

Visie document

Versie 2.0

1 juli 2011

Auteurs

Hennie Brugman, technisch coördator CATCHPlus
hennie.brugman@meertens.knaw.nl

Doel is, dat dit document uiteindelijk een visie formuleert, waar de volgende partijen achter kunnen staan:

- CATCHPlus projectbureau (opdrachtgever)
- Gridline (uitvoerder)
- Beeld en Geluid (gebruiker UPR, opdrachtgever ZieOok. Namens deze Siem Vaessen)
- Amsterdam Museum (gebruiker UPR en ZieOok)

Projectgegevens

Projectmanagement: Gridline

Ontwikkelmethode: Agile (Gridline gebruikt een eigen variant van scrum). Hennie Brugman kan daarbij voor het User Profile Repository gedeelte als vertegenwoordiger van de opdrachtgever aan (iteration planning en review) meetings deelnemen. Siem Vaessen representeert daarbij de andere gebruiker van de te ontwikkelen dienst, Beeld en Geluid.
Met betrekking tot de Art Recommender neemt NN namens het Amsterdam Museum deel.

Stuurgroep: bovengenoemde partijen formeren samen een stuurgroep. Taken en bevoegdheden:

- ...
-

Aanvangsdatum: 15 juli 2011?

Streefdatum laatste release: 1 januari 2012.

Wat willen we bouwen?

1. User Profile Repository

De User Profile Repository slaat ten behoeve van gebruikers *gebruikersprofielen* op: naast persoonskenmerken kan het ook beschrijvingen van en verwijzingen naar voorkeursobjecten en voorkeursconcepten betreffen. Per voorkeursobject of – concept kan een persoonlijke rating worden opgeslagen. De UPR kan deze ratings per object ook geaggregeerd terugleveren. Daarnaast kan de UPR opslaan hoe vaak een gebruiker een bepaald object heeft bekeken of afgespeeld (statistics).

De UPR biedt eenvoudig toegang aan collectie-aanbieders en dienstverleners door middel van REST interfaces. Voor deze toegang is wel authenticatie nodig.

Eigenaren van gebruikersprofielen hebben ook interactief toegang tot hun eigen profielinformatie door middel van een in het kader van dit project te bouwen webapplicatie. Deze webapplicatie heeft, net als andere diensten, toegang tot gebruiksprofielen via de REST interfaces.

Tenslotte omvat het UPR project een tweetal widgets, die door website bouwers kunnen worden toegepast. Het ene widget maakt het mogelijk je vanaf een willekeurige webpagina als gebruiker bij de UPR te authenticeren, het andere stelt gebruikers in staat persoonlijke ratings aan objecten op web-pagina's te koppelen ("0 tot 5 sterretjes").

Samenvattend: het UPR project omvat de volgende onderdelen:

- UPR repository / database
- UPR REST service, met functionele groepen
 - o Aanmaken en beheren van gebruikers en profielen
 - o Toevoegen en beheren van voorkeursobjecten, en -concepten (beschrijvingen en verwijzingen), ratings en gebruiks-statistieken
 - o Toevoegen en beheren (expliciete) kruisverbanden
 - o Tijdelijk delegeren van rechten aan externe services
 - Ten behoeve van opslag van (ratings voor) objecten/concepten
 - Ten behoeve van raadplegen van user profiles en ratings
- Front end voor profielbeheer door profieleigenaren zelf
- Front end voor beheer van de UPR zelf
- Authenticatie en Autorisatie module
- Widgets
 - o Registratie en inloggen
 - o Rating van objecten

Op langere termijn kan de UPR voor erfgoedgebruikers een centrale plek worden om in te loggen en eigen voorkeuren te bewaren ('cultural bookmarks'). De gebruikers kunnen collectie aanbieders en recommendation services tijdelijk en ge-anonimiseerd toegang verlenen tot deze informatie, zodat deze daarmee gepersonaliseerde diensten aan kunnen bieden.

2. Art Recommender

In het kader van CATCHPlus wordt een recommender service gebouwd op basis van de collectie-gegevens van één of meer deelnemende erfgoedinstellingen. Primair wordt deze recommender service toegepast op collectie-data van het Amsterdam Museum. Deze recommender service maakt voor wat betreft gebruikersprofielen gebruik van de UPR.

De recommendation service moet in staat zijn om, gegeven een (set van) collectie-object(en), een set objecten aan te bevelen, die de gebruiker mogelijk ook interessant vindt. De service wordt geïmplementeerd als toepassing van ZieOok en ge-tuned op de AM erfgoedcollectie(s). Daarmee heeft ZieOok ook een extra toepassing erbij.

