

# The Search for the Searcher:

## User Identification Using Web Search Log Mining

Author: Tom Moore

Supervisor: Prof. Séamus Lawless

Degree: Msc. in Computer Science

Year: 2017

### **Abstract**

User grouping is the act of treating users that share common characteristics as groups that can be modeled. These characteristics can be called user features. Transactional logs are records of interactions between users and systems. User features can be derived from transactional logs. It must therefore be possible to derive user groups from transactional logs. One method of forming groups from data points is with data clustering. As user features are shared characteristics about users, it may be possible to form groups of users using data clustering techniques on their use features.

However, is it possible to extract user features from transactional logs and then cluster those features using data clustering techniques to form clusters from which user groups can be derived? This thesis is an experiment to see if such an idea is possible.

To that end this dissertation proposes a pipeline that will take transactional logs as an input and will output data clusters from which user groups can be derived. The outputted clusters will be validated and analysed using data analysis techniques such as silhouette scores to see if it is possible to derive user groups from them.