

Tu delft

Groep 5

Salim Hadri  
Kaj Oudshoorn  
Viresh Jagesser

IKE Project

Inhoud

Table of Contents

Inleiding 3

2 Ontwerp 4

Probleem beschrijving 4

Ontwerp beschrijving 4

Functionele beschrijving 4

Use case 6

Workflow 7

3 Implementatie 8

4 Evaluatie 10

Tussen product(wat is er eigenlijk bereikt) 10

Verloop van implementatie (Is het volgens plan gegaan?) 10

5 Test uitvoering 11

Test senario’s 12

Welk test-scenario is uitgevoerd? 12

Wat was de begintoestand van de database en/of de applicatie? 12

Welke testgegevens zijn ingevoerd? 12

Wat is het waargenomen gedrag van de software? 12

Wat is het verwachte gedrag? 12

Waarin wijkt het waargenomen gedrag precies af van het verwachte gedrag? 12

Hoe kan de fout gereproduceerd worden? 12

In welke versie van de software is de fout aangetroffen? 12

Wat heeft u gedaan om mogelijke oorzaken van de fout te achterhalen? 12

# Inleiding

De opbouw van dit rapport is als volgt. Hoofdstuk 2 beschrijft globaal de oplossing voor de recommandatie systeem d.m.v. een probleembeschrijving, ontwerpbeschrijving, functionele beschrijving en een ruwe oplossing met daarbij een workflow. In hoofdstuk 3 komen de details van het ontwerp aan de orde. De evaluaties van dit product worden besproken in hoofdstuk 4. Tot slot staat er in hoofdstuk 5 de tests die zijn uitgevoerd.

# 2 Ontwerp

## Probleembeschrijving

De gebruiker moet worden geadviseerd met andere muziek, nadat hij zelf een nummer heeft ingevoerd (hierbij ook de artiesten naam). Voorlopig werkt dit met een enkel nummer, later worden dit er meerdere. Hierbij moet naar overeenkomsten tussen de nummers worden gezocht, waardoor er vergelijkbare nummers teruggegeven kunnen worden. Daarnaast moet het aantal geadviseerde nummers niet te groot zijn (dan wordt het onoverzichtelijk) of te klein (niet genoeg keuze).

## Ontwerpbeschrijving

Het programma wordt in java geschreven, omdat de programmeurs ervaring hebben met Sesame (een library voor het werken met SPARQL end points). Later bleek wel dat veel databases voor PHP gemaakt waren en dat een juiste database moeilijk te vinden was. Maar omdat uiteindelijk toch een database van Kasabi werd gevonden, is besloten om toch bij Java te blijven.

