

# Exploring the use of Semantic Techniques in Online Dating

Emily Wilson / Kenneth B. Kent

University of New Brunswick

Faculty of Computer Science

l4re6@unb.ca ken@unb.ca

## Outline

- Implementation of semantic reasoning in an existing online dating system
- Use of geographical data to infer new properties about customers and their potential suitors
- A four level matching architecture

## Motivation

Users may seek out online dating websites as a means of finding potential partners quicker. Producing a set of matches for a person that they will deem suitable is one of the major challenges in the online dating industry. Users enter information about themselves that is then used to identify other users that considered appropriate matches. However, many users become frustrated at the list of potential matches that they are returned as they are deemed unsuitable. This leads to customers cancelling their memberships and giving the company's website poor reviews.

In this contribution we present an application of semantic technologies to online dating matching techniques. These techniques consider more than just parameter based criteria. Information provided is used not only to eliminate matches deemed unsuitable but also to infer additional properties about the user, that will be used in the matching process.

## Objectives

- Find a suitable rule language and reasoning/inferencing engine in terms of performance and scalability.
- Incorporate geographical aspects of the relationship into the matching process.
- Improve matching results based on details extracted from users as well as rules extracted from user input to provide matches of higher semantic quality.

## Methodology

The matching system uses a level based architecture. There are four levels in total. Matches are generated for users in the system with a score based on level. Members answer questions from numerous levels (depending on their membership), and the system performs the matching process based on a pre-defined set of rules in the form of background knowledge.



Level 1 matches a user based on their basic information (gender, looking for, etc.). Any other users that do not meet the level 1 requirements for a match are eliminated from the potential matches. Next, a user is matched by their answers to the 5 questions in level 2 that they choose to answer regarding priorities and given a score representing the accuracy of the match. Information from level 2 is then used in level 3 and 4 matching. Based on their answers in Level 2 users are given a set of level 3 questions to answer regarding psychological aspects.

In level 4 data is inferred based on a user's geographical area. From geographic information we can infer the following things:

- Their housing
- Their income level
- Their language
- Their level of education
- Their major field of study
- Their employment status and occupation

This information is inferred by using Statistic Canada's 2006 Census of Population data. Also, information provided by other users who live in the same geographical area is used to infer information about other users who live in that area. The data produced by geographical inferences only affects a match's score and does not eliminate any potential matches from level 2.

## Implementation

Each user in the matching system has an RDF triple that identifies who they are, as well as RDF triples that identify all of the information that they have entered. An SQL database populated with Statistics Canada data is used to add additional triples about a person. Rule-based reasoning is then performed on those triples to generate additional information about a user to be used in the matching process.



Commercial Partner:

