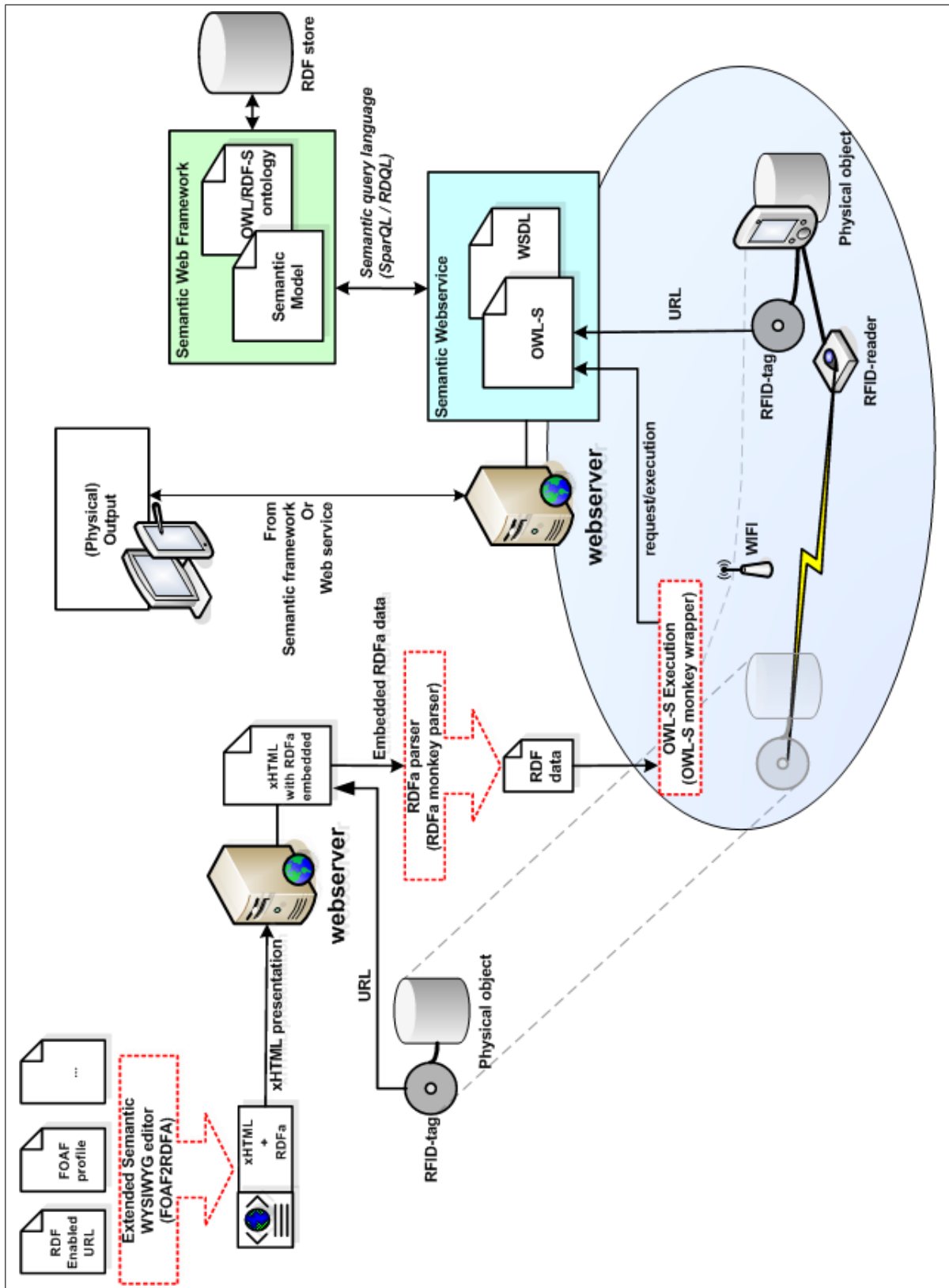
- Persoonlijke gegevens (interessegebied, vrienden, ...) eenvoudig gebruiken - een identiteitskaart (met RFID-chip) zou bijvoorbeeld uitgebreid kunnen worden met een verwijzing naar een bepaald FOAF-profiel (indien men dit wenst, met privacy issues wordt voorlopig geen rekening gehouden). Musea zouden hier kunnen op inspelen doordat de identiteitskaart van de bezoeker dienst zou kunnen doen als fysieke en digitale representatie van zichzelf doorheen het museum. Objecten in het museum zouden op hun beurt dienst kunnen doen als web services die bijvoorbeeld extra informatie kunnen bieden afhankelijk van interessegebied. Dit kort voorbeeld toont eenvoudig aan dat fysieke objecten uit andere domeinen, transparant bruikbaar worden in geheel andere toepassingsgebieden (musea) zonder hier ad hoc oplossingen bij te kijken komen.

6.5 Aannames

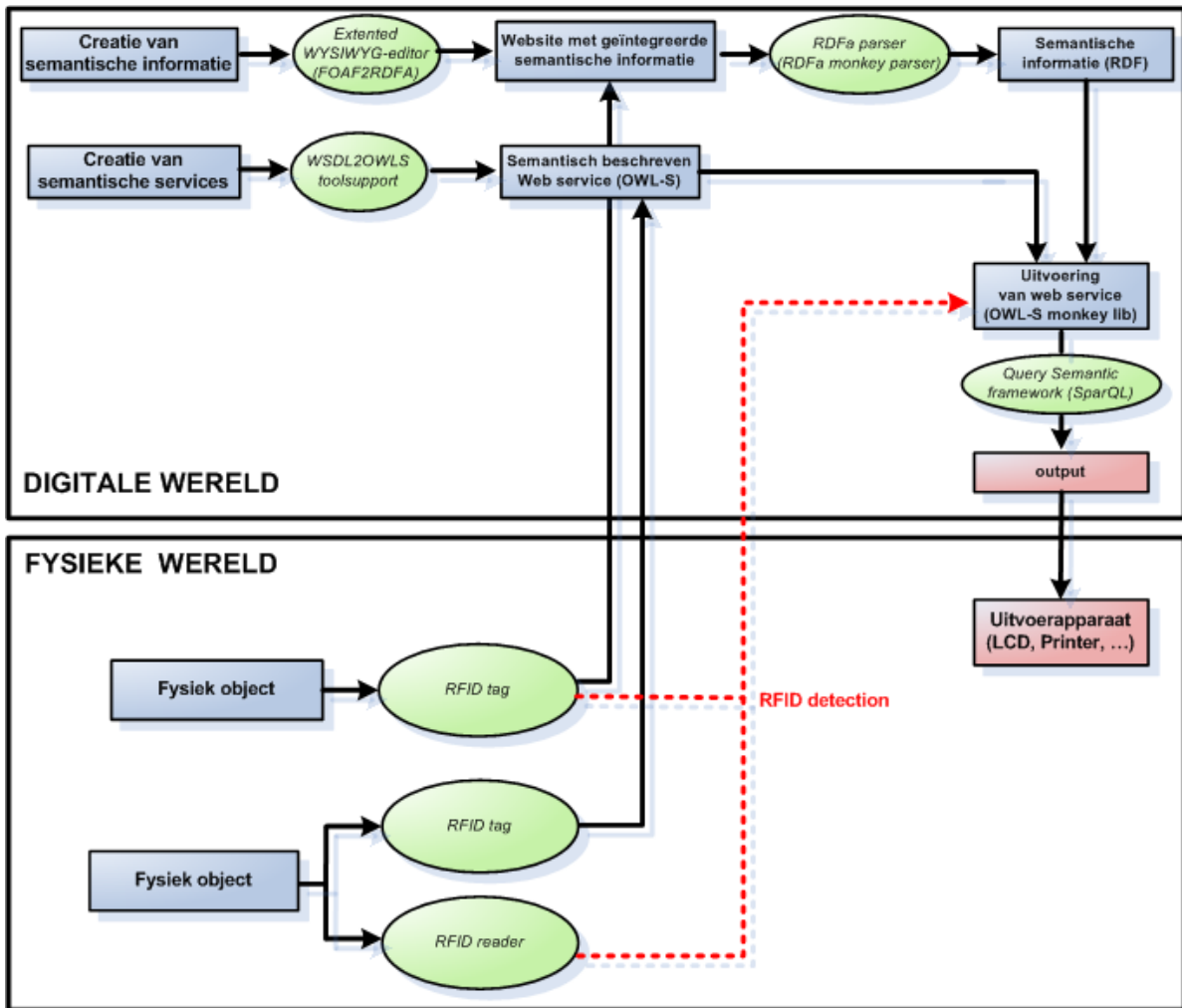
Deze conceptaanpak steunt op enkele aannames die het geheel mogelijk maken.

- Ten eerste wordt er verondersteld dat fysieke objecten in de toekomst *getagged* worden met een *uniforme identifier*. In deze thesis wordt er gebruik gemaakt van *RFID* en *Semacode tags* als identificatietechnologie. Doch zou andere bestaande of toekomstige technologie zonder problemen in dit concept gepast kunnen worden. *Atom tags*²⁰ zijn een voorbeeld van geavanceerde identificatie

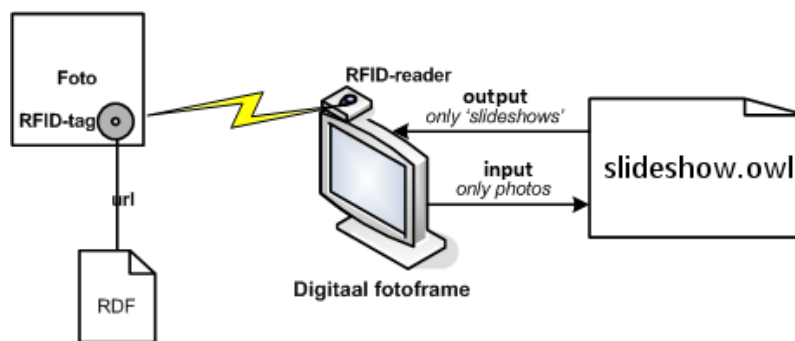
²⁰www.gotomobile.com/archives/next-generation-2d-code-atom-tag



Figuur 6.9: Overzicht van de volledige concept-architectuur



Figuur 6.10: Alternatief overzicht van de volledige concept-architectuur



Figuur 6.11: Voorbeeld van een mogelijke toepassing van de concept-architectuur

technologie die zonder problemen in deze architectuur zouden gepast kunnen worden. Merk op dat deze aanname reeds voor een groot deel voldaan is. Barcodes zijn alomtegenwoordig en kunnen gezien worden als een *uniforme identifier* voor fysieke objecten (zie figuur 6.12).



Figuur 6.12: Een universeel identificatiemiddel is reeds realiteit.

- Een tweede aanname bestaat erin dat deze *getagde* fysieke objecten (meer algemeen : augmented physical objects) op een bepaalde manier een verwijzing (meestal een URL) moeten bevatten naar een goed gedefinieerde computer interpreteerbare beschrijving (oftewel een semantische beschrijving). Dit lijkt op het eerste zicht ook een nogal zware aanname, doch als we kijken naar de explosie van het huidige WWW en de informatisering van onze hele maatschappij kan deze aanname sneller dan het aanvankelijk lijkt werkelijkheid worden.
- Web services kunnen inspelen op het feit dat ze semantische input- en outputtypes kunnen hebben. Ook dit is een probleem dat gerelateerd is met de vorige aanname. Wanneer het *Semantisch Web* populairder wordt, zullen web services hierop inspelen. Deze twee gerelateerde aannames zijn een gevolg van het 'kip en het ei'-probleem dat in sectie 6.3.1 beschreven werd.

6.6 Gerelateerd werk

De combinatie van RFID technologie en web services is niet onbestaande in de wetenschappelijke literatuur. Paliwal ea.[33] spitsen zich toe op het *compositen* van semantische webservices en passen dit toe in een grenscontrole scenario dat gebruik maakt van RFID technologie. In dit scenario worden afhankelijk van verschillende factoren, nieuwe webservices samengesteld en uitgevoerd. Hoewel er vrij generaliseerbare technieken worden gebruikt, wordt er geen aandacht besteed aan een veralgemeende architectuur die bruikbaar is in andere toepassingsdomeinen. Daarnaast wordt er geen link gelegd naar het gebruik van *augmented physical objects*. In [41] richt men zich meer op het zogenaamde '*Internet of Things*' door

middel van NFC²¹ en *Semantic web services*. De architectuur die in deze paper wordt voorgesteld bevat overeenkomsten met het concept dat in deze paper wordt voorgesteld, doch is het uitgewerkte prototype primitief. Deze paper bekijkt het volledige concept vanuit het oogpunt van een universele (mobiele telefoon) client die dienst doet als communicator tussen web services en fysieke objecten. Over het gebruik van meerdere fysieke objecten als input voor web services wordt niet gesproken. Fysieke *getagde* objecten worden eerder gezien als informatiebronnen die op een transparante manier geraadpleegd kunnen worden. Dit onderzoek maakt deel uit van het *Perci* project (*PERvasive ServiCe Interaction*) [5, 4] dat veel overlappings met de inhoud van deze thesis heeft. Toch legt dit onderzoek meer de nadruk op het proces waarbij user interfaces (semi-)automatisch kunnen worden afgeleid via semantische web services. Tevens legt men in deze studie weinig nadruk op de noodzakelijkheid van semantisch beschreven informatie op het huidige internet. In deze thesis wordt er meer de nadruk gelegd bij het praktisch gebruik en de generaliseerbaarheid van zulke *augmented objects* die worden ondersteund door *Semantic Web* technieken.

[39] levert een bijdrage aan het onderzoek naar *Semantic Web* services, via de realisatie van een uitgebreid en concreet scenario (nl. een *campus context-aware* applicatie). Een deel van dit onderzoek spits zich meer toe op het gebruik van semantic web services in combinatie met *context-awareness* technieken. Semantische web services worden gezocht (geselecteerd uit een lijst van services), gecombineerd met andere services en uiteindelijk gebruikt in zogenaamde *myCampus*-applicatie. Hierin verschilt *myCampus* met het concept in deze paper, waar we web services beschouwen als *augmented objects* die fysiek in onze wereld aanwezig kunnen zijn. Nogal wat onderzoek gaat uit naar de overbrugging van fysieke en digitale wereld, maar beperkt zich hierbij tot specifieke technologie zoals RFID[48].

Het volledige concept dat in deze thesis wordt voorgesteld *speelt* voortdurend met de grenzen van fysieke en digitale wereld, en tracht deze grenzen te vervagen. Tevens kan de architectuur worden gezien als een soort *interface*, een interface die toelaat dat menselijke gebruikers tegelijkertijd zowel kunnen interageren met fysieke als met digitale objecten. In *Being Digital*[29] beschrijft Negroponte de digitale toekomst en de relativiteit van ‘atomen’ en waarde van data (bits en bytes). Deze literatuur bevat veel raakvlakken met de architectuur die in deze thesis wordt voorgesteld. Negroponte bekijkt de overgang tussen digitale en fysieke wereld op een abstract niveau:

There will be specific points in space and time where bits get converted into atoms and the reverse.

Het is juist deze omzetting die de architectuur in deze thesis zo ‘eenvoudig’ mogelijk probeert te laten plaatsvinden. Deze overgang tussen digitale en fysieke wereld komt overeen met de *digital portal* die we in een eerder hoofdstuk hebben besproken (zie sectie 2.4).

