

# Linguistic Self-Expression on Dating Websites



# Elizabeth Dellea

Emerging Media, '16 • Ithaca College

Do generalizations about  
gender and language hold true  
in online communications?

Past linguistic studies and common beliefs tell us that men and women use language differently. Our research investigates such claims in an online context.

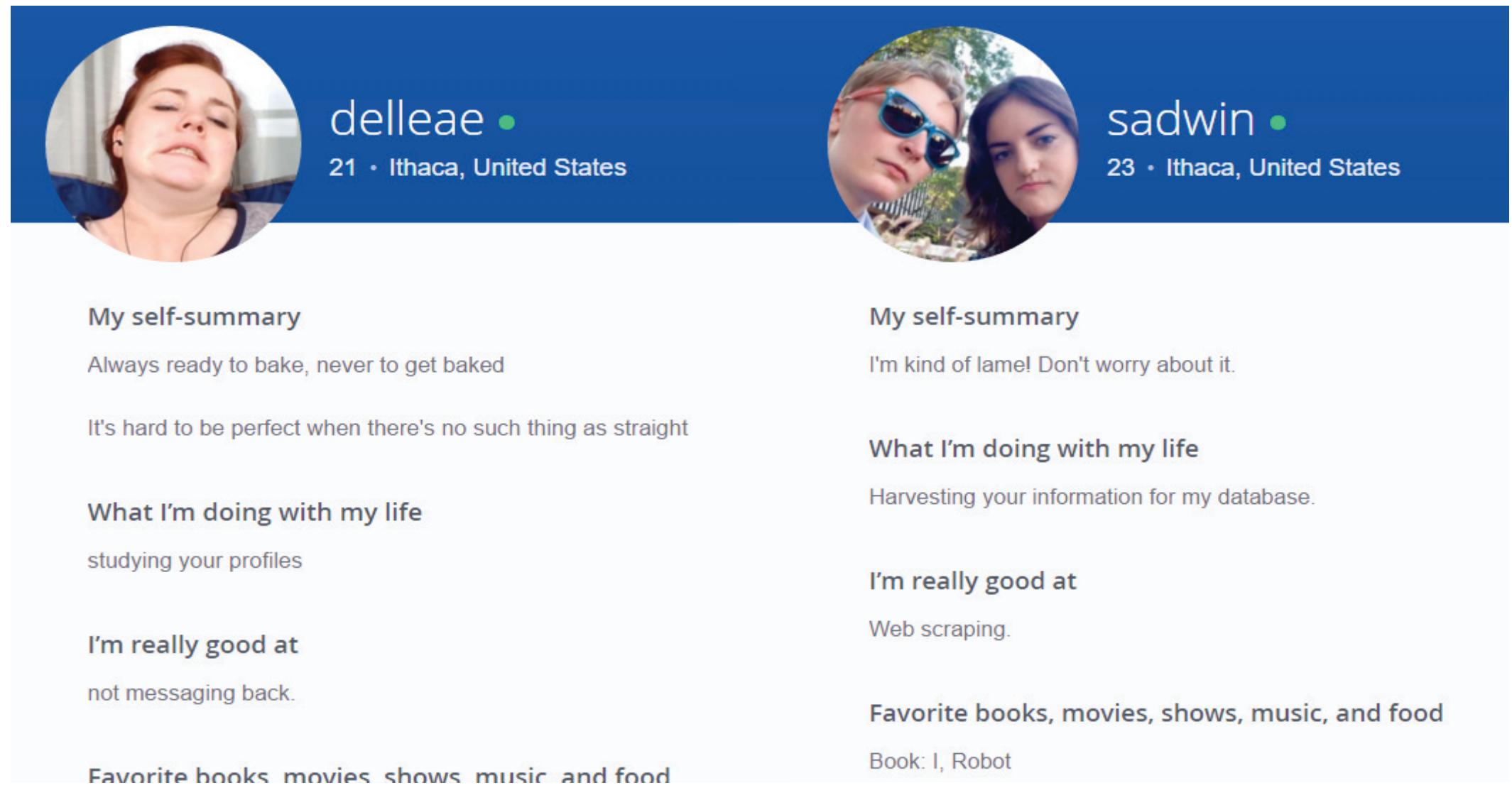
So, where can we find substantive samples  
of people writing about themselves?  
Dating websites, of course!

# Web Scraping to Gather Data

We created a pair of bots on OkCupid and cast as wide a net as possible for our potential “matches”.

