# Male Strategic Self-Representation Using Language on Online Dating Sites

Abhishek Pandit
Social Sciences Division, University of Chicago

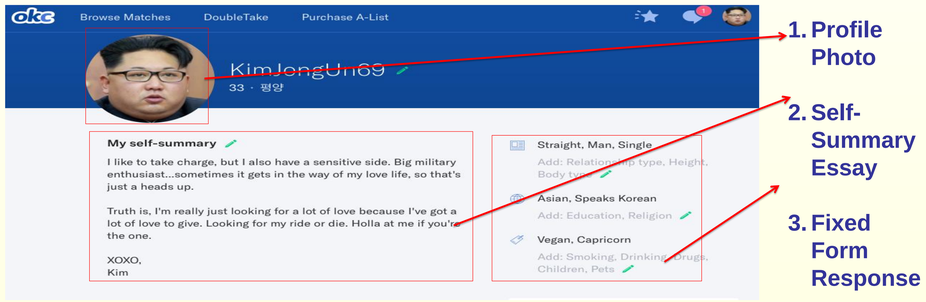
## Online Dating: New Platform for Old Human Question

•Online dating has developed into a $1.67 trillion industry with 340 million users. 1 in 3 American married couples claims to have met online.

•These recent advances fit with old philosophical questions on how humans change self-representation based on strategic objectives

•In online dating, this increasingly takes on the form of 'Relationshopping' where users market themselves as products (Heino, 2010)

## What Matters Most in Self- Representation?




•Profile photos continue to hold highest importance for both sexes. Fixed form responses are not correlated with attractiveness (Fiore, 2010)

• Women place more importance on men's self summary than men do on women's (Emory 2017). So men who may have been at a disadvantage due to aspects of their photos and fixed responses) can strategically 'compensate' in the way they write self-summaries

**QUESTION: Are there clear differences in men's language choice in their dating profile self-summaries across levels of specifics variables (like education, height, etc.) that may have disadvantaged them in terms of potential romantic prospects in the real world?**

## Data & Methods

•Public and anonymized profile data for 59,946 active users (filtered down to 18,830 men) of OkCupid within a 25 mile radius of San Francisco that were extracted with permission. The data set includes "typical user information, lifestyle variables, and text responses to 10 essay questions" (ibid).

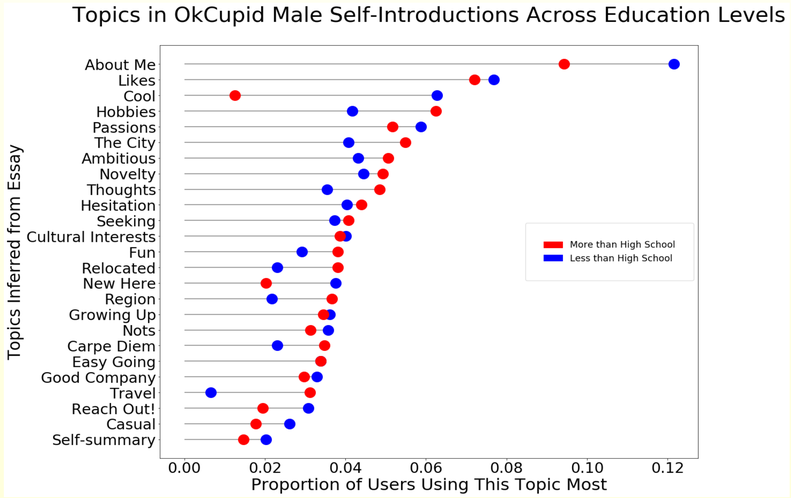
•This paper relies entirely on unsupervised machine learning algorithms applied on human-entered text data. Thus, it first required extensive cleaning.

## Approach 1: Concrete Representation

Topic Modeling using Nonnegative Matrix Factorization (NMF) generates 25 discernible topics, roughly following the identification used by Shishido et al (2016).

For each category of our variable of interest, (in this case, education) we then check for proportions of male users who use that topic most among all topics in their profile.

## Different Education, Same Topic Choice




There is no significant difference in the way in which educated men (with more than a high school diploma) use any of the 25 topics vis-à-vis those who do not. This pattern of similarity holds true for other variables like height, fitness level and race.

The exact words (unigrams, bigrams and trigrams) in the topics remain the same for the sub-groups based on the variable too.

## Male Profiles are Short & Simple

Men's profiles are on average highly skewed on most measures of writing complexity. Hence, we rely on median values below

| Group | Profile Length | Long Words Used | Reading Difficulty* |
|---|---|---|---|
| General Male Users | 80 | 16 | 6.73 |
| High School Diploma or Less | 66 | 12 | 5.63 |
| More Than High School Diploma | 85 | 18 | 7.10 |

* Based on Flesch-Kincaid Reading Level.

## Approach 2: Abstract Representation

The Doc2Vec modeling strategy allows for conversion of longform text to vector spaces. This allows for a geometrical interpretation of similarity of profiles, thereby helping with k- Means Clustering.

Reducing dimensionality via PCA, we see 5 clusters emerge. However, their interpretation is subjective.

## No 'Clustering' based on Education Level

There is no statistically significant association between levels of education and cluster membership, implying no dominant strategy within each level.

## Conclusion

Based on the available evidence and methods, there does **not** appear to be any systematic trend to how male users of the selected demographics include different content in their online profiles strategically.

## Limitations

•The study relies on 1 one of the 10 essays, which leaves out others aspects of self-presentation.

•We do not cover joint distributions of topics

•There are no details of interactions based on these profiles, so we cannot know which ones 'worked'

•Due to absence of any follow-up interviews, it is impossible to measure whether the specific choice of words was aimed at authenticity, or matches with an awareness of 'relationshopping'

## References

Emory, Leah(2017), Bustle. 'How Many People Who Meet On Dating Apps Get Married? Swiping Isn't Just For Hookups',

Fiore, A. T., Taylor, L. S., Mendelsohn, G. A., & Hearst, M. (2008, April). Assessing attractiveness in online dating profiles.

Heino, R. D., Ellison, N. B., & Gibbs, J. L. (2010). Relationshopping: Investigating the market metaphor in online dating. Journal of Social and Personal relationships, 27(4), 427-447.

Shishido, Juan, Jaya Narasimhan, and Matar Haller. "Tell Me Something I Don't Know: Analyzing OkCupid Profiles." (2016).