²¹Near Field Communication

Hoofdstuk 7

Applicaties

Inhoudsopgave

7.1	Inleiding	67
7.2	‘STalkingObjects’-scenario	68
7.2.1	Aanpak	68
7.2.2	Ontologie beschrijving	69
7.2.3	Beheer van semantische informatie	70
7.2.4	Domeingerelateerde toepassingen	71
7.2.5	3D virtuele shopping-omgeving	72
7.2.6	Communicatie server applicatie	75
7.2.7	Veronderstellingen	76
7.2.8	Bestaande shoppingconcepten	76
7.3	‘SemaNews’-scenario	79
7.3.1	Algemeen idee	79
7.3.2	Creatie van fysieke krant op basis van digitale informatie	80
7.3.3	Gebruik van digitale informatie in de fysieke wereld	82
7.3.4	Registratie en selectie van web services	84
7.3.5	Toegepast voorbeeldscenario	85
7.3.6	Vereisten	87

7.1 Inleiding

In hoofdstuk 6 werd dieper in gegaan op een veralgemeende architectuur die ons in staat stelt om *augmented physical objects* op een transparante manier te laten interageren met elkaar, en op die manier de kloof tussen digitale en fysieke wereld te verkleinen. In dit hoofdstuk trachten we de bruikbaarheid en generaliseerbaarheid van deze architectuur aan te tonen. Dit doen we aan de hand van twee specifieke applicaties die zich in geheel verschillende toepassingsdomeinen bevinden.

Een eerste applicatie bevindt zich in het domein van een *winkelomgeving*, vermits dit toepassingsdomein vandaag de dag te kampen heeft met veel praktische problemen die te wijten zijn aan de *fysiek-digitaal kloof*.

Een tweede applicatie is gerelateerd met de wereld van de informatie en data die zich op het internet bevindt. Vaak is deze informatie enkel beschikbaar wanneer een menselijke gebruiker in de nabijheid van een computer is. Mobiele apparaten proberen dit probleem te verhelpen, toch is het moeilijk en

tijdrovend om bepaalde informatie vervolgens te bewerken of te gebruiken voor andere doeleinden (zoals bijvoorbeeld het toevoegen van informatie aan een kalender). Dit tweede scenario tracht op basis van informatie afkomstig van het internet automatisch een fysieke krant te creëren waarbij informatie uit deze krant toegankelijk en bruikbaar is voor allerlei doeleinden.

Deze applicaties zijn gebaseerd op dezelfde basisideeën (afkomstig uit hoofdstuk 6), maar bevatten specifieke identificatietechnologie die beter binnen het toepassingsdomein past.

Naast het feit dat we trachten na te gaan hoe toepasbaar deze architectuur is, bestaat er eveneens een *nevendoel* dat we via deze implementaties willen bereiken. Na de studie van het *Semantisch Web* en de daarbij horende technologieën (hoofdstuk 4), werd namelijk duidelijk dat deze technologie veel te bieden heeft. Enkele conclusies die men vaak in de context van het *Semantisch Web* trekt :

- *Semantische Web* technologieën maken data toegankelijker en bruikbaar.
- *Semantische Web* technologieën laten toe data over meerdere domeinen te gebruiken en kunnen efficiënt worden gecombineerd met andere nieuwe technologieën.
- Het *Semantisch Web* is vandaag de dag reeds bruikbaar.

Veel bronnen prijzen het *Semantisch Web* omwille van deze eigenschappen, doch zijn er weinig reële uitwerkingen te vinden die deze conclusies staven. Een doel van een gedeelteEen deel van de implementaties in deze thesis heeft dan ook als doel het praktisch verifiëren van deze conclusies. Enkele vragen die we ons hierbij kunnen stellen : ‘*Komt het Semantisch Web zijn beloftes na, maakt het data wel degelijk bruikbaar en toegankelijker?*’, ‘*Hoe goed is het Semantisch Web combineerbaar met andere (nieuwe) technologieën en is het praktisch doenbaar om data over de domeingrenzen nuttig te gebruiken door het gebruik van het Semantisch Web ?*’.

7.2 ‘STalkingObjects’-scenario

7.2.1 Aanpak

In sectie 6.3.3 werd een algemene architectuur beschreven waarbij we *augmented physical objects* kunnen ontwikkelen op basis van enkele standaard concepten (zoals bijvoorbeeld identificatietechnologie,...). In deze sectie bekijken we een applicatie waarin we deze architectuur als basis gebruiken, maar een invulling doen met specifieke domeingelateerde technologie.

Binnen het domein van een winkelomgeving zijn er veel problemen die te wijten zijn aan de kloof die er bestaat tussen de digitale en fysieke wereld. Eén van die problemen waarmee winkelgebruikers heden-daags moeilijkheden hebben, is het beheer van een budget terwijl men winkelt. Een ander probleem zijn wachtrijen in de winkels, doordat kassa’s voor vertraging zorgen. Deze kassa’s zorgen dat fysieke objecten worden gerelateerd met prijzen die digitaal worden opgeslagen, dit is een typisch voorbeeld van een *fysiek-digitaal*-proces dat voor veel vertraging kan zorgen. Andere problemen (die voornamelijk te wijten zijn aan de *fysiek-digitaal kloof*) die zich voordoen in winkelomgevingen worden uitgebreid opgesomd in Appendix C.2. In dit scenario gaan we dieper in op een beperkte selectie van deze problemen, maar doen we dit op een weldoordachte manier waardoor ook andere problemen binnen dit domein eenvoudiger op te lossen zijn.

In volgende secties gaan we dieper in op de specifieke stappen die werden doorlopen om te komen tot uiteindelijke oplossingen voor deze problemen.

7.2.2 Ontologie beschrijving

Wanneer we het *Semantisch Web* willen gebruiken om een bepaald domein te ondersteunen, is het bestaan van een ontologie (zie sectie 4.5.3) die dit domein voldoende beschrijft steeds noodzakelijk. In het *STalkingObjects*-scenario is er bewust gekozen voor een domein dat nauwelijks beschreven was in de wereld van het *Semantisch Web*. Hierdoor werd het noodzakelijk om bij aanvang van de implementatie van het winkelscenario, een eigen ontologie te beschrijven. Volgende paragraaf geeft meer inzicht op de manier waarop dit is gebeurd in deze thesis.

De beschrijving van een ontologie start met een grondige analyse van het domein dat men wil beschrijven. In dit geval werd er een grondige analyse gedaan van de domeinconcepten van een winkelomgeving. Deze analyse houdt eigenlijk in dat men menselijke kennis (die reeds van kinds af wordt geleerd en continu wordt bijgestuurd) gaat overbrengen naar een computerbeschrijving. Doch is het onmogelijk alle concepten (en gerelateerde concepten) te beschrijven, daarom moet men een bepaalde filtering doen op basis van toekomstige bruikbaarheid en relevantie. Het is belangrijk dat een ontologie vrij generiek geschreven wordt, hierdoor kunnen andere ontwikkelaars deze ontologie hergebruiken. Het proces om een ontologie te beschrijven gebeurt op basis van ervaring en gezond verstand. Dit proces kan vergemakkelijkt worden door de ondersteuning van specifieke tools.

In deze thesis is er gebruik gemaakt van de ontologie editor *Protégé*¹, die vrij beschikbaar is en veel bijval kent binnen de community rond het *Semantisch Web*. *Protégé* maakt het mogelijk om ontologieën te beschrijven in RDFS en OWL (meer info zie sectie 4.5.4), afhankelijk van de mate van precisie waarin men de ontologie wil beschrijven. Er is bij de beschrijving van een ‘winkel’-domein gekozen voor een RDFS-gebaseerde beschrijving vermits de mogelijke voordelen van OWL op dit moment nog beperkt zijn.

In tegenstelling tot het gebruik van *normale* databases maken ontologieën het mogelijk relaties te beschrijven tussen verschillende concepten in het domein. Men kan bijvoorbeeld uitgaande van instanties data (die gebruik maken van een winkel-ontologie) afleiden dat een *winkel* bestaat uit *winkelwagens* en *winkelrekken*. Een ander voorbeeld is dat men uit zulke data kan concluderen dat data die een *winkelproduct* voorstelt, als relatie ‘*isLocatedIn*’ enkel een verwijzing naar data die *winkelrekken* of *winkelwagens* voorstellen kan bevatten. Logische relaties (althans logisch voor mensen) kunnen via een ontologie eenvoudig beschreven worden. Deze logische relaties resulteren eveneens in hiërarchische relaties, hierdoor wordt het mogelijk deze relaties voor te stellen als grafen².

Appendix A (p92) toont de verschillende onderdelen van de winkel-ontologie in verschillende afbeeldingen. Figuur A.1 (p92) geeft meer zicht op de beschrijving van een individueel product. Merk op dat het *Semantisch Web* toelaat andere ontologieën (zoals *FOAF* en *GEO*) te hergebruiken in de beschrijving van dit concept. In de ontwikkeling van deze ontologie wordt er tevens gebruik gemaakt van kennis die in meer detail werd beschreven in sectie 2.4.2 omtrent digitale tegenhangers. Figuur A.2 (p93) geeft meer inzicht op de koppeling tussen producten en winkelgerelateerde concepten zoals *winkelrekken* en *winkelwagens*. Figuur A.3 (p94) laat zien hoe een shopping uiteindelijk semantisch beschreven kan worden en geeft de relaties aan ten opzichte van andere domeinconcepten. In alle concepten in deze ontologie is er ruimte voorzien voor het toevoegen van een digitale identificatie ID, dewelke nodig is willen we deze ontologie later *mappen* naar objecten in onze fysieke wereld.

De voordelen van zo’n ontologie zijn misschien niet meteen duidelijk (een probleem dat vaak voorkomt bij technologie rond het *Semantisch Web*), toch zal doorheen dit hoofdstuk het nut duidelijk gemaakt

¹<http://protege.stanford.edu>

²<http://www.pandd.demon.nl/grafen.htm>

worden. Een eerste voordeel dat meteen duidelijk kan gemaakt worden, is het feit dat wanneer verschillende ontwikkelaars gebruik maken van dezelfde ontologie, de data die beschreven is in deze ontologie eenvoudig gekoppeld kan worden. In het voorbeeld van een winkel-ontologie wordt het bijvoorbeeld mogelijk om verschillende winkels op een transparante manier met elkaar te koppelen, en zo het mogelijk te maken om prijzen van producten uit verschillende shoppings eenvoudig te vergelijken. De graaf in figuur A.4 (p95) laat zien hoe zulke shoppings gekoppeld zouden kunnen worden via een uitbreiding op de bestaande shopping-ontologie. Meteen is hiermee ook een nieuw voordeel duidelijk geworden, het is namelijk mogelijk om bestaande ontologieën vrij eenvoudig uit te breiden met nieuwe domeinconcepten. Hierdoor zijn ontologieën flexibel en kunnen ze geperfectioneerd worden op een gedistribueerde manier (via het internet). Dit is ondenkbaar binnen het domein van relationele databases, waar de structuur van een database meestal verborgen is hoewel daar weinig reden toe is. Ook is het DB-schema van relationele databases min of meer vast, waardoor het moeilijk is deze te veranderen of te migreren naar andere databases.

7.2.3 Beheer van semantische informatie

Instanties (data) van zo'n ontologie kunnen (net zoals informatie in het traditionele web) dienst doen als data voor databases op het web. Er zijn tevens verschillende raamwerken (*frameworks*) voorhanden die het mogelijk maken deze RDF-data te beheren in zogenaamde modellen die gemakkelijk ondervraagbaar zijn. Er zijn een heel aantal bestaande frameworks vrij beschikbaar, enkele bekendere frameworks zijn *Redland RDF application framework*³, *Jena Semantic Web Framework*⁴ en *Sesame RDF-framework*⁵. In deze thesis is er gekozen voor het framework *RAP - RdfApi-PHP*⁶ omwille van twee redenen nl.

1. Het framework is *light-weight* en dus gemakkelijk te installeren. Dit is een belangrijke factor die maar al te vaak wordt geminimaliseerd door andere frameworks.
2. Het volledige framework is geschreven in PHP, waardoor het volledig webgebaseerd en bijgevolg toegankelijk is.

Als nadeel heeft *RAP* dat het minder features bevat ten opzichte van andere (grotere) frameworks. Toch bevat het de belangrijkste feature, namelijk de creatie en het beheer van één of meer RDF modellen. Deze graafgebaseerde modellen kunnen bijgevolg via specifieke querytalen ondervraagd worden, de twee meest bekende voorbeelden van querytalen zijn *SparQL*⁷ en *RDQL*⁸. In deze thesis is er gebruik gemaakt van de nieuwste variant namelijk *SparQL*. In de uiteindelijke implementatie van het winkelscenario maakt het *RAP*-framework het dus mogelijk om via het web een RDF model (dat gebruik maakt van de shopping-ontologie) te ondervragen via SparQL queries.

Figuur 7.1 laat een voorbeeld zien van een SparQL-query die gebruik maakt van de gecreëerde shopping-ontologie in deze thesis. Dit voorbeeld maakt duidelijk dat het beschrijven van een ontologie het mogelijk maakt om op een natuurlijke wijze vrij geavanceerde *queries* op te stellen. De SparQL-query in figuur 7.1 geeft van producten in een bepaald winkelrek (dat via een RFID-tag wordt geïdentificeerd) de naam en de plaats van de producenten van deze producten terug. Bij het gebruik van relationele databases zou deze query vrij complex zijn, terwijl het gebruik van *Semantic Web* technologie deze queries vrij tot zeer goed leesbaar maakt.

³<http://www.redland.opensource.ac.uk>

⁴<http://jena.sourceforge.net>

⁵<http://www.openrdf.org>

⁶<http://sites.wiwiss.fu-berlin.de/suhl/bizer/rdfapi/>

⁷<http://www.w3.org/TR/rdf-sparql-query/>

⁸<http://www.w3.org/Submission/RDQL/>


```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>;
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dc:    <http://purl.org/dc/elements/1.1/>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX shop:  <http://www.avthasselt.be/shopping/shop.owl#>
PREFIX country: <http://www.daml.org/2001/09/countries/fips-10-4-ont#>

SELECT ?companyName, ?companyCountry WHERE
{
?shelf shop:hasIdentificationTag "SHELVE1001"
?product shop:isPlacedIn ?shelf
?product shop:isMadeBy ?company
?company foaf:name ?companyName
?company country:name ?companyCountry.
}';

```

Figuur 7.1: Concreet voorbeeld van een SparQL query die gebruik maakt van de shopping-ontologie

Na de beschrijving van een ontologie en het opzetten van een Semantisch framework, bekomen we een stevige basis waarop een toegepast scenario (in een winkelomgeving) met *augmented physical objects* kan worden gerealiseerd op basis van de architectuur uit 6.3.3.