## Functionele beschrijving

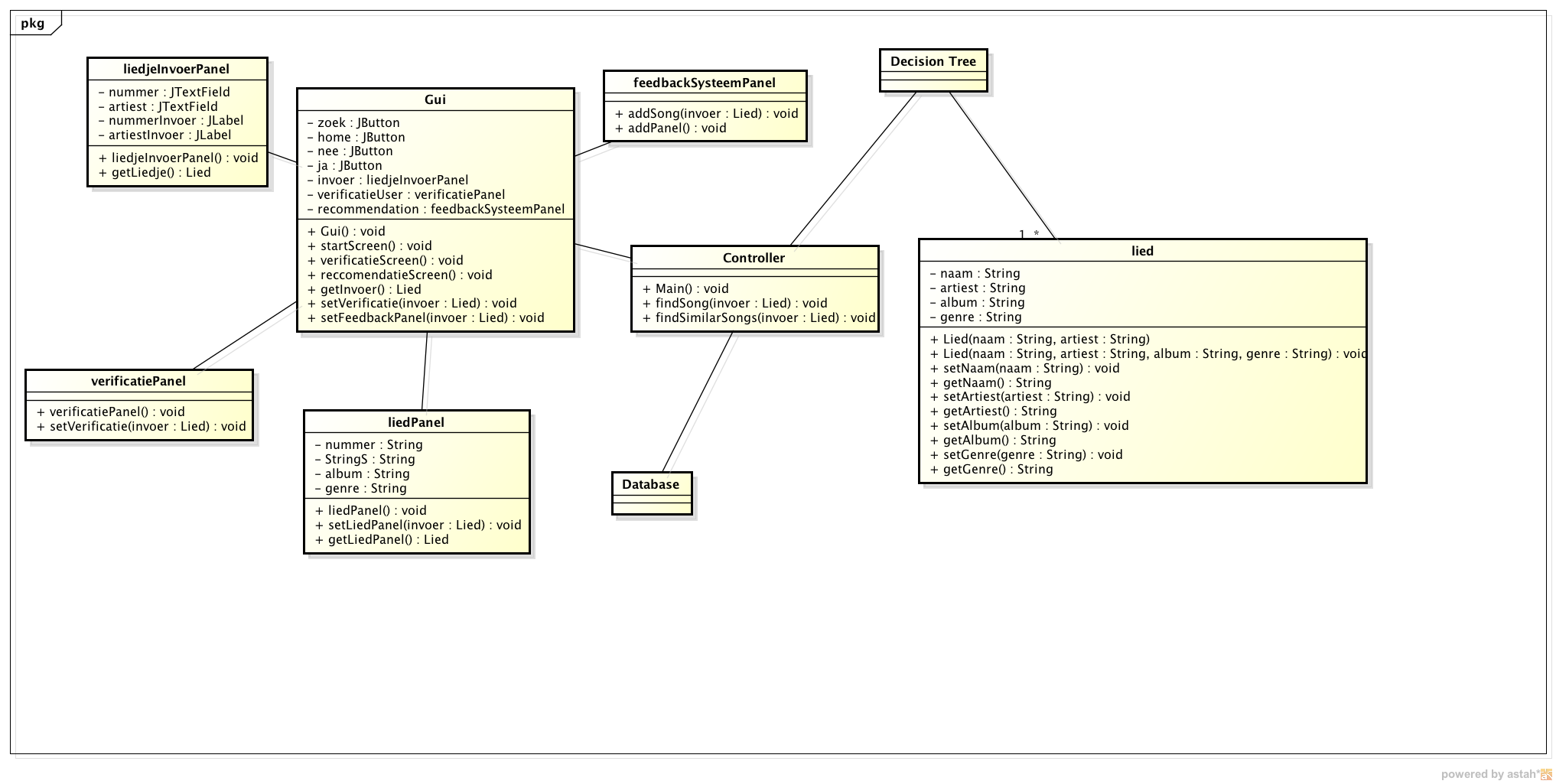
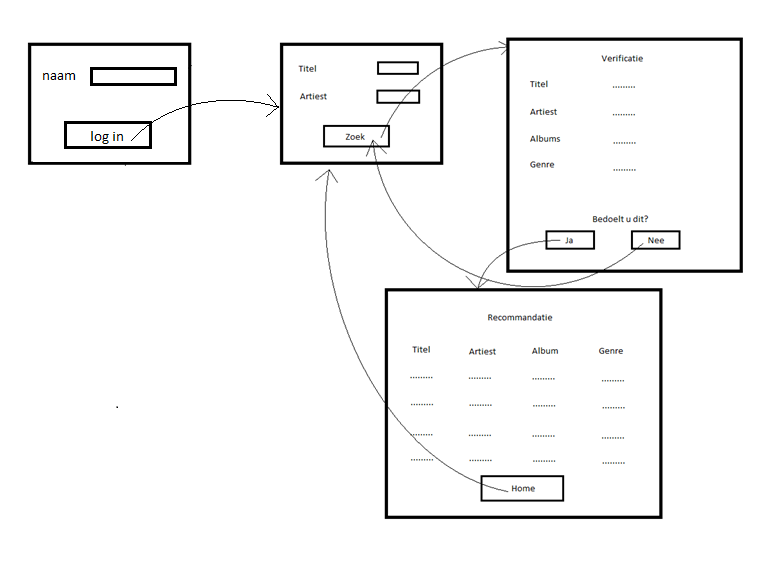
Het systeem moet de volgende dingen kunnen:

* Systeem moet gegeven nummer + artiest herkennen.
* Gevonden nummer verifiëren: met nummer, artiest, genre, album.
* Recommandaties teruggeven van nummers die vergelijkbaar zijn met het gegeven nummer.
* Bij recommandaties nog extra info geven (zoals een review).

De volgende dingen zijn gewenst:

* Geen lange wachttijd (niet langer dan een paar seconden)
* De optie om opnieuw een recommandatie te vragen, zonder het programma opnieuw te hoeven opstarten.

