# System U: Automatically Deriving Personality Traits from Social Media for People Recommendation

Hernan Badenes[1], Mateo N. Bengualid[1], Jilin Chen[2], Liang Gou[2], Eben Haber[2], Jalal Mahmud[2], Jeffrey W Nichols[2], Aditya Pal[2], Jerald Schoudt[2], Barton A Smith[2], Ying Xuan[2], Huahai Yang[2], Michelle X. Zhou[2]

[1]IBM, Buenos Aires, Argentina, {hbadenes, benguali}@ar.ibm.com

[2]IBM Research – Almaden, San Jose, CA, USA, {jilinc, lgou, ehaber, jumahmud, jwnichols, apal, jschoudt, barton.smith, yxuan, hyang, mzhou }@us.ibm.com
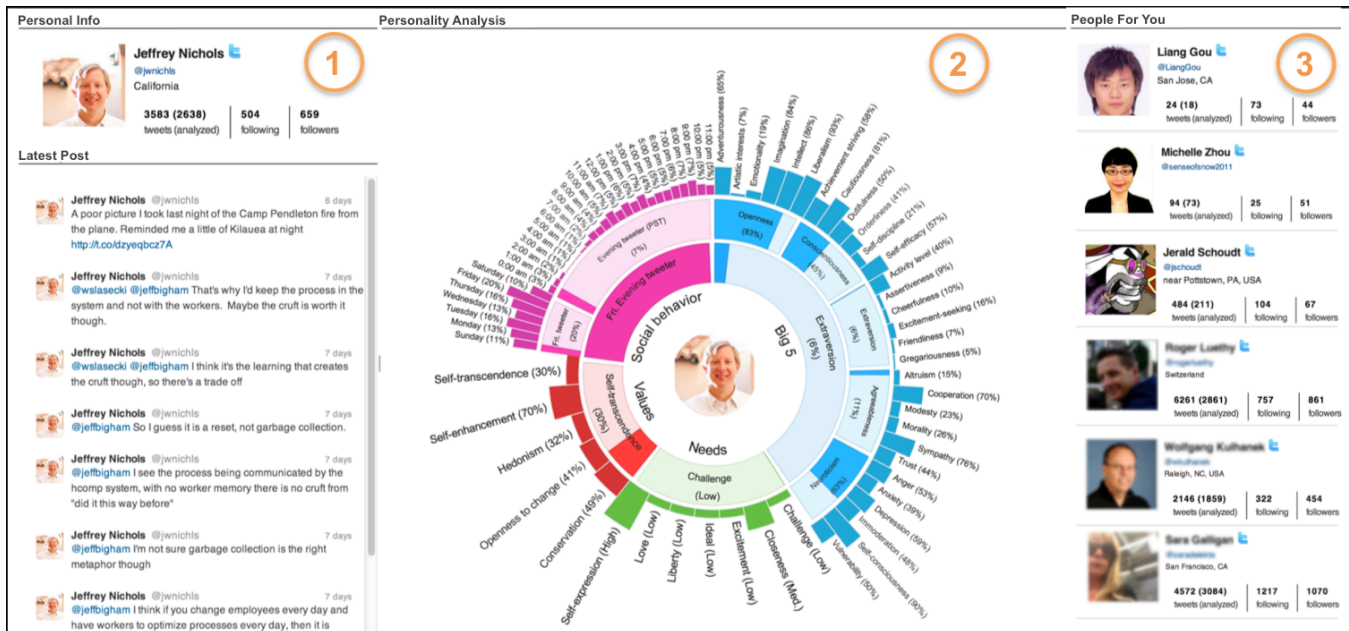
Figure 1: System U interface: (1) Basic information of a Twitter user and recent tweet posts from this user; (2) A visualization of personality portrait, including Big Five personality [8], fundamental needs [7], and human values [1]; (3) A panel showing recommended people for different tasks.

ABSTRACT

This paper presents a system, System U, which automatically derives people's personality traits from social media and recommends people for different tasks. The system leverages linguistic signals appearing in a person's social media activities to compute the personality portraits including Big Five personality, fundamental needs and basic human values. This system and technology can be used in a wide variety of personalized applications, such as recommending people to answer questions.

Categories and Subject Descriptors

H.1.2 User/Machine Systems: Human factors.

Keywords

System U; Big Five personality; basic human values; fundamental needs; social media; recommendation system.

1. INTRODUCTION

People have different personality traits, such as openness to experience new activities, needs of closeness with families. Psychology and Behavioral Economic studies show that personality influences a person's behavior in the real world, e.g. occupational proficiency [2] and economic decisions [7]. However, it is impractical to obtain these traits for a large population with traditional psychometric surveys and tests.

Fortunately, a number of studies show personality traits can be detected from individuals' language use based on psycholinguistic analysis [2][3][8]. For example, the agreeableness of a person is highly correlated with their language choices of inclusive words, such as with, together, and us. Meanwhile, the public footprints left in social media offers us an opportunity to derive people' personality traits at scale.

In this paper, we demonstrate a system, System U, which automatically derives people's personality traits from social media with psycholinguistic analysis. We currently focus on three types of traits, including Big Five personality [2][8], fundamental needs [7] and human values [1]. We present the system architecture, key technologies and user interfaces. We also show how these traits can be used to recommend people who share similar interests.