7.2.4 Domeingerelateerde toepassingen

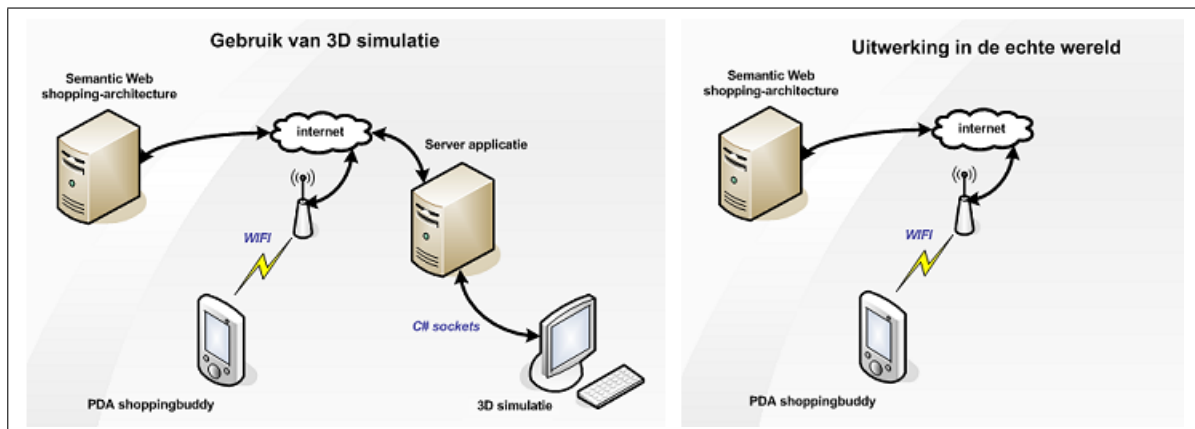
Wanneer we de architectuur uit hoofdstuk 6 dus willen *mappen* naar winkelomgevingen is er nood aan *augmented physical objects*, *identificatietechnologie* en *specifieke web services*. In dit scenario zullen winkelartikelen worden uitgebreid met RFID-tags (= identificatietechnologie), hierdoor ontstaan er potentiële *augmented physical objects*. Eveneens zal er een zogenaamde *shoppingbuddy* worden ontwikkeld die op de winkelkar wordt geplaatst, en specifieke winkelgerelateerde web services kan aanbieden en gebruiken.

De domeingerelateerde toepassingen die in deze sectie beschreven worden, trachten oplossingen te bieden voor enkele praktische problemen die zich vandaag de dag in shoppings voordoen. Vermits de praktische middelen voor deze thesis beperkt waren en het tevens onmogelijk is een scenario in een reële winkelomgeving uit te testen, is er voor gekozen een 3D testsimulatie te ontwikkelen. Het gebruik van virtuele testomgevingen is een vorm van *Rapid prototyping*, een techniek die eerder met succes werd uitgevoerd in verscheidene projecten[44, 40].

Deze volledige testsimulatie bestaat uit drie grote delen waarmee getracht wordt een zo natuurlijk mogelijk beeld te scheppen, dat tracht te bewijzen dat de implementatie in *real-life* wel degelijk mogelijk is. De drie grote delen die deze testsimulatie ondersteunen :

1. 3D virtuele shopping-omgeving
2. PDA ‘*Shoppingbuddy*’-applicatie
3. Communicatie server applicatie

Figuur 7.2.4 toont de verschillen tussen een mogelijke uitwerking in de echte wereld en de implementatie van het shopping-scenario in deze thesis. Globaal gezien zijn de verschillen tussen beiden vrij beperkt, er werd getracht een uitwerking te creëren die zo realistisch mogelijk overkomt.



Figuur 7.2: Structuur van de *STalkingObjects*-implementatie ten opzichte van een uitwerking in de echte wereld

In de volgende secties worden de drie grote delen (die eerder aangehaald zijn) in deze algemene structuur uit de doeken gedaan.

7.2.5 3D virtuele shopping-omgeving

Er is gekozen om deze omgeving in 3D te ontwikkelen, vermits deze manier het dichtst aansluit bij een natuurlijke uitwerking in de fysieke wereld. Deze implementatie is gebeurd in C# met behulp van de 3D-engine *TrueVision3D*⁹. Figuur 7.3 toont enkele screenshots van de 3D simulatie.



Figuur 7.3: Screenshots 3D virtuele shoppingomgeving

Deze applicatie laat de gebruiker toe om op een realistische manier door een shoppingomgeving te wan-

⁹<http://www.truevision3d.com/>

delen. Deze omgeving bevat de normale elementen die men in een winkel verwacht nl. producten, rekken, winkelwagens,... Er is getracht om afstanden tussen de verschillende elementen in deze omgeving realistisch te maken. Dit is van belang vermits er rekening gehouden moet worden met de eigenschappen van de gebruikte identificatietechnologie (zie hoofdstuk 3). Eveneens bevat deze applicatie mogelijkheden om zich sneller en trager doorheen de wereld te begeven, een winkelkar achter te laten, de zogenaamde ‘*digital aura*’ [10] te tonen,...

PDA ‘Shoppingbuddy’-applicatie

In een zogenaamde *next-gen shopping* (een shopping die tracht nieuwe digitale oplossingen te voorzien voor domeingerelateerde problemen) zal de (winkel)gebruiker communiceren met de digitale wereld en de *augmented objects* in de fysieke wereld via een handheld device. Dit kan bijvoorbeeld een PDA zijn of een ander soort embedded device. In enkele bestaande implementaties (oa. ‘*IBM’s Shoppingbuddy*’, *Peapod’s ‘Stop & Shop*’,...) gebruikt men de benaming ‘*shoppingbuddy*’ hiervoor.

Vermits zulke devices beperkt zijn op meerdere vlakken (schermgrootte, processorkracht,...), is er gekozen om een reële PDA-applicatie hiervoor te ontwikkelen. Deze applicatie is gekoppeld met de 3D simulatie (zie sectie 7.2.6) en zal in *realtime* nuttige informatie tonen terwijl de gebruiker zich door de virtuele omgeving begeeft. Eveneens heeft de gebruiker de mogelijkheid om zelf te interageren met de omgeving en contextuele informatie op te vragen. In de volgende paragrafen overlopen we de mogelijkheden die deze applicatie biedt en de wijze waarop dit is geïmplementeerd.

Winkelwagen informatie

De winkelwagen bevat zodus een *shoppingbuddy* die eveneens een RFID lezer bevat. Wanneer artikels van deze winkel binnen het bereik van de RFID lezer worden gebracht (in de winkelkar worden gelegd), zal een web service die gerelateerd is aan de *shoppingbuddy* worden aangesproken. Dit principe is volledig gebaseerd op de architectuur in hoofdstuk 6. De web service voegt de producten in de winkel toe aan een lijst, die door de shoppingbuddy in real-time wordt getoond. Wanneer er een product wordt toegevoegd of weggehaald, zal de shoppingbuddy dit meteen aangeven. Op die manier kan men aan budgetbeheer doen zonder zijn winkelgedrag te moeten aanpassen. Figuur 7.4 toont de output die zichtbaar is op de shoppingbuddy van de inhoud van de winkelwagen op een bepaald moment. In een soort van slideshow zullen alle producten die in de winkelwagen liggen in meer detail getoond worden. De gebruiker heeft eveneens de mogelijkheid om de eigenschappen van een product van dichterbij te bekijken.

2D Localisatie

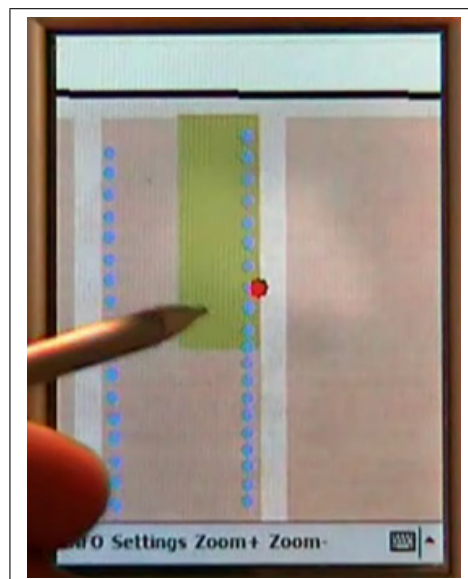
De shoppingbuddy is tevens in staat om de positie van een gebruiker in de winkel te tonen op een 2D plattegrond. Hoe dit precies wordt gerealiseerd, ligt buiten het bestek van deze thesis. Een mogelijkheid is door aan indoorlocalisatie te doen via vaste RFID-punten in de winkel. De shoppingbuddy voorziet verschillende *zoom* en *panning* mogelijkheden om een beter zicht te krijgen op de plattegrond van de winkel. Wanneer een gebruiker in de buurt komt van een nieuw winkelrek, zal de plattegrond dit eveneens grafisch aanduiden. Figuur 7.5 toont de manier waarop de shoppingbuddy 2D localisatie aanbiedt.

Mashups van winkelrekken - Productinformatie

Een gedeelte van de informatie die bekomen kan worden via de shoppingbuddy wordt rechtstreeks gehaald uit de ondervraging van het semantische framework. Dit semantische framework beheert een (graaf)model dat gebruik maakt van de eerder beschreven shoppingontologie. De implementaties van de shoppingbuddy



Figuur 7.4: Shoppingbuddy output : 'winkelwagen overzicht' afkomstig van semantische web service

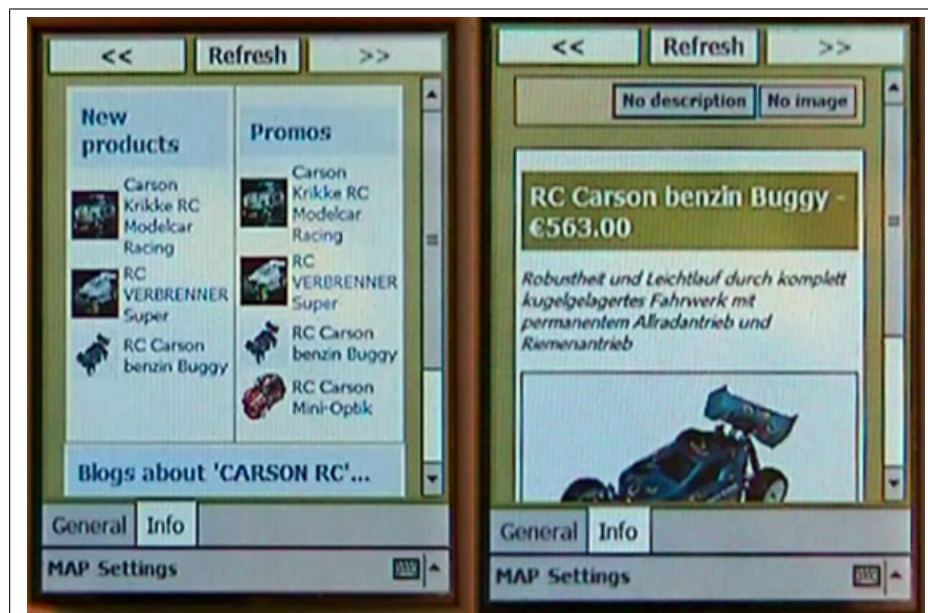


Figuur 7.5: Shoppingbuddy output : 2D plattegrond van de winkel

hieromtrent trachten de geavanceerde mogelijkheden van het gebruik van een graafmodel in combinatie met ontologieën aan te tonen.

Zo is het mogelijk dat de gebruiker in real-time een zogenaamde *mashup*¹⁰ [47] laat samenstellen van de artikelen in een bepaald winkelrek. Deze mashup kan bijvoorbeeld alle producten uit een rek die in promotie zijn tonen of alle nieuwste producten uit een bepaald rek. Eveneens wordt deze mashup aangevuld met internetblogs die producten uit het winkelrek bespreken. De linkerzijde van figuur 7.6 toont de mashup die de shoppingbuddy in real-time kan samenstellen.

Elk product kan eveneens in meer detail getoond worden, ook hiervoor worden er SparQL-queries *on-the-fly* samengesteld en uitgevoerd. De rechterzijde van figuur 7.6 toont de details van een bepaald product. Merk op dat door het maken van *mashups* van producten uit specifieke winkelrekken (dus niet uit de volledige winkel), we eveneens het *information overload* probleem (besproken in sectie 2.5.2) trachten te omzeilen. Daarnaast wordt er op basis van het profiel van de gebruiker relevante informatie geselecteerd. Zo toont de mashup bijvoorbeeld enkel nieuwe producten of producten die in promotie zijn, aangepast aan het profiel van de gebruiker. Dit kan gezien worden als een vorm van *profile matching* [43], een onderzoeksgebied dat vrij populair is de laatste jaren.



Figuur 7.6: Shoppingbuddy output: mashups en productinformatie

7.2.6 Communicatie server applicatie

Ten slotte is er een server-applicatie ontwikkeld om de interactie tussen de 3D simulatie en de PDA shoppingbuddy applicatie te voorzien. Deze applicatie is verantwoordelijk voor de communicatie tussen de 3D simulatie en de PDA-applicatie. Wanneer men een uitwerking doet in een echte shoppingomgeving, is deze applicatie en bijhorende communicatie (zoals dit wordt duidelijk gemaakt in figuur 7.2.4) overbodig. Deze

¹⁰A mashup is a website or application that combines content from more than one source into an integrated experience.

serverapplicatie is ontwikkeld in C# en maakt gebruik van .NET-sockets. Deze applicatie is geïntegreerd met de 3D simulatie applicatie, waardoor het gemakkelijker wordt om de volledige simulatie te laten lopen.

Tijdens de uitwerking van deze applicatie is duidelijk geworden dat het .NET-framework niet foutloos is vermits het nog steeds bugs¹¹ bevat in de component die socket-communicatie voorziet.

7.2.7 Veronderstellingen

In dit scenario worden enkele veronderstellingen gemaakt :

- Aanwezigheid van producten voorzien van RFID-tags in plaats van barcodes. Hoewel er nogal wat tegenstanders bestaan die uit onwetendheid angst hebben voor een wereld vol RFID-tags, zullen de voordelen van 'Radio Frequency Identification'-technologie in de loop der jaren verder doordringen.
- Aanwezigheid van indoor localisatie-technologie. Aan deze veronderstelling wordt reeds vandaag de dag hard gewerkt zodat er reeds sprake is van enkele oplossingen. Sommige daarvan maken gebruik van RFID-technologie[55, 54] maar ook via andere technologie[56, 8]. probeert men tot een oplossing te komen. Op dit moment bestaan er ook reeds commerciële¹²¹³ oplossingen.

7.2.8 Bestaande shoppingconcepten

De afgelopen jaren is er veel onderzoek gericht in totaaloplossingen die digitale technologie binnen het shoppingdomein introduceren. In de volgende secties wordt een overzicht gegeven van bestaande concepten en wordt aangegeven waarin deze verschillen van het *STalkingObjects*-scenario. Het is belangrijk in het achterhoofd te houden dat *STalkingObjects* geen oplossing biedt voor alle domeingerelateerde problemen, maar eerder tracht een *voedingsbodem* te zijn waarop kan worden verder gewerkt. De bestaande concepten worden eveneens vanuit dit standpunt bekeken.

