# Gepersonaliseerde suggesties aan de hand van voorkeuren voor genres

## De invoer

De invoer die gebruikt wordt om de suggesties voor een gebruiker te berekenen, bestaat uit 3 onderdelen, namelijk een dataset van Yahoo, gebruikersvoorkeuren en gebruikerswaarderingen.

### Algemene data van Yahoo

In de dataset van Yahoo die gebruikt is voor dit project, was een koppeling te maken van een hoeveelheid stemmen voor bepaalde genres. Deze stemmen vormen een basisscore voor enkele van de genres, zodat er in het systeem redelijk wat data aanwezig is zonder input van de gebruiker. Deze data wordt puur gebruikt als initiële scores voor genres, zodat enkele relaties wat betreft waardering tussen genres alvast vaststaan.

### Gebruikersvoorkeuren

Na de registratie bij het systeem, komt de gebruiker als eerste een pagina tegen waarop hij gevraagd wordt om zelf aan enkele vooraf gedefinieerde genres een score toe te kennen. Deze scores vormen naast de geïmporteerde data uit de Yahoo-dataset een zeer algemene indruk van wat een gebruiker graag zou luisteren.

### Gebruikerswaarderingen

De gebruikerswaarderingen vormen het belangrijkste onderdeel van het systeem. Zodra het systeem enkele suggesties gegeven heeft, kan de gebruiker aangeven of deze suggesties goede suggesties zijn of niet. Deze scores worden opgeslagen in een database, zodat er na verloop van tijd een gedetailleerd beeld ontstaat van wat de gebruiker zijn muzieksmaak is.

## Het achterliggende systeem

Nu het verzamelen van de gegevens geregeld is, moet er ook nog wat mee gedaan worden. In dit deel wordt een deel van het datamodel besproken, en de verwerking van de data.

### De graaf

Er is gekozen om de relaties tussen verschillende genres vast te leggen in een graaf. De nodes in de graaf stellen de genres voor, edges tussen deze nodes geven aan hoe sterk de relatie is tussen 2 nodes (of geven aan dat er geen relatie is als er geen edge tussen 2 nodes is).