Concreet wordt gebouwd (in te vullen door Siem en/of Gridline):

- Evaluatie en eventuele aanpassing OAI data provider van AM

- (eventueel) Mechanisme voor aanleveren logging gegevens. In een vooronderzoekje wordt nader bepaald of dit zinvol is, en binnen de scope van het project past.
- Tuning ZieOok op AM collectie, liefst interactief en door het AM zelf met behulp van het ZieOok Dashboard.
- Inbouwen ZieOok recommender widget in AM website
- inbouwen login- en rating-widgets van de UPR in AM website

Waarom willen we de UPR bouwen?

Verschillende partijen hebben verschillende mogelijke belangen bij het beschikbaar zijn van de User Profile Repository.

Gebruikers van erfgoed-websites:

- Zijn baas over hun eigen profiel-informatie, en bepalen zelf actief wie daar gebruik van mag maken, in welke mate en hoe lang.
- Gebruiker kan anoniem blijven, profiel informatie is niet te herleiden tot specifieke gebruikers.
- Geen 'filter bubble': gebruikers bepalen zelf wanneer met welke profiel-informatie een service wordt gebruikt. Ze zijn zelf baas over of en hoe resultaten gefilterd worden teruggegeven.
- Verzamelde informatie is in principe bruikbaar voor verschillende collecties en voor verschillende services.

Aanbieders van erfgoedcollecties en -websites:

- Hoeven geen gebruikersadministratie bij te houden, en geen statistieken of ratings bij te houden.
- Desondanks kunnen ze toch inzicht krijgen in gebruikers- en gebruiksgegevens. Ze hebben vrij toegang tot geaggregeerde gegevens, en met toestemming van de eigenaren, ook tot (geanonimiseerde) individuele gegevens.
- Hebben mogelijk ook toegang tot profiel-informatie, die is gebaseerd op gebruik van andere collecties.

Toolbouwers en dienstenaanbieders:

- nette interfaces en een modulaire opbouw
- kant en klare bouwstenen

Use cases

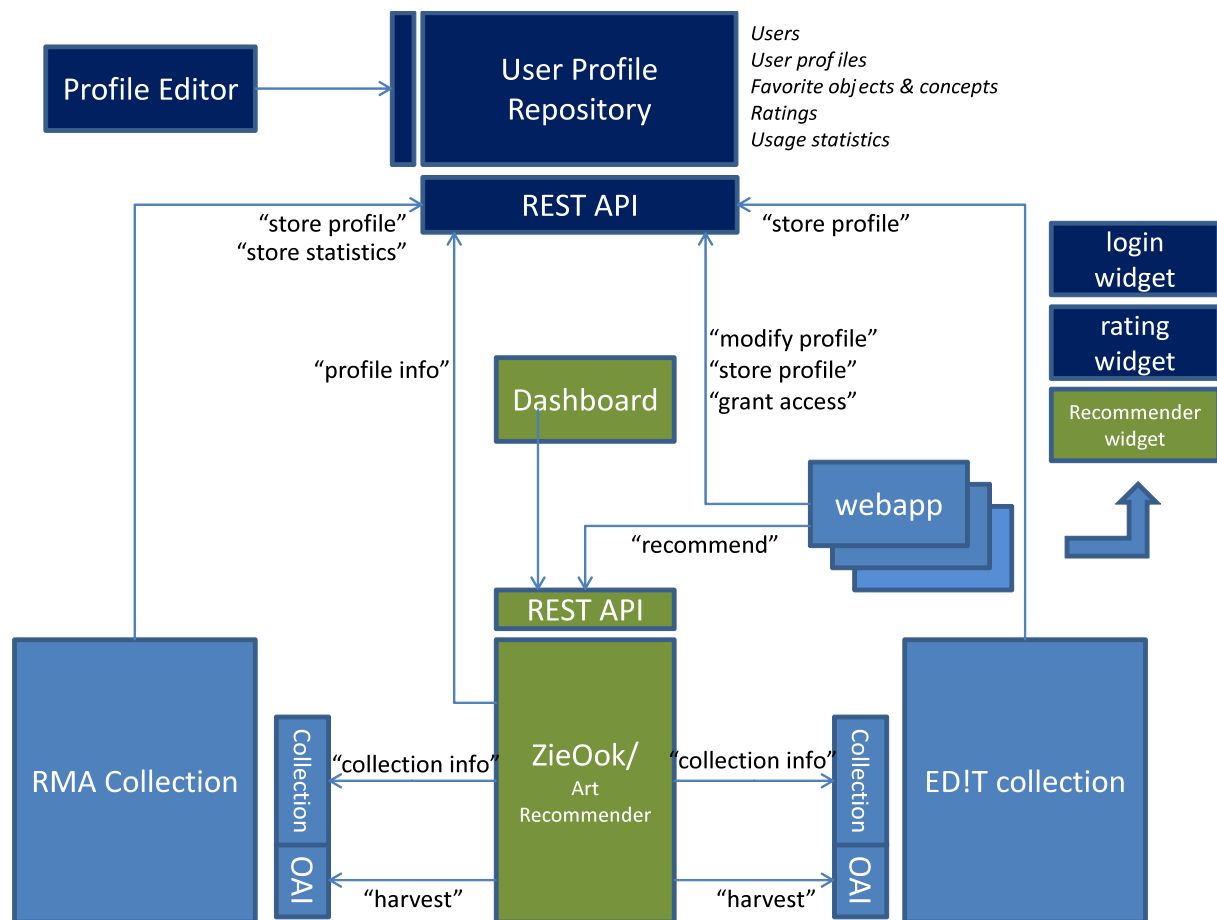
Enkele nader uit te werken en aan te vullen use cases:

- ik (beheerder) wil een nieuwe gebruiker met een gebruikersprofiel kunnen aanmaken (of verwijderen)
- ik (profiel-eigenaar) wil mijn gebruikersprofiel kunnen invullen, inspecteren en aanpassen
- ik (profiel-eigenaar) wil vanaf een collectie-webpagina kunnen inloggen, om toegang tot mijn profiel-informatie te krijgen
- ik (profiel-eigenaar) wil vanaf een collectie-webpagina verwijzingen naar objecten en object-beoordelingen aan mijn profiel kunnen toevoegen
- ik (profiel-eigenaar) wil op grond van een object op een collectie-webpagina aanbevelingen krijgen met betrekking tot andere objecten
 - o op grond van mijn eigen profiel

- op basis van beoordelingen en gebruik door alle andere collectie-bezoekers
 - ik (collectie-aanbieder) wil, op nog niet nader gedefinieerde wijze, website-bezoekers gepersonaliseerd iets aanbieden op basis van hun gebruikersprofiel
 - ik (ZieOok) wil informatie uit de UPR ophalen en gebruiken om recommendations voor een bepaalde collectie te kunnen aanbieden.
 - ...
- Wat zijn de randvoorwaarden?

Architectuur

De UPR is onderdeel van een samenhangende architectuur, zoals bijvoorbeeld in de onderstaande schets is aangegeven



Autorisatie service

Voorstel is om autorisatie te baseren op OAuth 2.0.

OAuth 2.0 kent een aantal rollen: client, resource owner, authorization server en resource server. De client is een applicatie, die toegang wil tot een resource op de resource server. De resource owner bepaalt wie welke rechten heeft op de resource. Dat kan gebeuren in interactie met een menselijke gebruiker, dat kan ook namens zo'n persoon automatisch gebeuren op grond van opgeslagen toegangrechten (of Access Control Lists). De authorization server geeft namens de gebruiker tijdelijke

'toegangsbewijzen' aan de client, die daarmee de gewenste toegang tot de resource krijgt.

Concreet te implementeren:

- Ieder van de clients moet volgens een OAuth workflow toegang tot resources gaan verkrijgen. Mogelijke clients zijn voor onze gebruiks-scenario's: UPR rating widget, ZieOok recommender widget, profile editor, (mogelijk ZieOok zelf?).
- Een Authorization server: naast login voor clients en resource owners geeft deze AS z.g. authorization grants en authorization tokens uit. Bovendien heeft deze AS toegang tot specifieke informatie over mogelijke soorten toegang tot resources op de resource server.
- Voor niet-interactieve toegang moet de UPR access control lists bijhouden voor de beschikbare profielen. Mogelijk is in geen van onze gebruiks-scenario's sprake van niet-interactieve toegang. Dan vervalt dit onderdeel.

Een en ander kan waarschijnlijk worden geïmplementeerd op basis van beschikbare open source libraries. Mogelijk kan de software voor de Authorization Server worden gedeeld met het CATCHPlus Workspaces deelproject.

Alle authenticatie voor onze scenario's vindt plaats bij de Authorization Service. Deze is daarmee een goede plek om in een later stadium eventuele single sign on oplossingen aan te koppelen.

Niet-functionele requirements

Er moeten afspraken worden gemaakt over gebruikte basistechnologiën.

Eventuele tussenreleases, interne en externe afhankelijkheden, externe deadlines en dergelijke worden door alle stakeholders samen zo vroeg mogelijk in kaart gebracht.

Oplevering en exploitatie na afloop

In overleg nader in te vullen.

Open source en publicatie daarvan (GPL 3.0, github)
Hosting, open service gebruik?
Business model?