Fujitsu Wireless U-Scan Shopper

Dit draadloos shopping cart¹⁴ systeem bestaat uit een mini-computer die op de winkelwagen geplaatst wordt. De services en informatie die dit systeem biedt, bestaan uit:

- Artikels scannen/ Onbemande kassa's
- Gepersonaliseerde aanbiedingen
- Aanbiedingen in de winkel die gelinkt zijn aan de locatie van de winkelswagen
- Opvragen van prijs en locatie van artikels
- Plaatsing van bestellingen in de winkel

Ondanks het feit dat dit systeem met de beperkingen van barcodes als identificatiemiddel moet leven, biedt het een heleboel interessante toepassingen. Er is weinig informatie te vinden over de onderliggende architectuur en aanpak. De implementatie in deze thesis steunt in dit gebied op het *Semantisch Web* en

¹¹<http://forums.microsoft.com/MSDN/ShowPost.aspx?PostID=871666&SiteID=1>

¹²<http://www.indoorgps.com/>

¹³<http://www.u-blox.com/technology/supersense.html>

¹⁴<http://www.physorg.com/news3143.html>

zijn technologieën. In Fujitsu's systeem is het *mashuppen* van de huidige staat van een rek niet ondersteund, eveneens zijn *on-the-fly* combinaties van verschillende producten niet aanwezig. Daarnaast is het niet mogelijk om een product te localiseren en de eigen positie in de winkel te bekijken. Het grootste nadeel blijft de manier waarop de interactie tussen producten gebeurt, dit moet namelijk via een hand-scanner gedaan worden. Figuur 7.7¹⁵ toont de shoppingbuddy van het Fujitsu Wireless U-Scan Shopper systeem.



Figuur 7.7: Fujitsu Wireless U-Scan Shopper

IBM customer Stop & Shop

*IBM's Stop & Shop (Shopping Buddy)*¹⁶ bestaat (net zoals Fujitsu's systeem) uit een draadloze mini-computer die op de winkelkar geplaatst wordt. Deze computer bevat eveneens een laserscanner die toelaat producten of klantenkaarten in te scannen. Grote voordeel van IBM's systeem is dat het ondersteund wordt door een volwaardig framework (*IBM's Store Integration Framework software*) dat toelaat de gebruiker persoonlijke informatie te geven. Hiervoor heeft *IBM* een samenwerkingsakkoord met *Cuesol*, een bedrijf dat gespecialiseerd is in het werken rond *customer purchase history*. Een overzicht van enkele mogelijkheden van *IBM's Stop & Shop*:

- Shopping geschiedenis en favorieten + shoppinglijstjes die men kan emailen naar de winkel
- Notificatie van favoriete artikels wanneer een persoon in de buurt komt van deze artikels.
- Gepersonaliseerde aanbiedingen

Dit systeem biedt veel geavanceerde mogelijkheden, doch het is moeilijk te zeggen waar de grenzen van dit systeem zich bevinden. Hoe wordt er met data in de winkel omgesprongen? Kan deze data ook gebruikt worden in andere toepassingsdomeinen? De gebruiker is in dit systeem eveneens verplicht om handmatig producten te scannen, er is dus geen moderne identificatie-technologie aanwezig. Figuur 7.9¹⁷ toont de shoppingbuddy van het IBM customer Stop & Shop systeem.

¹⁵<http://xataka.com/images/>

¹⁶<http://www-03.ibm.com/industries/retail/doc/content/news/pressrelease/1162561101.html>

¹⁷http://money.cnn.com/2005/06/21/news/fortune500/supermarket_showdown/



Figuur 7.8: IBM customer Stop & Shop

METRO Group Future Store

De *METRO Group Future Store*¹⁸ is een volwaardige implementatie van een zogenaamde *next-gen shopping*. Dit initiatief is ontstaan uit een samenwerkingsverband van meerdere bedrijven waaronder IBM, Intel, T-Systems,... Het is dan ook niet verwonderlijk dat de shoppingbuddy die in de METRO-stores wordt gebruikt, qua eigenschappen vergelijkbaar is met IBM's Stop & Shop. Het grote verschil tussen dit concept en IBM's Shop & Stop is het feit dat deze shoppingbuddy werkt met behulp van RFID-identificatie. De METRO-store biedt naast een shoppingbuddy ook een heleboel andere *next-gen shopping* concepten waaronder oa. :

- *Information Terminals* die extra informatie over producten kunnen geven
- *Everywhere Display*, beelden die informatie over producten projecteren op de grond van de winkel
- *Smart Scales*, weegschalen die *zien* welke producten op de schaal liggen
- *Electronic Advertising Displays*, grote beeldschermen die meer informatie geven over producten
- *Electronic Shelf Labels*, elektronische prijslabels die automatisch vanop afstand kunnen aangepast worden
- *Personal Digital Assistant*, winkelbedienden kunnen via PDA verschillende eigenschappen van de winkel veranderen
- *Store Manager's Workbench*, dit is een systeem bestaande uit een groot beeldscherm dat een totaalbeeld kan geven van de winkel waarop bijvoorbeeld de *stock* kan bekeken worden.
- ...

De *METRO Group Future Store* is een uitgebreide shopping die veel nuttige concepten bevat. Hoewel het veel toepassingen kent binnen het shoppingdomein, bevat het weinig concepten die een overlapping proberen te maken met het dagelijkse leven van de shopper. Een voorbeeld hiervan is de mogelijkheid dat iemand anders van de familie (die niet aanwezig is in de shopping) de shoppinglist kan uitbreiden via

¹⁸www.future-store.org/

het internet. Een ander voorbeeld zou kunnen zijn : het zoeken van overeenkomsten tussen de agenda van een bepaald persoon en benodigde producten hiervoor. Dit zijn geavanceerde mogelijkheden waarvan onmogelijk gezegd kan worden of de architectuur van *METRO-store* dit toelaat. Door gebruik te maken van Semantische Web technologie liggen zulke oplossingen binnen de mogelijkheden. Figuur 7.9¹⁹ toont de zogenaamde *Personal Shopping Assistant* uit de METRO-stores.



Figuur 7.9: Metro's Personal Shopping Assistant

7.3 'SemaNews'-scenario

7.3.1 Algemeen idee

Een tweede scenario waarvoor een implementatie is gebeurd in deze thesis is getiteld '*SemaNews*'. Het basisidee bestaat eruit dat informatie (zoals nieuws, weerberichten, blogposts, ...) die hedendaags op het internet te vinden is, vrij beperkt beschikbaar is. Wanneer menselijke gebruikers deze informatie willen raadplegen is het namelijk noodzakelijk dat dit (in bijna alle gevallen) gebeurt in de nabijheid van een desktop computer of laptop. Ondanks de digitale revolutie van de laatste decennia, kiezen mensen in verschillende scenario's nog steeds voor 'ouderwetse' technologie boven nieuwe digitale. Een goed voorbeeld hiervan zijn '*e-books*' die ouderwetse kranten en boeken zouden verdringen in de toekomst. Hoewel *e-books* in vergelijking met de normale papieren versie veel voordelen hebben (informatie kan gelinkt worden met andere informatie, internet binnen handbereik,...), verkiezen de meeste mensen toch de papieren versies vermits het gebruiksgemak hoger ligt (leercurve is bijna onbestaande) en het lezen van een papieren versie simpelweg aangenamer is.

Het scenario dat in de volgende secties uit de doeken wordt gedaan, probeert te zoeken naar combinaties van digitale en fysieke technologie, dit startend vanuit de architectuur in hoofdstuk 6. Het scenario bestaat uit twee stappen namelijk:

1. Het creëren van een fysieke krant op basis van informatie afkomstig uit de digitale wereld. Hierbij wordt op basis van Atom en RSS feeds automatisch een krant gecreëerd en aangevuld met speciale barcodes.

¹⁹<http://www.best-practice-business.de/blog/>

2. Het gebruik van digitale informatie in de fysieke wereld.

7.3.2 Creatie van fysieke krant op basis van digitale informatie

Algemeen proces

De idee achter het creëren van een fysieke krant op basis van digitale informatie is dat digitale informatie tegenwoordig deel uitmaakt van ons dagelijks leven maar in veel gevallen enkel via computer geraadpleegd kan worden. Net zoals een *normale* krant, zou zulke informatie ook bruikbaar moeten zijn in andere omgevingen zoals bijvoorbeeld de ontbijttafel, de trein,...

Om dit te verhelpen is er in deze thesis een implementatie gebeurd die op basis van zowel RSS- als Atom-feeds automatisch een PDF-document kan samenstellen. De layout van dit document kan volledig aangepast worden naar de noden van de gebruiker. Hiernaast kan de gebruiker ook nog typische eigenschappen specificeren zoals het filteren van images, de grootte van images (indien aanwezig),... Eenmaal het PDF-document is gecreëerd, wordt dit document automatisch afgedrukt. Dit volledige proces kan op een tijdstip naar keuze worden gedaan, bijvoorbeeld elke ochtend om 7u.

Verschillende initiatieven (FeedJournal²⁰, SimplyHeadlines²¹,...) bieden reeds de mogelijkheid aan om dit proces te ondersteunen. Doch verschilt de implementatie in deze thesis erin dat de krant wordt aangevuld met speciale barcodes genaamd *Semacode tags*²² (meer informatie zie sectie 3.3). Deze *Semacode tags* doen dienst als een link tussen de fysieke informatie en digitale informatie die online te vinden is. Met andere woorden de *Semacode tags* bevatten een URL die verwijst naar de informatie online. *Semacode tags* kunnen eender welke informatie bevatten gaande van URL's tot stukken tekst. De informatie in zulke *Semacode tags* kan eenvoudig worden gedecodeerd door middel van *image-recognition algoritmes* [37]. In volgende paragraaf wordt uitgebreid beschreven waarvoor deze semacodes in de krant uiteindelijk worden gebruikt.

Implementatie details

Voor de realisatie van het zonet beschreven proces wordt gebruik gemaakt van een combinatie van eigen implementatie en verschillende externe applicaties. Figuur 7.10 geeft een zicht op het proces dat het creëren van een speciale krant op basis van RSS- en Atom-feeds mogelijk maakt.

In de *SemaNews-component* kan de gebruiker een layout definiëren en eigen RSS- en Atom-feeds aangeven. Wanneer het volledige proces wordt gestart, zullen de verschillende *feeds* geraadpleegd worden op het internet en samengevoegd worden tot één HTML-bestand (zie figuur 7.10 '*SemaNews RSS/Atom parser*'). Eveneens worden van elke *feed* verschillende semacodes aangemaakt (zie figuur 7.10 '*SemaNews Semacode generator*'), die eveneens een plaats krijgen op een HTML-pagina. Vervolgens worden de resulterende HTML-bestanden uit beide subcomponenten, samen met een opmaakbestand (CSS) doorgegeven, aan de '*HTML-to-PDF component*'. Deze component bestaat uit een applicatie genaamd *HTMLDoc*²³ die een HTML-document kan omzetten naar een PDF-bestand. Het resulterende PDF-bestand kan vervolgens automatisch worden afgedrukt via de command-line tool *NirCmd*²⁴.

²⁰<http://www.feedjournal.com/>

²¹<http://simplyheadlines.com/>

²²<http://www.semacode.org>

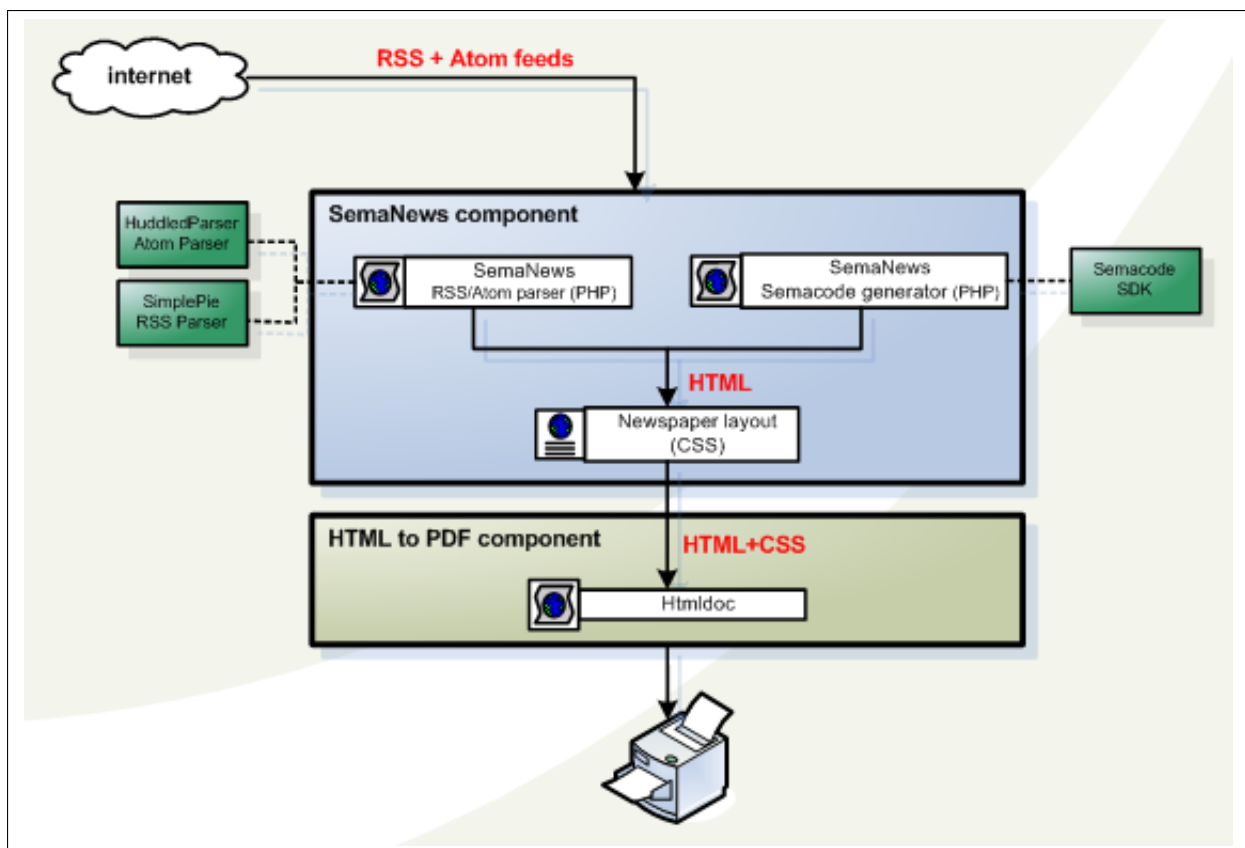
²³<http://www.htmldoc.org/>

²⁴<http://www.nirsoft.net/utills/nircmd.html>

Dit volledige proces wordt uitgevoerd in een uitvoerbaar (*.bat*)bestand dat als *geplande taak* automatisch kan worden uitgevoerd. De PHP-pagina's uit de *SemaNews-component* worden aangeroepen via *cURL*²⁵, een command-line tool die toelaat *output* van webpagina's op te vangen.

Er worden overigens enkele specifieke bibliotheken gebruikt in de *SemaNews-component* :

- *SimplePie*²⁶, RSS/Atom-parser. Vermits deze parser problemen leverde met Atom-parsing (o.a. Gmail²⁷ biedt Atom-feeds aan), is er gekozen om Atom-feeds via een andere parser af te handelen.
- *HuddledParser*²⁸, Atom-parser.
- *Semacode Tagger SDK for Web 1.5*²⁹, creatie van Semacodes.



Figuur 7.10: SemaNews : creatie van fysieke krant op basis van RSS en Atom feeds.