2. SYSTEM OVERVIEW

2.1 System Architecture

Figure 2 shows the system architecture of System U. The social media crawlers first obtain people's public digital footprints (tweets, blogs and Facebook updates etc.), from various social media sources. This content is saved to a storage layer for later analysis. The module of people analytics then computes three types of traits from the textual content, and there are three models including Big Five personality, basic human values, and fundamental needs in this module. With the computed traits, the module of people recommender can suggest a group of people who can help on certain tasks. Also, some other personalization modules can be plugged into this to offer individualized services by leveraging the derived personality traits. Finally, the user interface allows users to interact with the personality traits and personalized services through interactive visualization.
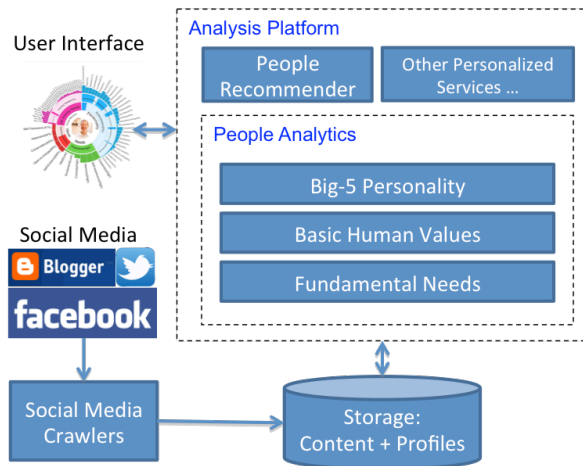
Figure 2. The system architecture of System U.

2.2 People Analytics

We currently model three types of personality traits with a lexicon-based, psycholinguistic analysis, including Big Five personality, basic human values, and fundamental needs. The first model is Big Five personality that consists of five unique and enduring personality characteristics, including *Openness*, *Conscientiousness*, *Extraversion*, Agreeableness, and *Neuroticism*. It influences one's life from many aspects, including job proficiency and team performance [2][8]. The second model concerns a person's fundamental needs with 12 dimensions (6 of them are currently shown in System U), which often influences one's brand and product choices [7]. The last one is the basic values model capturing people's beliefs and motivators along five major dimensions [1].

A lexicon-based analysis is used to build the trait models by analyzing correlations between personality traits and the usage of certain words or word categories. Specifically, our Big Five personality and basic values [6] use words and word categories derived from the LIWC dictionary [6], and our needs model [7] uses a custom dictionary. We developed a custom dictionary with a hybrid empirical and computational approach. The idea is to analyze the correlation of set of words extracted from written samples with each trait dimension by feature-based, statistical models. The extracted words construct a custom dictionary. More details can be found in [1][7].

2.3 People Recommender

A people recommender is designed to suggest a group of people who can help on certain tasks by assessing the task suitability of an individual. These tasks include finding

- Answerers to questions: We built a smart social QA system that automatically routes questions to suitable employees who are willing, able, and ready to provide answers. Our optimization-based approach also recommends both top-matched active (seed) and inactive (prospect) answerers for a given question [5].

- Information Spreaders: We built a recommendation system that actively identify and engage the right strangers at the right time on social media to help effectively propagate intended information within a desired time frame [4].

2.4 User Interface

The user interface of System U is shown in Figure 1. It includes three panels of user twitter information, derived personality portrait, and recommended people. Specifically, Panel 1 shows basic information of a Twitter user, including recent tweet posts from this user. Panel 2 shows a visualization of personality portrait, including Big Five personality, fundamental needs, human values and social activities. The visualization enables users to expand a trait of interest to see subtraits. For example, a user can have a better understanding of "*Neuroticism*" by expanding this sector to look at its subtraits, such as "*Anxiety*", "*Anger*", "*Depression*" and "*Self-consciousness*". Panel 3 shows a list of people who are recommended for certain tasks.

3. CONCLUSION

In this paper, we present a system, System U, which automatically derives people's personality traits, including Big Five personality, basic human values, and fundamental needs, from social media. We also demonstrate the capability of applying such technology for various potential personalized services such as recommending people to answer questions.

4. REFERENCES

[1] Chen, J., Hsieh, G., Mahmud, J., & Nichols, J.:. Understanding individuals' personal values from social media word use. *CSCW'14*.

[2] Golbeck, J., Robles, C., Edmondson, M., &Turner, K.: Predicting personality from twitter.*SocCom'11*, 149-156.

[3] Gou, L., Zhou, M. X., & Yang, H: KnowMe and ShareMe: Understanding Automatically Discovered Personality Traits from Social Media and User Sharing Preferences. *CHI'14*.

[4] Lee, K., Mahmud, J., Chen, J., Zhou, M.X., Nichols, J.: Who will retweet this?: Automatically Identifying and Engaging Strangers on Twitter to Spread Information. *IUI'14*: 247-256

[5] Luo, L, Wang, F., Zhou, M.X., Pan, Y., Chen, H.: Who have got answers?: growing the pool of answerers in a smart enterprise social QA system. *IUI'14*: 7-16.

[6] Pennebaker, J. W., Francis, M. & Booth, R.: Lin-guistic inquiry and word count: LIWC 2001.

[7] Yang, H., & Li, Y. Identifying user needs from social media. *IBM Tech. Report'13*.

[8] Yarkoni, T.: Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, 44, 363-373, 2010.