

Samenvatting

Met het opkomen van Web 2.0, zijn de verschillende instituten en organisaties begonnen hun eertijds geïsoleerde en heterogene collecties open te stellen ten einde het publiek optimaal toegang te verschaffen. Semantische Web technologieën zijn daarbij instrumenteel voor de integratie van rijke collecties van metadata en het definiëren van ontologieën die de variëteit aan representatie schemata omvatten en inconsistenties tussen de verschillende vocabulaires oplossen. Met het oog op de grote verscheidenheid aan metadata met een complexe semantische structuur is het van groot belang een zorgvuldige selectie te maken van de informatie die aan bezoekers wordt gepresenteerd. In deze context werkt het CHIP¹ project (Cultural Heritage Information Personalization), vanaf 2005, gesponsord door NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek), als onderdeel van het Nederlandse CATCH² programma (Continuous Access to Cultural Heritage). CHIP is een samenwerkingsproject, met als partners het Rijksmuseum Amsterdam³, de Technische Universiteit Eindhoven⁴ en het Telematica Instituut⁵.

De probleemstelling die aan het onderzoek dat in dit proefschrift beschreven wordt ten grondslag ligt luidt als volgt: *Kunnen we bezoekers ondersteunen met gepersonaliseerde toegang tot de semantisch verrijkte collecties?* Ten einde deze vraag te bestuderen en te beantwoorden hebben we gekozen voor cultureel erfgoed (en in het bijzonder musea) als domein van toepassing, en dient de semantisch rijke achtergrondkennis van de museum collecties als uitgangspunt voor ons onderzoek. Met die kennis als basis, hebben we daarenboven gebruik gemaakt van gebruikersmodellerings- en aanbevelings-technieken teneinde gepersonaliseerde diensten te kunnen ontwikkelen voor museumbezoekers. De belangrijkste onderzoeksbijdragen van dit proefschrift zijn, samengevat: (i) we hebben een interactieve *rating*-dialoog voor kunstwerken en concepten ontwikkeld, om snel een CHIP gebruikersprofiel te kunnen instantiëren, dat gerealiseerd is als een specializatie van FOAF⁶ en tevens afgebeeld kan worden op een bestaande *event model* ontologie SEM⁷; (ii) we hebben een hybride aanbevelingsalgoritme

¹www.chip-project.org/

²www.nwo.nl/catch

³www.rijksmuseum.nl/

⁴w3.tue.nl/

⁵www.novay.nl/en/

⁶www.foaf-project.org/

⁷semanticweb.cs.vu.nl/2009/11/sem/

voorgesteld, waarin zowel expliciete als impliciete relaties die bestaan in de semantische structuur van de collectie gecombineerd worden. Ten behoeve van de presentatie, voor de eind-gebruikers, hebben we drie *tools* ontwikkeld: *Art Recommender*, *Tour Wizard* en de *Mobile Tour Guide*. Met een *user-centered* ontwerp cyclus als leidraad, hebben we een reeks evaluaties met museum bezoekers uitgevoerd om de effectiviteit van aanbevelingen die voortkomen uit de *rating*-dialoog te testen, alsook de verschillende manieren waarop een gebruikers-profiel (*user model*) geconstrueerd kan worden en de voorspellings-accuraatheid van het hybride aanbevelingsalgoritme.

In hoofdstuk 1 worden de onderzoeksvragen en onze aanpak besproken, alsook de indeling van dit proefschrift.

In hoofdstuk 2 wordt een overzicht gegeven van het werk dat ter voorbereiding dient van het latere werk. Dit overzicht omvat (i) de semantische verrijking van de Rijksmuseum collectie, en de afbeelding daarvan op drie Getty vocabulaires⁸ (ULAN, AAT, TGN) en de Iconclass thesaurus⁹; (ii) de minimale gebruikersprofiel (*user model*) ontologie, gedefiniëerd als een specialisatie van FOAF, waarin in eerste instantie alleen *ratings* opgeslagen worden; (iii) de eerste implementatie van het inhoud-gebaseerde (*content-based*) aanbevelingsalgoritme van de CHIP *Art Recommender*.