²⁵<http://curl.haxx.se/>

²⁶<http://simplepie.org/>

²⁷<http://www.gmail.com>

²⁸<http://huddledmasses.org/jaykul/huddledparser2/>

²⁹<http://semacode.org/software/>

7.3.3 Gebruik van digitale informatie in de fysieke wereld

Algemeen proces

Vermits de digitale informatie (komende van RSS- en Atom-feeds) nu in de fysieke wereld beschikbaar is geworden, verliest het zijn digitale karakter. Dit heeft enkele nadelen, het wordt namelijk moeilijk om deze digitale informatie te gebruiken voor andere doeleinden. Een voorbeeld hiervan kan zijn dat iemand die in de krant een aankondiging leest over een bepaald concert deze gebeurtenis wil toevoegen aan zijn/haar persoonlijke digitale kalender. Meestal zal de toevoeging van zulke gebeurtenissen aan een digitale kalender handmatig gebeuren, waarbij de gebruiker informatie (uit de fysieke wereld komende van de krant) handmatig overbrengt naar de digitale wereld (in dit geval de online kalender).

Het scenario in deze thesis probeert dit probleem te verhelpen op een zo transparant mogelijke manier die uitbreidbaar en scaleerbaar is en eveneens gebaseerd op *W3C*-standaarden. Volgend op het proces in paragraaf 7.3.2, waar een krant gecreëerd werd op basis van RSS- en Atom-feeds, kan een gebruiker informatie uit deze krant eenvoudig gebruiken als volgt:

1. Via de *SemaNews* mobiele telefoonapplicatie maakt de gebruiker een foto van een bepaalde *Semaooode tag*. Elk nieuwsbericht in de krant is gerelateerd met een *Semaooode tag*.
2. De mobiele applicatie toont vervolgens een selectie van bruikbare web services voor een bepaald nieuwsbericht. Vervolgens kan de gebruiker een keuze maken uit deze selectie van bruikbare web services. Eenmaal dit gebeurd is zal de web service uitgevoerd worden. Deze web services kunnen van uiteenlopende aard zijn, gaande van het toevoegen van informatie aan een online kalender tot het bekijken van een wegenkaart.

Om de mogelijkheden van dit proces beter te kunnen begrijpen, beschrijven we in de volgende secties de achterliggende werking. Dit doen we op basis van figuur 7.11 die een overzicht geeft van de technische uitwerking van dit geheel proces.

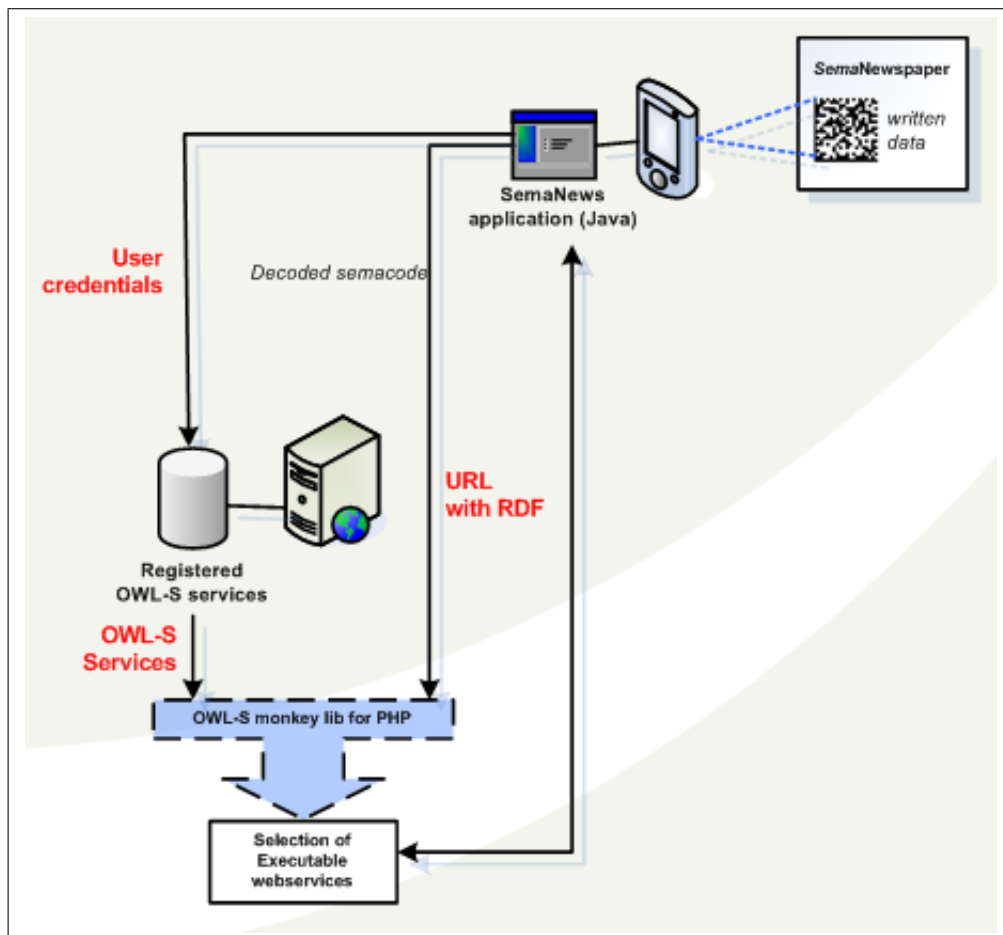
De drie belangrijkste componenten in het proces om digitale informatie te gebruiken in een fysieke wereld zijn de ‘*SemaNews (mobiele telefoon) applicatie*’, de (in sectie 6.3.2 beschreven) ‘*OWL-S wrapper*’ en de ‘*Registered OWL-S services database*’. Via de mobiele telefoon applicatie wordt de data van een bepaalde *Semaooode tag* geëxtraheerd. Deze data stelt een URL voor die verwijst naar een bepaalde webpagina die RDF-data bevat, oftewel een *semantisch beschreven webpagina*.

Elke gebruiker van *SemaNews* heeft een eigen online account. Hiermee kan de gebruiker een selectie aangeven uit een bepaalde hoeveelheid geregistreerde OWL-S web services. Wanneer een bepaalde service bijvoorbeeld nooit geschikt is voor een bepaalde gebruiker, kan deze gebruiker dit aangeven op de *SemaNews*-website. Hierdoor zal het uiteindelijke (automatisch) selectie-proces veel gepersonaliseerder zijn en zal de gebruiker nog eenvoudiger informatie uit de krant kunnen gebruiken.

Op basis van de RDF-data (die is bekomen via het nemen van een foto van een bepaalde *Semacode tag*) en een selectie van OWL-S web services (die specifiek is voor elke *SemaNews*-gebruiker), kan de *OWL-S wrapper* nagaan welke web services bruikbaar zijn voor specifieke informatie die in de krant te vinden is. De manier waarop de OWL-S wrapper dit nagaat werd reeds in sectie 6.3.2 beschreven.

Men heeft bijvoorbeeld volgende selectie van web services geregistreerd op de *SemaNews* website :

- Toevoegen van data aan Google Notebook
- Toevoegen van events aan Google Calendar



Figuur 7.11: SemaNews : gebruik van digitale informatie in de fysieke wereld.

Wanneer men in dit geval een *Semacode tag* decodeert die gerelateerd is met ‘een gedicht’, zal de *OWL-S wrapper* automatisch de ‘*Google Calendar*’-service weghfilteren. Vermits deze component kan ‘zien’ of bepaalde RDF-data al dan niet geschikt is voor bepaalde web services.

De gebruiker kan uiteindelijk een bepaalde web service uitvoeren. Ook dit wordt afgehandeld door de *OWL-S wrapper*. In sectie 7.3.5 wordt een toegepast scenario beschreven dat dit gehele proces verduidelijkt.

7.3.4 Registratie en selectie van web services

In de implementatie van deze thesis zijn drie verschillende web services geschreven die door het *SemaNews framework* gebruikt kunnen worden.

- Toevoegen van events aan Google Calendar³⁰
- Toevoegen van data aan Blogger³¹
- Tonen van locaties op Yahoo! maps³²

Registratie van web services

Het *SemaNews* framework werkt volgens het principe dat *service agents* OWL-S services kunnen aanbieden (registreren) via de *SemaNews* website. Wanneer een ontwikkelaar bijvoorbeeld een service heeft geschreven die aan de hand van een locatie het huidige weer kan teruggeven, kan deze bekend gemaakt worden via de *SemaNews* website. In de toekomst zou het beter zijn, mochten zulke services op een meer ‘erkende’ plaats gecentraliseerd worden zoals nu reeds het geval is met WSDL-services³³. Deel 1 uit figuur 7.12 toont het registratieformulier op de *SemaNews* website.

Selectie van web services

De *SemaNews*-gebruikers kunnen een selectie maken uit de geregistreerde OWL-S services. Als een gebruiker bijvoorbeeld geen *Blogger*-account heeft, is het nutteloos deze service op te nemen in deze selectie. Deel 2 uit figuur 7.12 toont het formulier dat dit proces op eenvoudige wijze toelaat. Deel 3 uit dezelfde figuur toont de voorpagina van de *SemaNews*-website.

Implementatie details

De mobiele telefoon applicatie is geschreven voor het *Java ME* platform³⁴ en maakt gebruik van de *Semacode Reader SDK*. Deze applicatie is uitvoerbaar op een groot aantal (Java-compatibele) mobiele telefoons en smartphones die een digitale camera bevatten.

De implementatie details van de *OWL-S wrapper* werden reeds in sectie 6.3.2 beschreven. Deze wrapper werd uiteindelijk gebruikt in een specifiekere PHP-pagina die aangepast is aan het *SemaNews*-scenario.

De realisatie van de verschillende web services is gebeurd via het PHP-pakket *PEAR::SOAP*, dat automatisch een WSDL beschrijving kan aanmaken op basis van standaard PHP-functies. Vervolgens zijn via een specifieke applicatie ‘*WSDL-to-OWLS*’ (die zich in de bibliotheek van OWL-S Mindswap bevindt) de verschillende WSDL-services omgezet naar OWL-S services. Dit proces neemt weinig tijd in beslag maar

³⁰<http://calendar.google.com>

³¹<http://www.blogger.com>

³²<http://maps.yahoo.com>

³³<http://www.xmethods.net>

³⁴<http://java.sun.com/javame>



Figuur 7.12: De belangrijkste onderdelen van de SemaNews website

vereist nog te veel kennis van zaken, in dit onderzoeksgebied is er nog veel werk aan de winkel. Hoofdstuk 5 gaat dieper in op de gebruikte technologieën omtrent web services en de achterliggende redenen voor het gebruik van deze services.

7.3.5 Toegepast voorbeeldscenario

Wanneer Eline 's morgens opstaat heeft haar computer reeds een gepersonaliseerde krant gedrukt. Deze krant is volledig gebaseerd op informatie die digitaal ter beschikking wordt gesteld. Eline heeft de inhoud (RSS- en Atom-feeds) en de layout reeds op voorhand bepaald. Op deze manier kan Eline nu digitale informatie op een natuurlijke wijze raadplegen terwijl ze ontbijt. Figuur 7.13 toont een overzicht op dit proces.

Via deze dagelijkse krant kan ze haar e-mail berichten eenvoudig raadplegen (op basis van de feed die naar haar GMail-mailbox verwijst). De krant van vandaag toont kalenderinformatie over het *Play Festival* waar ze naartoe wil gaan met haar vriendin. Graag zou ze de informatie over dit festival willen toevoegen aan haar *Google Calendar*³⁵, doch heeft ze weinig tijd en zin om deze informatie handmatig in te geven in haar *Google Calendar*. Ook moet ze zich dan naar haar computer begeven, vermits haar smartphone beperkte inputmogelijkheden heeft (geen toetsenbord, kleine resolutie,...).

Via het *SemaNews*-framework kan ze dit probleem verhelpen. Nadat ze de *SemaNews* applicatie op haar smartphone heeft gestart, maakt ze een foto (zie figuur 7.14) van de *Semacode tag* die hoort bij de kalenderinformatie over het '*Play Festival*'.

Eenmaal de *SemaNews* applicatie op haar smartphone deze semacode heeft gedecodeerd, krijgt ze een lijstje van mogelijke web services die ze kan gebruiken (Zie figuur 7.15 links). Het *SemaNews* framework heeft rekening gehouden met het soort van de data, en toont daarom enkel web services die relevant

³⁵<http://calendar.google.com>



Figuur 7.13: Overzicht van generatie-proces van gepersonaliseerde SemaNews-krant



Figuur 7.14: Het vastleggen van een Semacode tag door middel van de SemaNews mobiele telefoonapplicatie



Figuur 7.15: Visualisatie van de mogelijk uit te voeren web services

zijn. Zo zal het *SemaNews*-framework het automatisch zien als er voor een bepeeld event een locatie is opgegeven. Deze locatie is dan ook meteen toonbaar op een digitale wegenkaart via één van de geregistreerde web services.

Wanneer Eline een keuze heeft gemaakt uit de automatisch geselecteerde web services, zorgt het *SemaNews*-framework ervoor dat deze service wordt uitgevoerd. Het is het *SemaNews*-framework zelf dat bepaalt welke inputparameters gebruikt (en doorgegeven) moeten worden. De rechterzijde van figuur 7.15 toont bijvoorbeeld de output wanneer Eline er voor heeft gekozen de locatie van het 'Play Festival' op *Yahoo! maps* te tonen.

7.3.6 Vereisten

Dit scenario stelt enkele vereisten en doet tevens enkele aannames om het geheel praktisch werkbaar te maken:

- Webpagina's bevatten semantische informatie in de vorm van RDF of RDFa. Dit is een veronderstelling die op dit moment zwaar is en waar deze thesis (onder andere in sectie 6.3.1) zijn steentje aan probeert bij te dragen. Gebruikers hoeven hiervoor geen specifieke domeinkennis te hebben, technisch kan dit alles mooi afgeschermd worden. De opbouw van een zogenaamd *Semantisch Web* kan enkel realiteit worden wanneer dit op grote schaal collectief gebeurt.
- In het *Semantisch Web* is het mogelijk om soortgelijke dingen op meerdere manieren te beschrijven. Onderzoek naar 'ontology matching' probeert deze problemen te verhelpen. In de toekomst zullen naar alle waarschijnlijkheid enkele ontologieën de bovenhand nemen waardoor de nood aan *ontology matching* beperkt blijft. In het *SemaNews* scenario is er gewerkt met de *ROR*-ontologie. *Resources of a Resource*³⁶ is een ontologie die toelaat een ruim domein gaande van blogs, kalender-events, jobs tot personen, muziek, films, . . . semantisch te beschrijven. De *ROR*-ontologie wordt reeds op verschillende websites gebruikt en is met RDF gerealiseerd.
- In dit scenario is er gekozen voor *OWL-S* web services die zogenaamde *WSDL-groundings* bevatten. Hoewel beide technologieën goed op elkaar aansluiten, bestaat in dit domein nog geen uitgesproken standaard. In de toekomst zal pas duidelijk worden welke web service technologieën de bovenhand nemen. Doch is het doel van deze thesis niet aan te tonen wat dé beste technologie is op dit moment. Wel trachten we te zoeken naar meer oplossingen die onafhankelijk zijn van welbepaalde technologie. De vertaling van *WSDL* services naar semantisch beschreven (*OWL-S*) web services wordt overgelaten aan de ontwikkelaars van web services.