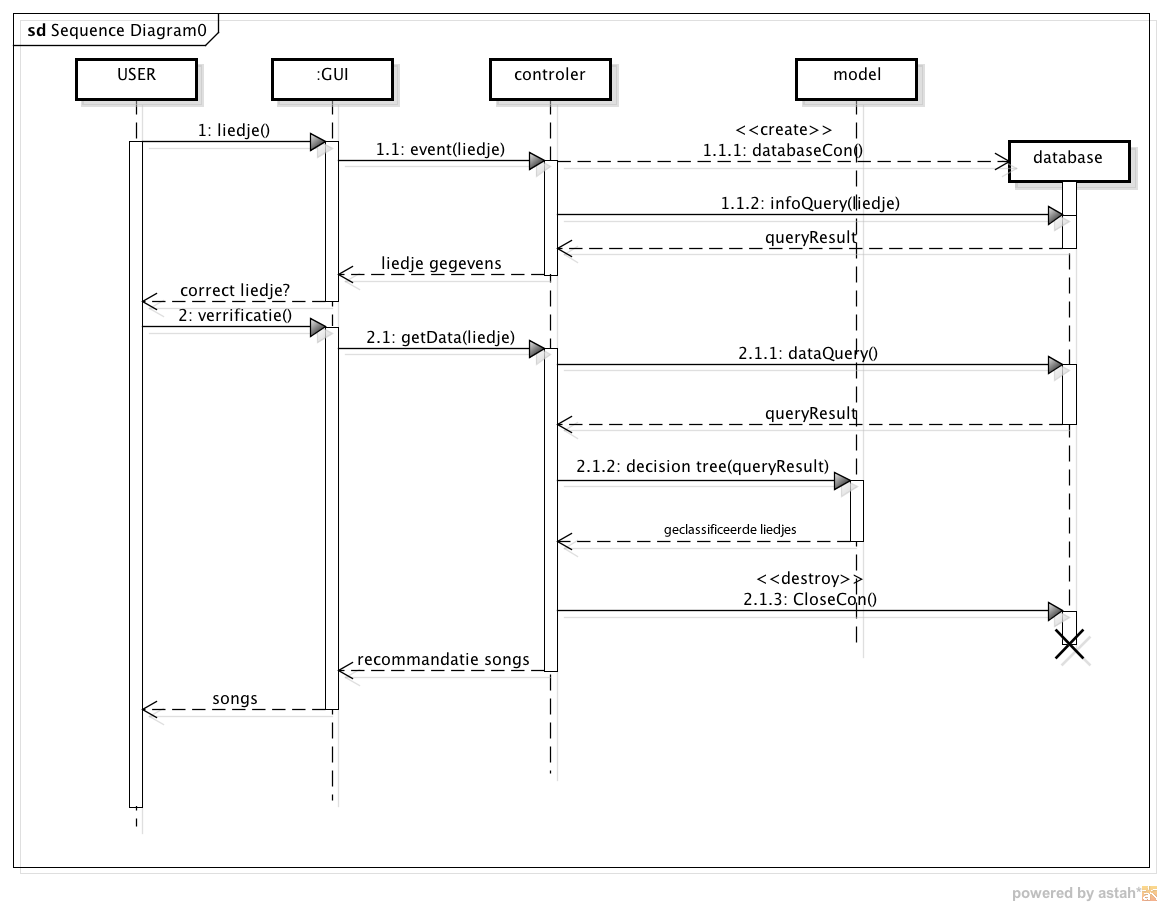
Interface model

 Klassendiagram

## 

## Use case

## Workflow



In dit sequence diagram is de workflow van het programma te zien, horende bij de use case van succesvol een liedje opzoeken en suggesties opvragen. Het diagram zelf spreekt grotendeels voor zich. Hou hier wel rekening mee dat model een verzameling van klassen is die de liedjes van de database verwerkt tot een lijst van aanbevolen liedjes.

Ruwe oplossingen Kaj verbeteren O\_\_\_O

Er wordt gebruik gemaakt van de lastFM database. Met deze database kunnen tracks, artiesten en albums worden opzoeken. Met de methode getSimilar (uit de lastFM API) worden er vergelijkbare tracks gevonden en deze worden geretourneerd.

Later implementeren wij de getSimilar functie zelf m.b.v. een Decision Tree. Tracks met overeenkomsten komen hierbij bij elkaar te liggen en vervolgens wordt er rondom het gegeven nummer gezocht naar tracks.