In hoofdstuk 3 worden de twee andere tools gepresenteerd: de *Tour Wizard* en de *Mobile Tour Guide*. Op basis van gebruikers-*ratings*, geeft de web-gebaseerde *Tour Wizard* aanbevelingen over mogelijke museum *tours* bestaande uit aanbevolen kunstwerken, die momenteel voorhanden zijn in de museum expositieruimte. De *Mobile Tour Guide* vertaalt deze aanbevelingen naar een formaat geschikt voor mobiele apparaten (bv. een PDA) die gebruikt kunnen worden in de (fysieke) museum ruimte. Om te kunnen reageren op interactie door de gebruikers hebben we een vertaling gemaakt van het gebruikersprofiel zoals opgeslagen in RDF naar een XML formaat dat leesbaar is voor het mobiele apparaat, zodat een dynamische synchronisatie mogelijk is tussen de *online* en *on-site* gebruikersprofielen.

In hoofdstuk 4 wordt de tweede generatie van de *Mobile Tour Guide* gepresenteerd, waarin een *real time* routingssysteem is opgenomen, geschikt voor verschillende apparaten (waaronder de iPod). In vergelijking met de eerste generatie, kan deze versie museum *tours* aanpassen op basis van de *rating* van kunstwerken en concepten door de gebruiker, de locatie waar de gebruiker zich in het (fysieke) museum bevindt en de coördinaten van de kunstwerken en tentoonstellingszalen in het museum. Bovendien ondersteunen we in deze versie een afbeelding van het CHIP gebruikersprofiel naar een bestaande *event model* ontologie SEM. Behalve *ratings* laat dit model toe ook andere gebruikersactiviteiten op te slaan, zoals het volgen van een tour en het bekijken van kunstwerken.

In hoofdstuk 5 worden een aantal semantische relaties geïdentificeerd, zowel binnen een vocabulaire (bv. een concept heeft een generaliserend/specialiserend concept) als over de grenzen van vocabulaires heen (bv. een kunstenaar is geassocieerd met een kunststijl). We hebben deze relaties in combinatie met ele-

⁸www.getty.edu/research/conducting_research/vocabularies/

⁹www.Iconclass.nl/libertas/ic?style=index.xml

mentaire kunstwerk eigenschappen toegepast in inhoud-gebaseerde aanbevelingen (*recommendations*) en de verschillende relaties met elkaar vergeleken in termen van bruikbaarheid (*usefulness*). Dit onderzoek geeft ons verder mede zicht op het gebruik van kenmerken van kunstwerken in combinatie met meervoudige semantische relaties, en de afleiding van typische patronen in gebruikersnavigatie.

In hoofdstuk 6 wordt de taak van gepersonaliseerde aanbevelingen nauwkeurig gedefiniëerd en deze taak vervolgens opgesplitst in een aantal inferentiestappen gegeneraliseerd naar ontologie-gebaseerde aanbevelingssystemen, vanuit een perspectief van *knowledge engineering*. We hebben een hybride aanpak voorgesteld, waarin zowel expliciete als impliciete aanbevelingen gecombineerd worden. Tot de expliciete relaties behoren kenmerken van kunstwerken alsook semantische relaties met initiële gewichten, afgeleid van de evaluatie beschreven in hoofdstuk 5. De impliciete relaties hebben betrekking op (kunst) concepten en zijn verkregen middels *instance-based ontology matching*.

In hoofdstuk 7 wordt een voorbeeld gegeven van het hergebruik van de uit een applicatie verkregen gebruikers interactie data in een andere applicatie, om op die wijze *cross-application* aanbevelingen te kunnen geven. In dit voorbeeld worden gebruikers *tags* over culturele *events*, verzameld door iCITY¹⁰, gebruikt om het gebruikersprofiel van de *CHIP Art recommender* te verrijken. Om een volledige *tagging* interoperabiliteit te realiseren, hebben we de problemen die zich voordoen bij de afbeelding van gebruikers *tags* naar domein ontologieën nader onderzocht, en additionele mechanismes voorgesteld, zoals het gebruik van SKOS *matching* operatoren, om mogelijke *mis-alignments* tussen *tags* en domein-specifieke ontologieën op te vangen.

In hoofdstuk 8 is een samenvattende bespreking van de oorspronkelijke probleemstellingen opgenomen en de mate waarin de afzonderse onderzoeksvragen beantwoord zijn.

¹⁰iCITY.di.unito.it/dsa-dev/