³⁶<http://www.rorweb.com>

Deel IV
Conclusie

Hoofdstuk 8

Conclusie

Deze thesis spitst zich toe op de kloof die er bestaat tussen onze fysieke wereld en de *nieuwe* digitale wereld. Vanuit theorie, opgedaan in meerdere onderzoeksgebieden, wordt er stap voor stap gewerkt naar een nieuwe manier om tegen deze problemen aan te kijken.

Op basis van een studie naar verschillende technologieën (die zich in verschillende onderzoeksgebieden bevinden), wordt er een generieke architectuur beschreven. Deze architectuur kan dienst doen als basis voor meerdere toepassingsdomeinen waarin we geconfronteerd worden met problemen die te wijten zijn aan een *fysiek-digitaal* kloof. Het *Semantisch Web* is samen met het gebruik van *Augmented Physical Objects* een belangrijke peiler in dit concept.

Om deze architectuur werkbaar te maken, is er een *OWL-S wrapper* geschreven die het mogelijk maakt dat web services op een meer automatische manier kunnen worden gebruikt. Daarnaast zijn er twee implementaties geschreven, die een bijdrage leveren aan algemene ideeën achter de architectuur. *FOAF2RDFS* tracht de uitbreiding van het *Semantisch Web* te ondersteunen door FOAF gebaseerde informatie op een eenvoudige manier herbruikbaar te maken. *RDFS monkey parser* is een implementatie van een parser die in staat is om semantische informatie (in RDFS formaat) te extraheren uit HTML-pagina's.

Verder worden er twee toegepaste implementaties beschreven die zich volledig baseren op de eerder beschreven architectuur. Deze beide implementaties maken gebruik van de conceptuele architectuur maar onderscheiden zich van elkaar door het gebruik van specifieke (domeingelateerde) identificatie-technologie.

Een eerste scenario (*STalkingObjects*) tracht aan de hand van *Radio Frequency Identification technologie* oplossingen te bieden voor problemen die zich afspelen binnen winkelomgevingen. Voor dit scenario werd een 3D simulatie-omgeving ontwikkeld in combinatie met een gesynchroniseerde PDA-applicatie die dienst doet als zogenaamde '*shoppingbuddy*'. Eveneens wordt er hierbij gebruik gemaakt van een Semantisch framework, dat instaat voor het beheer van de data in winkelomgevingen. Dit Semantisch framework werkt op basis van een nieuwe shopping ontologie die het mogelijk maakt geavanceerde queries uit te voeren en op die manier geavanceerde oplossingen te bieden. Hierdoor wordt het bijvoorbeeld mogelijk om de gebruiker in *real-time* een *mashup* van digitale informatie te laten zien op basis van producten uit een bepaald winkelrek.

Een tweede scenario (*SemaNews*) probeert digitale informatie beschikbaar te maken in de fysieke wereld. Dit gebeurt door digitale informatie samen te voegen tot een fysieke krant waarbij de informatie zijn

digitaal karakter blijft behouden. Om dit te bereiken wordt er gebruik gemaakt van de conceptuele architectuur in combinatie met specifieke identificatietechnologie nl. *Semacode tags*. Via een mobiele telefoon applicatie kunnen gebruikers informatie uit deze krant gebruiken om vervolgens een selectie geavanceerde web services aan te spreken. Deze selectie is geautomatiseerd door gebruik te maken van *Semantische Web* technologie en kan verder worden gespecificeerd via de *SemaNews*-website.

Toekomstig werk kan zich richten op verschillende onderzoeksgebieden die overlappingen vertonen met de implementatie binnen deze thesis. Het uitbreiden van de huidige architectuur naar een architectuur die korter aansluit bij eerder beschreven architecturen omtrent het *Internet of Things*.

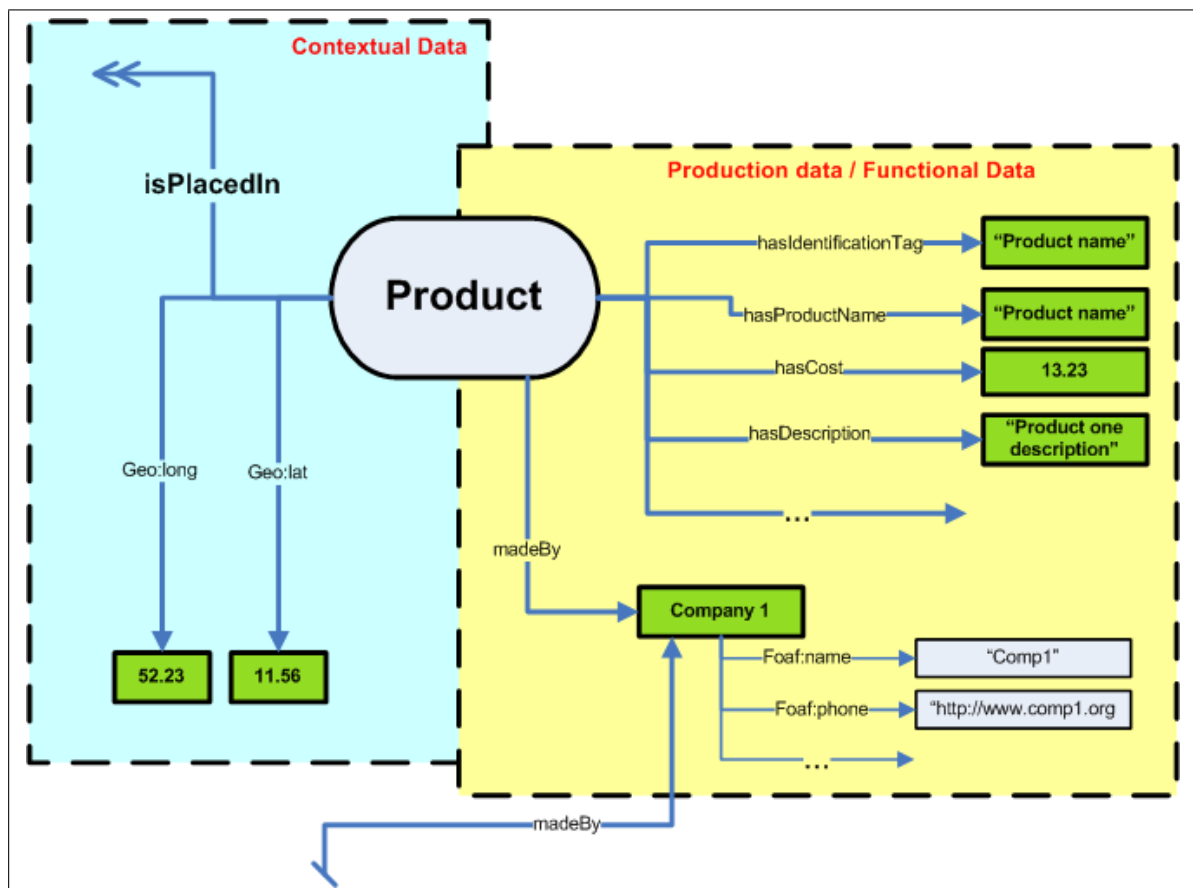
Ander verder onderzoek kan zich richten op specifieke problemen binnen de wereld van het *Semantisch Web*. De verbetering van het *matching proces* om de combinatie van verschillende ontologieën bruikbaar te maken in de praktijk behoort hier toe.

Het grootste deel van het toekomstig werk bevindt zich niet in het technische aspect, wel in de manier waarop we tegen bepaalde technologie (oa. het *Semantisch Web*) uit deze thesis aankijken. De manier waarop zo'n technologie wordt voorgesteld en uitgelegd, is van cruciaal belang voor de doorbraak in de praktijk naar een groter publiek.

