Information Retrieval - Project Report

# Problem Statement

There is an abundance of music available in today’s world. Numerous songs, artists and even genres have been popping up all over the world over the last couple of decades. The sheer amount of musical material sometimes makes it difficult for a person to find new music that he or she would really like. To remedy this problem, services have begun to implement recommender systems to help people discover artists or songs that suits them based on their or similar users’ taste in music.

This is where our project comes into play. We are going to build a recommender system for Last.FM, one of these services, because we would like to know which method works best (user-based, item-based or a combination of both) for accurately recommending music artists to listen to. Last.FM is an online service that allows users to listen and share any music that they would like. Since their inception in 2003, over 100,000,000,000 tracks have been listened to on the service. Our recommender system will utilise a provided dataset containing the activity of users in regards to which artists they have listened to and how much they have listened to those artists to determine similar users for the user-based recommender. For the item-based recommender, we will make use of a provided dataset of tags that have been given to songs and artists to determine similar songs or artists. We believe this system could prove fruitful in helping improve the algorithm for discovering new music on the Last.FM service. What combination of recommender techniques gives the best results in recommending artists to users?

This project is directly derived from the research proposal submitted by Karel Beckeringh (s2600358). We felt that he hit the nail on the head with his proposal, since the subject matter and dataset is rather unique and appeals to all of us on a personal level and the project would offer a welcome challenge for us as a group. As such, we unanimously decided to perform his proposal.