De Decision Tree vergelijkt nummers op basis van bijbehorende tags. Om er voor te zorgen dat er niet een eindeloze lijst van tags hoeft te worden afgelopen, worden alleen de meest gebruikte tags opgevraagd m.b.v. de methode getTopTags.

Verificatie wordt gedaan door info over het gegeven nummer terug te sturen, waarmee de gebruiker kan beslissen of hij dat nummer daadwerkelijk bedoeld.

Beide databases zijn remote, de data hoeft dus niet worden ingeladen bij het opstarten van het programma en ligt vanaf het begin al klaar om opgehaald te worden. Hierdoor is er geen lange laadtijd.

# 3 Implementatie

## Vorige iteratie

Als eerste werd de GUI geïmplementeerd, hiervoor werd gebruik gemaakt van de awt java library. Vervolgens werd de LastFM library geïmporteerd. In de LastFM klasse werden de functies findSong en FindSimilar geïmplementeerd. Zoals de naam al doet vermoeden kan je met de findSong functie op basis van een naam van een nummer en artiest het liedje vinden in de database van lastFM. Dit was heel makkelijk te implementeren aangezien lastFM een heel duidelijke API heeft. Bij het implementeren van findSimilar songs wordt er gebruik van een standaard functie die op basis van een lied al recommandaties kan doen. Dit is echter tijdelijk, aangezien er bij het implementeren van de { Ontology…..} geen werkende connectie met een database kon worden gemaakt. Daarom werd er een zijsprong naar de last Fm database gemaakt.

Als laatst is er ook nog een Controller klasse die geïmplementeerd is, dit verliep ook soepel . De controller klasse zorgt voor het dataverkeer tussen de GUI en de last Fm database.

## Huidige iteratie

Deze iteratie zijn er een aantal functies toegevoegd. De recommandatie van songs wordt nu gedaan door middel van clusteren. Van het ingevoerde liedje worden de bijbehorende tags opgevraagd. Per tag worden vervolgens de meest populaire liedjes opgevraagd. Zo ontstaat een lijst van liedjes die minimaal eenzelfde tag hebben als het gegeven liedje. Vervolgens wordt per liedje de overlap van tags bepaald en hierop worden de liedjes gesorteerd. Als laatste worden de liedjes met de meeste overlap teruggegeven aan de user (hierbij wordt wel een maximum van 10 liedjes aangehouden).

Bij het recommanderen hangt ook nog een voorwaarde: er mogen maar max 2 liedjes van dezelfde artiest worden gerecommandeerd. Dit wordt gecontroleerd door tijdens het toevoegen van liedjes een lijst van artiesten bij te houden met bijbehorende aantal. Als dit aantal gelijk is aan 2, betekent dit dat er van die artiest geen liedjes meer mogen worden toegevoegd.

Een user kan gekoppeld worden aan een username en hiervan wordt bijgehouden welke liedjes hij heeft ingevoerd, welke tags hij belangrijk vind en welke niet. Momenteel wordt hier nog niets mee gedaan, dit is voor de volgende iteratie.

Als een user een liedje heeft ingevoerd wordt vervolgens in het verificatie scherm (zie interface model) gevraagd welke van de bijbehorende tags hij belangrijk vindt. Deze tags worden dan gebruikt voor het recommanderen zoals hierboven beschreven.

Er kwamen bij het recommanderen nog 2 fouten voor. Als een liedje niet gevonden kon worden stuurde lastFM een null pointer terug, maar het programma probeerde hierdoor nog verder te werken, wat problemen opleverde. Dit is opgelost door terug te gaan naar het invoer scherm wanneer een liedje niet gevonden kan worden en een melding hiervan te geven.

Het andere probleem had te maken met de tags. Van sommige tags kon je niet de populairste liedjes opvragen (dit leidde tot een exceptie van lastFM). Het is onbekend waarom dit gebeurde, maar waarschijnlijk kwam dit omdat de tags bijna nergens voor gebruikt werden.

Clusteren, tags kiezen, inloggen, opslaan van keuzes nog bezig, error handling, restrictie: niet meer dan 2x dezelfde artiest in recommandatie.

# **4** Evaluatie

## Eind product

Op dit moment is de recommandeer systeem af. Om te beginnen moet een gebruiker inloggen met zijn gebruikers naam. Dit wordt gedaan om de zoek geschiedenis bij te houden van elke gebruiker. Met deze functie zal het systeem betere recommandaties geven na mate de gebruiker meerdere nummers heeft gezocht. In de volgende stap kan elke gebruiker een of meerdere nummers invoeren. Ook kan de gebruiker die eerder gebruik heeft gemaakt van dit systeem, de “quick recommand” functie gebruiken. In dit geval kijken dit systeem naar de zoek geschiedenis van deze gebruiker en zal op basis hiervan een recommandatie geven. Wanneer een gebruiker nummer(s) handmatig in heeft gevoerd, zal deze komen in de verificatiescherm. Hier wordt gevraagd welke tags bij de smaak van de gebruiker past. De aangevinkte tags samen met de tags die de gebruiker in vorige zoek opdrachten heeft ingevoerd, worden meegenomen in de recommandatie. Hier geldt dat elke recommandatie max. 10 nummers heeft.

## Verloop van implementatie

Het is uiteindelijk niet gelukt om een werkende end point te vinden met de informatie die gezocht werd. Daarom wordt er nu alleen nog gebruik gemaakt van lastFM.

LastFM maakt alleen gebruik van tags, daarom moesten er ook een aantal aanpassingen worden gemaakt voor het vergelijken van nummers. Verder wordt de zoek geschiedenis van elke gebruiker bijgehouden.

## Verbeter punten

Dit systeem werkt helaas niet altijd even goed. De nummers die lastFM stuurt naar dit systeem zijn “toptracks”. Dat wil zeggen dat dit systeem geen onbekende nummer als recommandatie kan geven. Dit zou opgelost kunnen worden door gebruik te maken van een database, die per tag, alle nummers kan geven i.p.v. alleen “populaire” nummers.

Het leerproces van dit systeem is niet optimaal, omdat de recommandaties niet beter worden na mate er meer gebruikers, gebruik maken van dit systeem. Dit systeem ziet elke gebruiker als individu en gebruikt geen informatie van andere gebruikers.

## Manier van testen

## Om te meten hoe goed onze recommandaties zijn, zal er gebruik gemaakt worden van playlists en proefpersonen. Om de playlists te testen zal bij elke playlist een deel weggelaten worden. De recommandatie systeem zal daarna het weggelaten deel moeten recommanderen. Hierna kan worden gemeten hoeveel van de weggelaten nummer worden terug gevonden. Om te meten hoe goed dit systeem het doet bij proefpersonen, zal deze persoon liedjes invoeren. Daarna zal deze persoon een oordeel geven over de recommandatie die hij/zij heeft gekregen.**5** Test uitvoering

Dit systeem heeft een aantal functionaliteiten.

• Het Systeem moet nummer + artiest herkennen.

• Verificatie: met nummer, artiest, genre, album.

• Recommandaties teruggeven van nummers die vergelijkbaar zijn.

• Bij recommandaties nog extra info geven (zoals een review).

Deze functionaliteiten zullen getest worden met behulp van de userinterface. Dit zal in het begin door de programmeurs zelf worden gedaan en later door een onafhankelijk persoon. Daarbij letten we op een aantal aspecten. Om te beginnen met het gedrag van het systeem en hoe zinvol dit systeem reageert op foute invoer. Verder wordt er gekeken of de output de door de tester verwachte data bevat. Tot slot moet dit systeem makkelijk te gebruiken zijn, dus wordt er getest op gebruiksvriendelijkheid.

Het testen van de recommandatie wordt echter op een andere manier gedaan. Hier worden de resultaten van de lastfm getSimilar functie vergeleken met onze resultaten, die zijn gevonden met behulp van onze cluster algoritmen.

## Test senario’s

Op de ”<http://www.applinet.nl/artikelen/acceptatietest.html>” site is een vragenlijst gegeven. We zullen dit gebruiken om een aantal test senario’s uit te voeren.

## Welk test-scenario is uitgevoerd?

## Wat was de begintoestand van de database en/of de applicatie?

## Welke testgegevens zijn ingevoerd?

## Wat is het waargenomen gedrag van de software?

## Wat is het verwachte gedrag?

## Waarin wijkt het waargenomen gedrag precies af van het verwachte gedrag?

## Hoe kan de fout gereproduceerd worden?

## In welke versie van de software is de fout aangetroffen?

## Wat heeft u gedaan om mogelijke oorzaken van de fout te achterhalen?