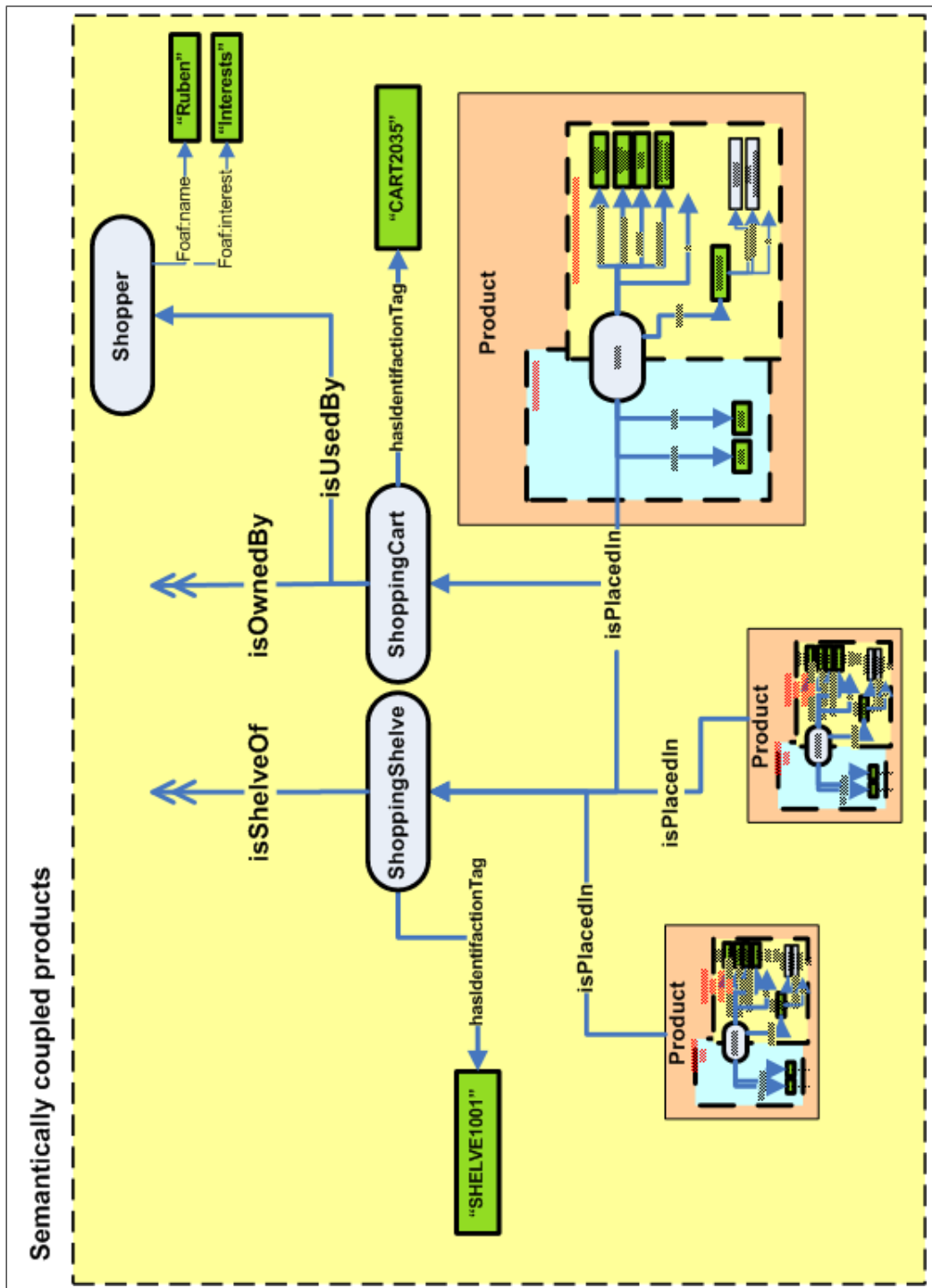
Deel V
Appendices

Bijlage A

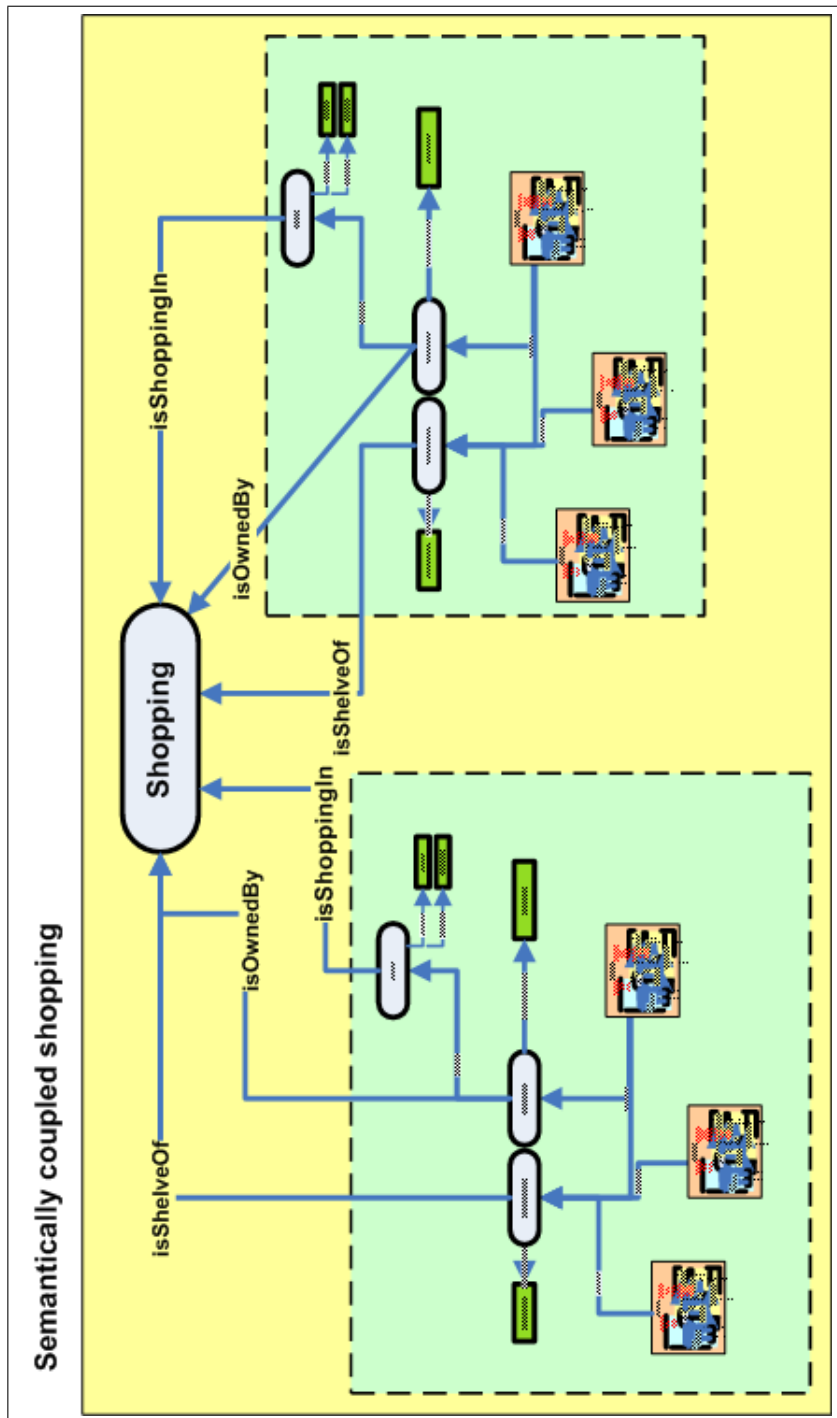
Afbeeldingen



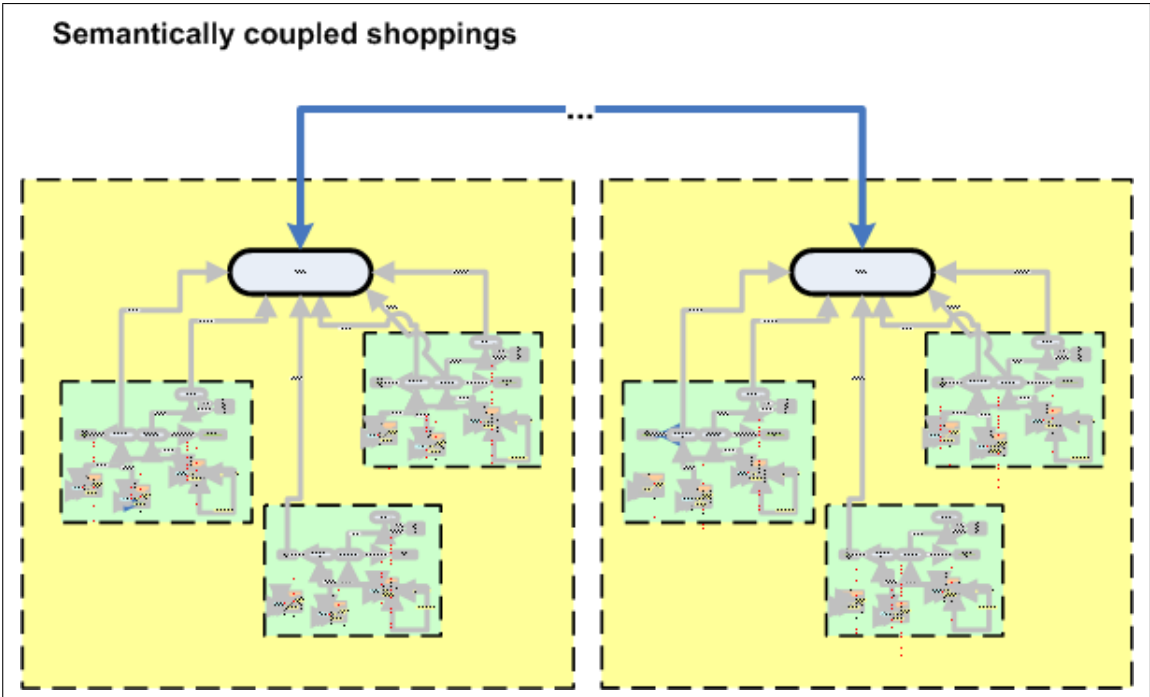
Figuur A.1: Graafmodel een semantische beschrijving van een product



Figuur A.2: Graafmodel van een beschrijving van semantisch gekoppelde producten



Figuur A.3: Graafmodel van een semantisch beschreven shopping



Figuur A.4: Graafmodel van semantisch gekoppelde shoppings

Bijlage B

Afkortingen

- AJAX : Asynchronous JavaScript and XML
- API : Application Programming Interface
- CSS : Cascading Style Sheets
- DAML : DARPA Agent Markup Language
- FOAF : Friend of a Friend
- FTP : File Transport Protocol
- HTTP : Hypertext Transfer Protocol
- lib : library = bibliotheek
- MIME : Multi-purpose Internet Mail Extensions
- N3 : Notation 3
- OWL : Web Ontology Language
- OIL : Ontology Inference Layer
- PDF : Portable Document Format
- PHP : PHP Hypertext Preprocessor
- SAWSDL : Semantic Annotations for WSDL
- STalkingObjects : Semantic Talking Objects
- RAP : RDF API for PHP
- RDF : Resource Description Framework
- RDFS : RDF Schema
- RDQL : RDF Data Query Language
- REST : Representational State Transfer

- RF : Radio Frequency
- RFID : Radio Frequency Identification
- RSS : Rich Site Summary / Really Simple Syndication
- SOAP : Simple Object Access Protocol
- SPARQL : SPARQL Protocol and RDF Query Language
- UDDI : Universal Description, Discovery and Integration
- URI : Uniform Resource Identifier
- URL : Uniform Resource Locator
- W3C : World Wide Web Consortium
- WiFi : Wireless Fidelity
- WSDL : Web Service Description Language
- WSMO : Web Service Modeling Ontology
- WYSIWYG : What you see is what you get
- WWW : World Wide Web
- XHTML : Extensible Hypertext Markup Language
- XML : Extensible Markup Language

Bijlage C

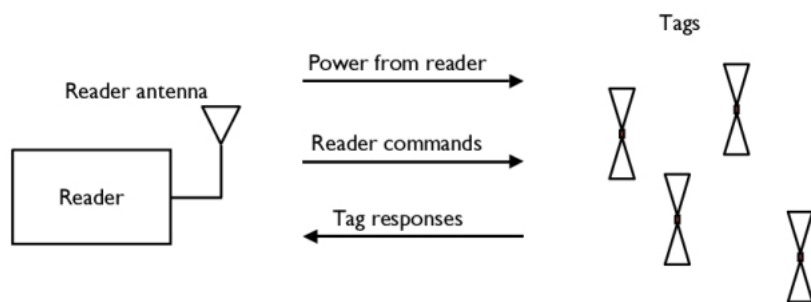
Extra's

C.1 Radio Frequency Identification

In deze sectie bekijken we de grondprincipes en standaarden achter *Radio Frequency Identification* in meer detail. Dit kan gezien worden als een aanvulling op hoofdstuk 3.2 die noodzakelijk was om latere veronderstellingen te maken.

C.1.1 Grondprincipes

RFID technologie staat toe dat er communicatie plaats kan vinden tussen een lezer en bepaalde tags. In volgende secties wordt de werking van passieve tags uit de doeken gedaan, vermits hier gebruik van gemaakt wordt in deze thesis. Figuur C.1¹ geeft een eenvoudig overzicht van de werking van passieve RFID-tags. Zoals reeds eerder vermeld gebruiken de passieve tags de energie die afkomstig is van de RF-lezer, meerbepaald de antenne van deze lezer. Vervolgens start deze RF-lezer de communicatie en wacht de lezer op een antwoord van de passieve RFID-tags.



Figuur C.1: Eenvoudig overzicht op de werking van passieve RFID-tags

Er bestaan een heleboel verschillende grondbeginselen die voor de werking van RFID systemen zorgen.

¹naar <http://www.rfid-handbook.de>

De twee belangrijkste principes worden in de volgende secties uit de doeken gedaan.

Inductieve koppeling

Inductieve koppeling (ook wel magnetische koppeling) is het overdragen van energie via het elektrische-inductie-fenomeen. Dit fenomeen wordt gebruikt bij passieve RFID-tags. In elektronische schakelingen (zoals die van een RF-lezer) kan resonantie² worden bereikt door een spoel en een condensator met elkaar te verbinden, dit wordt een LC-kring genoemd. RFID technologie is gebaseerd op zulke LC-kringen die op een bepaalde resonantie frequentie door de RF-lezer worden afgesteld (fR).

Omdat de golflengte van de gegeneerde LC-kringen enkele keren groter is dan de afstand tussen de antenne van de RF-lezer en die van de RFID-tag, kan het elektro-magnetisch veld gezien worden als een 'veranderend magnetisch veld' tussen beide antennes. Hierdoor vindt er inductie³ plaats wanneer een RFID-tag het uitgezonden magnetisch veld binnendringt. Deze energie kan uiteindelijk gebruikt worden als energietoevoer voor de microchip van de RFID-tag.

Backscatter koppeling De RFID-tags maken tevens gebruik van het principe van *backscattering* om informatie over te brengen naar de RF-lezer. *Backscatter* (letterlijk vertaalt 'terug verspreiden') staat voor het proces waar (radio)golven of signalen worden gereflecteerd naar hun oorspronkelijke richting, in dit geval de positie van de RF-lezer.

Inzicht hebben in de maximale afstand die de identificatietechnologie RFID ons kan bieden, is belangrijk wanneer we deze technologie in een uiteindelijke implementatie willen gebruiken (later in deze thesis). In [31, 7] wordt er na gegaan wat de theoretische grenzen van RFID technologie. In deze studies gaat men na wat de afstand is die RFID tags kunnen overbruggen. Men besluit dat passieve RFID tags tot 4 meter theoretisch zouden kunnen overbruggen. Hieruit kunnen we besluiten dat RFID technologie bruikbaar is in bijvoorbeeld een winkel, waar afstanden meestal binnen dit bereik blijven.

De RFID-technologie moet zich net zoals andere technologieën houden aan zogenaamde EMC-vereisten (ElektroMagnetische compatibiliteit vereisten). Interferentie ontstaat doordat elk elektrisch of elektronisch product elektromagnetische straling uitzendt. Andere producten kunnen deze straling dan ontvangen, waardoor ongewenste effecten in het functioneren van dat product optreden.

EMC⁴ is het vakgebied dat elektromagnetische interferentie in en tussen elektrische en elektronische producten en systemen voorkomt en bestrijdt. RFID-technologie wordt op dit vlak dus wel tegengehouden door menselijke grenzen, in praktijk zal daarom de afstand variëren van 10cm tot 2 meter.

C.1.2 Standaarden

Standaarden zijn noodzakelijk om de communicatie tussen een RF-lezer en een RFID-tag binnen enkele vastgelegde protocollen te laten verlopen. Wanneer RFID-tags gebruikt worden in een open keten (zoals het framework in deze thesis beschreven) is het van groot belang dat alle tags dezelfde 'taal' spreken.

Door een wildgroei aan standaarden in de beginjaren van de RFID technologie, is de situatie tegenwoordig dat er veel standaarden bestaan die vaak niet compatibel zijn. De gebruikte standaarden kan men

²Resonantie is een natuurkundig verschijnsel dat voorkomt bij trillingen. Een trillend voorwerp zal bij een ander voorwerp resonantie teweegbrengen, als dit voorwerp met de trillingen gaat meetrillen, sterker dan men op grond van de aanstoting zou verwachten

³Inductie is het natuurkundig verschijnsel waarbij in een geleider elektrische spanning wordt opgewekt wanneer de geleider zich bevindt in een veranderend magnetisch veld of wanneer een geleider beweegt in een magnetisch veld

⁴http://nl.wikipedia.org/wiki/Elektromagnetische_compatibiliteit

opdelen in drie grote categorieën nl. technologiestandaarden en datastructuurstandaarden en applicatiestandaarden.

Technologiestandaarden De technologiestandaarden leggen de grote lijnen vast van de technische specificaties waarbinnen RFID technologie zich verder kan ontwikkelen. De belangrijkste standaard op dit gebied is de ISO⁵ 18000 standaard. Deze standaard beschrijft een communicatieprotocol voor elke specifieke frequentieband. ISO 18000 probeert te werken aan een standaard waarin de verschillende frequenties onderling zo weinig mogelijk interferentie hebben en men naadloos kan overschakelen van één frequentie naar een andere.

Datastructuurstandaarden

De datastructuurstandaarden beschrijven verschillende aspecten van de gegevensorganisatie en zijn niet afhankelijk van de gebruikte technologie. De ISO/EIC-standaarden zijn op dit vlak de meest belangrijke standaarden. Een kort overzicht van enkele bekendere EIC-substandaarden :

- ISO/EIC 15962 biedt een protocol aan met algemeen gegevensformaat en commandostructuur, foutafhandeling, collisiondetection procedures,...
- ISO/EIC 15961 bepaalt gemeenschappelijke functionele commando's en karakteristieken (gebruikte type RFID-tag, formaat van gegevensopslag...)
- ISO/EIC 15963 garandeert dat elke RFID-tag een uniek nummer bevat. Deze standaard bevat een specifiek nummeringssysteem en een heleboel registratieprocedures.

Toch zijn deze datastructuurstandaarden nog lang niet gemeengoed en worden er veel verschillende standaarden door mekaar gebruikt.

Applicatiestandaarden De applicatiestandaarden beogen het gebruik van RFID bepaalde toepassingen eenvormig te maken door een vooraf bepaalde technische oplossing voor te stellen. De Electronic Product Code (EPC⁶) streeft naar een unieke identificatie voor elk uniek product. Deze standaard heeft vooral betrekking op de inhoud van de gegevens op RFID-tags. Tevens probeert de EPC-standaard vroegere standaarden (onder andere de EAN/UCC-standaard die bij barcodes gebruikt wordt) te integreren.

C.2 Fysiek-Digitaal problemen in winkelomgevingen

Dezer dagen worden winkels geteisterd door veel problemen die een oorzakelijk verband hebben met de alomtegenwoordige digitale kloof (beschreven in sectie C.2).

Deze problemen spelen zich in twee grote gebieden af :

1. Consumentenproblemen
2. Bedrijfsgerichte problemen

In de volgende secties worden twee scenario's beschreven die deze problemen uit de doeken doen.

⁵International Organization for Standardization

⁶<http://www.epcglobalinc.org/home>

C.2.1 Scenario : consumentenproblemen

Petra, een voltijds-werkende moeder met 2 kinderen, gaat wekelijks vergezeld van haar familie naar de supermarkt waar ze een heleboel verschillende goederen koopt. Daar kopen ze goederen gaande van voeding en drank voor de familie, speelgoed voor de kinderen, ICT-gerelateerde artikelen, ... Jan, de man van Petra, koopt meestal ICT-gerelateerde producten nadat hij op internet thuis heeft gekeken wat andere mensen over de producten denken. Jan heeft daarenboven vaak problemen met het terugvinden van een bepaald artikel. Deze week zocht hij zich dood naar een USB-kabel voor zijn nieuwe printer. . .

Petra heeft vaak moeite om zich binnen het budget van de familie te houden, vermits ze pas aan de kassa het totaalbedrag te zien krijgt. Hierdoor wordt het winkelen een onaangename bezigheid vermits de kinderen ook graag speelgoed willen en ze continu moet bijhouden of ze binnen het budget blijft. Daarenboven hebben beiden weinig zin om veel tijd te spenderen aan boodschappen doen.

Doordat ze wordt afgeleid door zoveel dingen tegelijk, ontgaan de promoties haar vaak. Petra maakt ook elke week een lijstje van producten die ze zeker moet aankopen. Jammer genoeg merkt ze maar al te vaak dat ze producten vergeten is wanneer ze met haar overvolle winkelkar aan de kassa wacht. Vermits de overgrote meerderheid van de mensen winkelen op hetzelfde tijdstip als Petra en haar familie, verliest ze vaak veel tijd in de wachtrij van de kassa.



Figuur C.2: Enkele alomtegenwoordige consument-gerelateerde problemen in de supermarkt.

Het beschreven scenario toont meteen een heleboel problemen die we momenteel in supermarkten vinden. Goede oplossingen voor deze problemen zouden voor de consument veel verschillende voordelen met zich meebrengen nl.

- Tijdswinst (geen of kortere wachtrijen)

- Beter budgetbeheer
- Gemakkelijker toegang tot informatie van producten
- Meer aandacht voor secundaire zaken (promoties, boodschappenlijstjes, . . .)

C.2.2 Scenario : bedrijfsgerichte problemen

Bart, het hoofd van de supermarkt, heeft een heleboel problemen die niet meteen zichtbaar zijn voor de consument en het personeel, waar hij geen directe oplossing voor kan bieden. Een eerste probleem bestaat erin dat consumenten soms een aankoop overwegen, en uiteindelijk het product terug leggen op de verkeerde plaats. Doordat de consument afgeschrikt wordt om producten te kopen die niet op de juiste plaats staan, zijn deze producten nutteloos en worden ze bijgevolg niet verkocht. Het opsporen van deze verloren gelegde producten is een hels werk voor de winkelbedienden vermits deze nog een heleboel andere taken hebben.

In de winkel wordt er ook vaak om versterking geroepen om de lange wachtrijen aan de kassa te verminderen. Daarenboven worden sommige consumenten door deze wachttijden afgeschrikt om naar de supermarkt te gaan wanneer ze maar enkele producten willen kopen.

Het beheer van de stock is een enorm tijdrovend werk waar veel fouten in optreden. Het grootste deel van de tijd is men bezig met het verifiëren van digitale cijfers met de werkelijkheid in de winkel. Eveneens is de boekhouding van de supermarkt een afschuwelijk karwei. Dit is te wijten aan meerdere factoren : producten die aan de kassa niet juist worden aangerekend, diefstal van producten die niet beveiligd zijn, promoties die verkeerd worden toegekend, . . .

En zelfs als dit nog niet genoeg is, moeten er dagelijks etiketten gedrukt met nieuwe prijzen. Deze etiketten moeten vervolgens verspreid worden over heel de winkel. Dit kost enorm veel tijd voor de winkelbediendes. In dit proces gebeuren er overigens regelmatig fouten, waarbij verkeerde prijzen zijn gedrukt of etiketten op de verkeerde plaats worden toegekend.

Ook op bedrijfsvlak zijn er blijkbaar veel problemen die voor een groot deel te wijten zijn aan de digitale kloof. Het oplossen van deze problemen zou eveneens een aantal voordelen met zich meebrengen :

- Winkelbediendes krijgen meer tijd voor andere primaire zaken die tegenwoordig secundair zijn geworden. Een voorbeeld hiervan is het persoonlijke contact met de consument. Tegenwoordig zijn winkelbediendes meer bezig met de problemen die hierboven beschreven zijn waardoor elk persoonlijk contact verdwijnt.

Dit is een duidelijk voorbeeld van een situatie waarin digitale technologie kan leiden tot meer persoonlijk contact, dit in tegenstelling met wat vaak wordt aangenomen.

- Consumenten vinden het aangener om te winkelen vermits artikels meestal op de juiste plaats staan en de prijzen up-to-date zijn.
- De digitale cijfers van de stock komen beter overeen met die van de fysieke wereld. Hierdoor zal ook de boekhouding beter gaan kloppen.

Reeds in april 2004 rapporteerde *VRTNieuws.NET*⁷ het volgende : 'De Belgische supermarkten willen meer open zijn'. Toch werd dit voorstel meteen van de kaart geveegd vermits een supermarkt 's avonds

⁷<http://www.vrtnieuws.net>



Figuur C.3: Bedrijfsgerelateerde problemen in de supermarkt.

openhouden veel personeel vereist. Naar aanleiding van dit nieuws kwamen er veel reacties in de trend van :

... Wie gaat de rekken vullen, administratie bijhouden, kassa's bedienen, klanten helpen, orders doorvoeren,...? Minder personeel impliceert minder service, minder service dan weer minder klanttevredenheid wat dan weer leegloop van het klantenbestand met zich meebrengt!...

Hierin zou nieuwe technologie een enorme hulp kunnen zijn, door secundaire taken te vergemakkelijken zou service op een evenwaardig (en zelfs hoger) niveau kunnen bestaan en bijgevolg ook de klanttevredenheid.

Bibliografie

- [1] Robert Adelman, Marc Langheinrich, and Christian Floerkemeier. A toolkit for bar-code-recognition and -resolving on camera phones – jump starting the internet of things. Workshop Mobile and Embedded Interactive Systems (MEIS'06) at Informatik 2006, October 2006.
- [2] ADM, Cisco Systems, Piper Rudnick, and Gray Cary. 'rfid, klaar voor de grote doorbraak?'. 2005.
- [3] Jürgen Bohn and Friedemann Mattern. Super-Distributed RFID Tag Infrastructures. In Panos Markopoulos, Berry Eggen, Emile Aarts, and James Crowley, editors, *Proc. 2nd European Symposium on Ambient Intelligence (EUSAI 2004)*, number 3295 in LNCS, pages 1–12, Eindhoven, The Netherlands, November 2004. Springer-Verlag.
- [4] Gregor Broll, Sven Siorpaes, Enrico Rukzio, Massimo Paolucci, John Hamard, Matthias Wagner, and Albrecht Schmidt. Supporting service interaction in the real world. *percom*, 0, 2006.
- [5] Gregor Broll, Sven Siorpaes, Enrico Rukzio, Massimo Paolucci, John Hamard, Matthias Wagner, and Albrecht Schmidt. Supporting mobile service usage through physical mobile interaction. *percom*, 0:262–271, 2007.
- [6] George Coulouris, Jean Dollimore, and Tim Kindberg. *Distributed systems (3rd edition): concepts and design*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.
- [7] NCR Retail Solutions Division. Physics of backscatter tags. 2005.
- [8] Marias Giannis F., Papazafeiropoulos Giorgos, Priggouris Nikos, Hadjiefthymiades Stathes, and Merakos Lazaros. An innovative gateway for indoor positioning. *EURASIP Journal on Applied Signal Processing*, 2006:Article ID 81714, 10 pages, 2006.
- [9] Schmid B. F. and Schmid-Isler. Konzepte von beat f. schmid ein Überblick. page 6, 2004.
- [10] A. Ferscha, M. Hechinger, R. Mayrhofer, M. dos Santos Rocha, M. Franz, and R. Oberhauser. Digital aura.
- [11] Kenneth P. Fishkin, Thomas P. Moran, and Beverly L. Harrison. Embodied user interfaces: Towards invisible user interfaces. In *Proceedings of the IFIP TC2/TC13 WG2.7/WG13.4 Seventh Working Conference on Engineering for Human-Computer Interaction*, pages 1–18, Deventer, The Netherlands, The Netherlands, 1999. Kluwer, B.V.
- [12] Armando Fox, Nigel Davies, Eyal de Lara, Mirjana Spasojevic, and William Griswold. Real-world ubicomp deployments: Lessons learned. *IEEE Pervasive Computing*, 5(3):21–23, 2006.
- [13] Gatenby and David Arthur Gray. Galatea : personalized interaction with augmented objects. Massachusetts, USA, 2005. MIT.

- [14] Neil Gershenfeld, Raffi Krikorian, and Danny Cohen. The internet of things. *Scientific American*, 291(4):76–81, October 2004.
- [15] KD Gottschalk, S Graham, H Kreger, and J Snell. Introduction to web services architecture. *IBM Systems Journal*, 2001.
- [16] Beverly L. Harrison, Kenneth P. Fishkin, Anuj Gujar, Carlos Mochon, and Roy Want. Squeeze me, hold me, tilt me! an exploration of manipulative user interfaces. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 17–24, New York, NY, USA, 1998. ACM Press/Addison-Wesley Publishing Co.
- [17] Andreas Heß, Eddie Johnston, and Nicholas Kushmerick. Assam: A tool for semi-automatically annotating semantic web services (draft). In *Proceedings of the 3rd International Semantic Web Conference (ISWC 2004)*, Hiroshima, Japan, 2004.
- [18] Ishii Hiroshi and Ullmer Brygg. Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 234–241, New York, NY, USA, 1997. ACM Press.
- [19] Lars Erik Holmquist, Friedemann Mattern, Bernt Schiele, Petteri Alahuhta, Michael Beigl, and Hans-Werner Gellersen. Smart-its friends: A technique for users to easily establish connections between smart artefacts. In *UbiComp '01: Proceedings of the 3rd international conference on Ubiquitous Computing*, pages 116–122, London, UK, 2001. Springer-Verlag.
- [20] O. Kasten and M. Langheinrich. First experiences with bluetooth in the smart-its distributed sensor network, 2001.
- [21] Cory D. Kidd, Robert Orr, Gregory D. Abowd, Christopher G. Atkeson, Irfan A. Essa, Blair MacIntyre, Elizabeth Mynatt, Thad E. Starner, and Wendy Newstetter. ‘easyliving: Ubiquitous computing & the role of geometry’. 2000.
- [22] Cory D. Kidd, Robert Orr, Gregory D. Abowd, Christopher G. Atkeson, Irfan A. Essa, Blair MacIntyre, Elizabeth D. Mynatt, Thad Starner, and Wendy Newstetter. The aware home: A living laboratory for ubiquitous computing research. In *CoBuild '99: Proceedings of the Second International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture*, pages 191–198, London, UK, 1999. Springer-Verlag.
- [23] Naohiko Kohtake, Kenta Matsumiya, Kazunori Takashio, and Hideyuki Tokuda. Smart device collaboration for ubiquitous computing environment. *Fifth International Conference on Ubiquitous Computing (UbiComp2003)Worksho on Multi-Device Interface for Ubiquitous Peripheral Interaction*, 10 2003.
- [24] Shin’ichi Konomi and George Roussos. Ubiquitous computing in the real world: lessons learnt from large scale rfid deployments. *Springer-Verlag London Limited*, 2006.
- [25] Matthias Kranz and Albrecht Schmidt. Prototyping smart objects for ubiquitous computing. In *In Proceedings of the International Workshop on Smart Object Systems in Conjunction with the Seventh International Conference on Ubiquitous Computing*, September 2005.
- [26] Yves Lafon. Web Services Activity Statement. <http://www.w3.org/2002/ws/Activity>, 2002.
- [27] Weiser Mark. The computer for the 21st century. *Scientific American*, September 1991.
- [28] B M Matthews. Integration via meaning: Using the semantic web to deliver web services. pages Article ID 81714, 10 pages, 2005.

- [29] Negroponte Nicholas. *Being Digital*. January 1995.
- [30] Thomas Nicolai, Florian Resatsch, and Daniel Michelis. The web of augmented physical objects. In *ICMB '05: Proceedings of the International Conference on Mobile Business (ICMB'05)*, pages 340–346, Washington, DC, USA, 2005. IEEE Computer Society.
- [31] Pavel V. Nikitin and K. V. S. Rao. Theory and measurement of backscattering from rfid tags. 2005.
- [32] Ernest H. Page, Arnold Buss, Paul A. Fishwick, Kevin J. Healy, Richard E. Nance, and Ray J. Paul. Web-based simulation: revolution or evolution? *ACM Trans. Model. Comput. Simul.*, 10(1):3–17, 2000.
- [33] Aabhas V Paliwal, Nabil Adam, and Christof Bornhovd. Semantic discovery and composition of web services for rfid applications in border control. In *3rd International Semantic Web Conference, ISWC'04 - Hiroshima, Japan, 2004*.
- [34] Massimo Paolucci, Katia Sycara, and Takahiro Kawamura. Delivering semantic web services. Technical Report CMU-RI-TR-02-32, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, May 2003.
- [35] Ravi Pappu and ThingMacig LLC. ‘the physics of rfid’. 2003.
- [36] Yvonne Rogers. Moving on from weiser’s vision of calm computing: Engaging ubicomp experiences. *Springer-Verlag Berlin Heidelberg*, pages 404–421, 2006.
- [37] M. Rohs. Real-world interaction with camera-phones, 2004.
- [38] Albrecht Schmidt, Frank Siegemund, Michael Beigl, Stavros Antifakos, Florian Michahelles, and Hans-Werner Gellersen. Mobile ad-hoc communication issues in ubiquitous computing – the smart-its experimentation platforms. In *Proc. 8th Intl. Conference on Personal Wireless Communication (PWC 2003)*, Springer-Verlag, LNCS 2775, pages 213–218, Venice, Italy, September 2003.
- [39] M. Sheshagiri, N. Sadeh, and F. Gandon. Using semantic web services for context-aware mobile applications. In *MobiSys 2004 Workshop on Context Awareness*, 2004.
- [40] Ali A. Nazari Shirehjini and Felix Klar. 3dsim: rapid prototyping ambient intelligence. In *sOc-EUSAI '05: Proceedings of the 2005 joint conference on Smart objects and ambient intelligence*, pages 303–307, New York, NY, USA, 2005. ACM Press.
- [41] Sven Siorpaes, Gregor Broll, and Massimo Paolucci. Mobile interaction with the internet of things. In *Embedded Interaction Research Group*, 2004.
- [42] Evren Sirin, Bijan Parsia, and James Hendler. Filtering and selecting semantic web services with interactive composition techniques. *IEEE Intelligent Systems*, 19(4):42–49, 2004.
- [43] Kern Steffen, Braun Peter, and Rossak Wilhelm. *MobiSoft: An Agent-Based Middleware for Social-Mobile Applications*. 2006.
- [44] Martin R. Stytz, Terry Adams, Brian Garcia, Steven M. Sheasby, and Brian Zurita. Rapid prototyping for distributed virtual environments. *IEEE Softw.*, 14(5):83–92, 1997.
- [45] Katia P. Sycara, Massimo Paolucci, Anupriya Ankolekar, and Naveen Srinivasan. Automated discovery, interaction and composition of semantic web services. *J. Web Sem.*, 1(1):27–46, 2003.
- [46] Kristof Thys, Ruben Thys, Kris Luyten, and Karin Coninx. Photofoaf: A community building service driven by socially-aware mobile imaging. *smap*, 0:97–102, 2006.

-
- [47] Jo Vermeulen, Kris Luyten, and Karin Coninx. Tangible mashups: Exploiting links between the physical and virtual world., 2007.
- [48] Roy Want, Kenneth P. Fishkin, Anuj Gujar, and Beverly L. Harrison. Bridging physical and virtual worlds with electronic tags. In *CHI*, pages 370–377, 1999.
- [49] Mark Weiser. Ubiquitous computing (abstract). In *CSC '94: Proceedings of the 22nd annual ACM computer science conference on Scaling up : meeting the challenge of complexity in real-world computing applications*, page 418, New York, NY, USA, 1994. ACM Press.
- [50] Mark Weiser and John Seely Brown. The coming age of calm technology. pages 75–85, 1997.
- [51] Pierre Wellner. The digitaldesk calculator: tangible manipulation on a desk top display. In *UIST '91: Proceedings of the 4th annual ACM symposium on User interface software and technology*, pages 27–33, New York, NY, USA, 1991. ACM Press.
- [52] Pierre Wellner. Interacting with paper on the digitaldesk. *Commun. ACM*, 36(7):87–96, 1993.
- [53] Pierre Wellner, Wendy Mackay, and Rich Gold. Back to the real world. *Commun. ACM*, 36(7):24–27, 1993.
- [54] Bao Xu and Wang Gang. Random sampling algorithm in rfid indoor location system. In *DELTA '06: Proceedings of the Third IEEE International Workshop on Electronic Design, Test and Applications*, pages 168–176, Washington, DC, USA, 2006. IEEE Computer Society.
- [55] Guang yao Jin, Xiao yi Lu, and Myong-Soon Park. An indoor localization mechanism using active rfid tag. In *SUTC (1)*, pages 40–43, 2006.
- [56] Vasileios Zeimpekis, George M. Giaglis, and George Lekakos. A taxonomy of indoor and outdoor positioning techniques for mobile location services. *SIGecom Exch.*, 3(4):19–27, 2003.