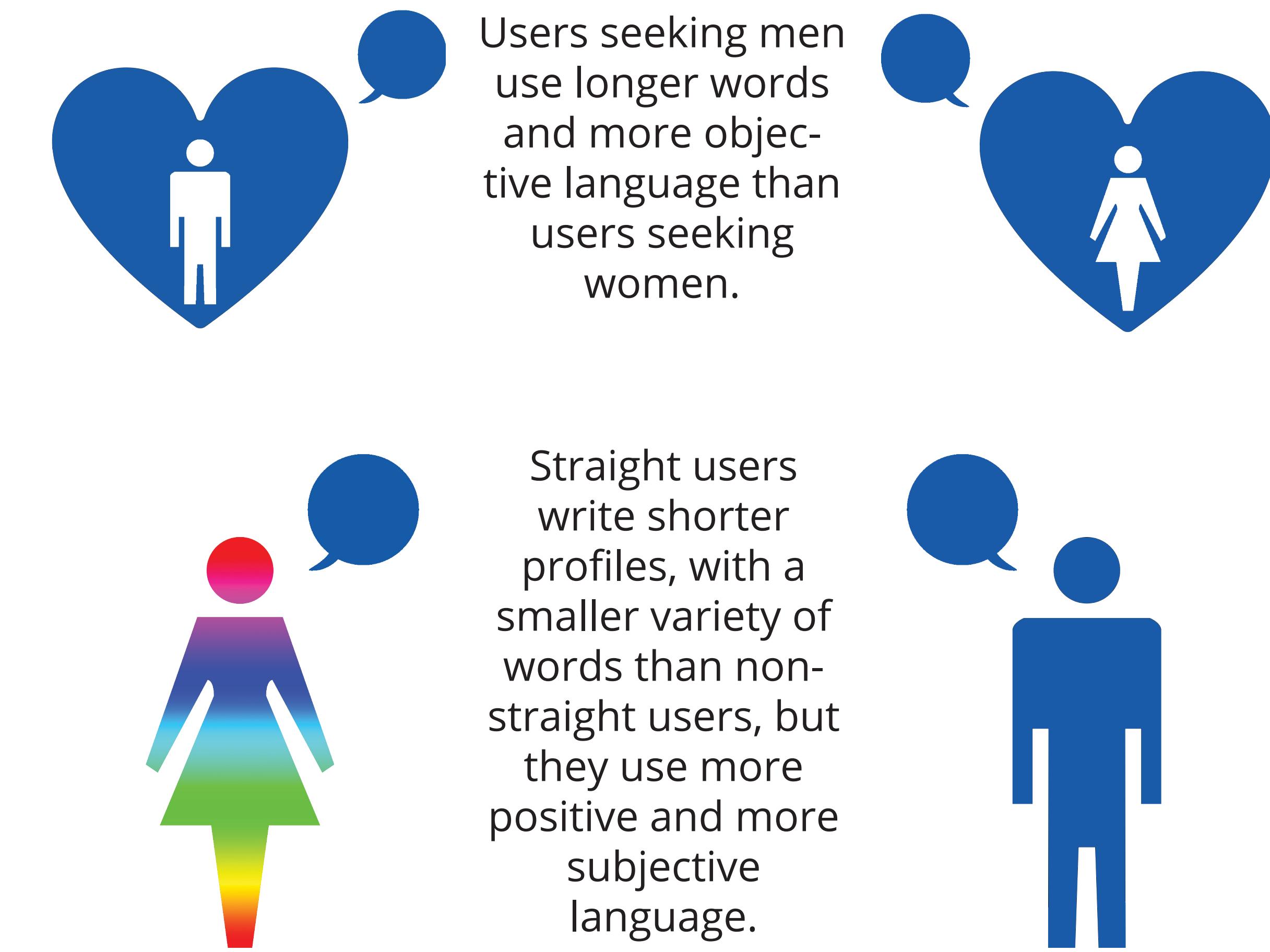
One male bot, one female bot  
Both bisexual and looking for everyone  
Ignore location for matches from all over the world  
Looking for ages 18-30  
Answered minimum number of survey questions

We designed a web scraper in Python to download over 1500 matches' profiles and their demographic tags.



# Word Clouds

Using ITC tf-idf to increase the importance of words unique to the demographic, we generated the following word clouds. We'll let them speak for themselves.



Interestingly enough, there were no statistically significant differences found between men and women.

# References and Special Thanks

Python modules used: wordcloud, beautifulsoup, textblob, nltk, matplotlib, and numpy

Iconography by Edward Boatman, Matt Brooks, and Hayashi Fumihiro ([nounproject.com](http://nounproject.com))

Special thanks to our faculty adviser, John Barr, Ph. D.

Thanks to Doug Turnbull, Ph. D., for course material on web scraping, tf-idf, and NLP.

scraping, tf-idf, and NLP.  
Additional thanks to Laurie Arliss, Ph. D., professor of  
Communications and Women's & Gender Studies for her insight