

# MyPersonality in Recommender System

Junhua Chen  
Wei Zeng  
2015.10.15

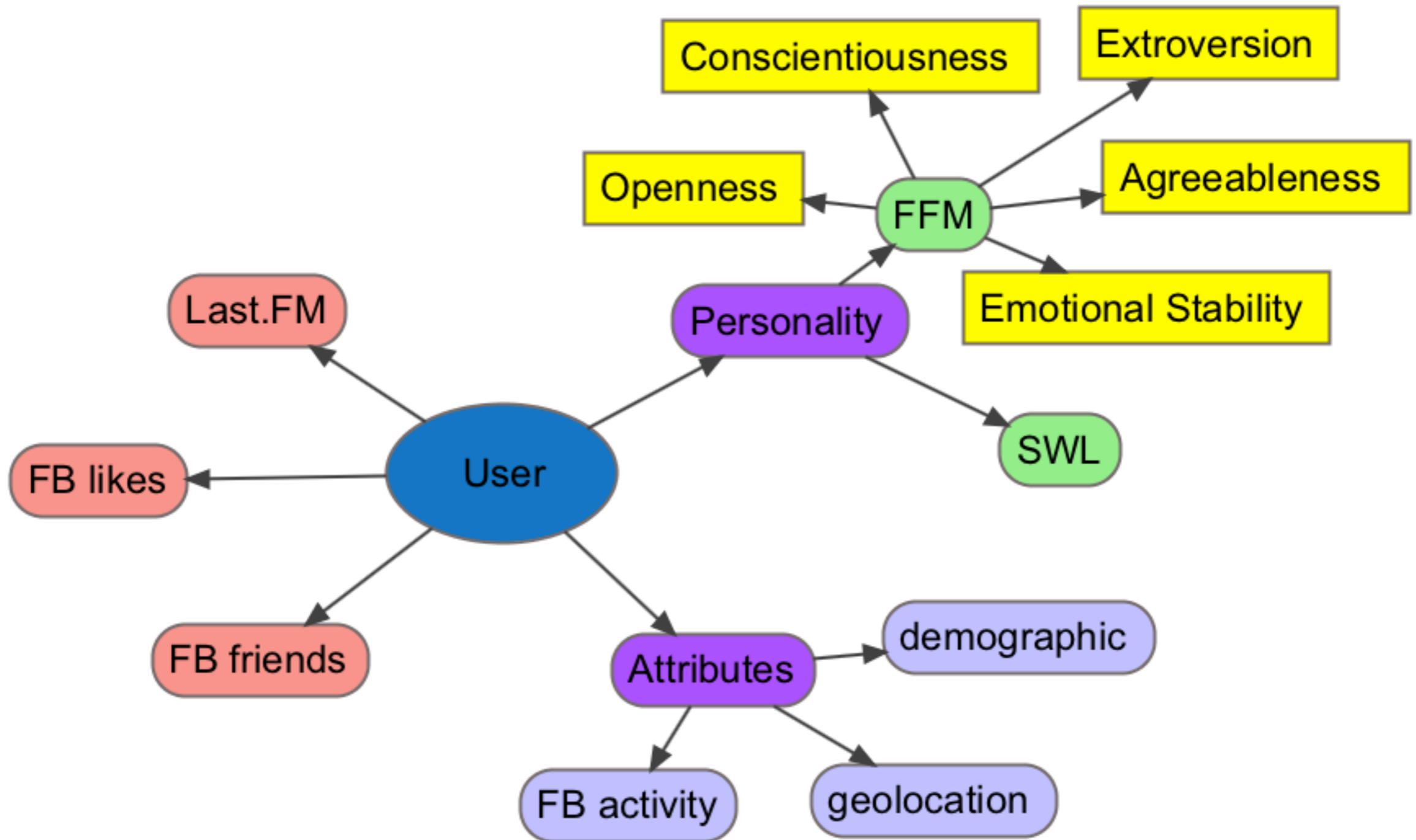


- What is MyPersonality?
  - MyPersonality is a project (<http://mypersonality.org/>) conducted by Michal Kosinski and David Stillwell. It collects users' digital footprints (e.g., Facebook likes, friends, and Last.FM) and users' demographic and psychological profiles (e.g., BigFive, IQ, SWL-satisfaction-of-life).
- What is Recommender System?
  - RSs concept is wide, what we are focusing on is to recover potential link given a user-item sparse matrix.



# Contents

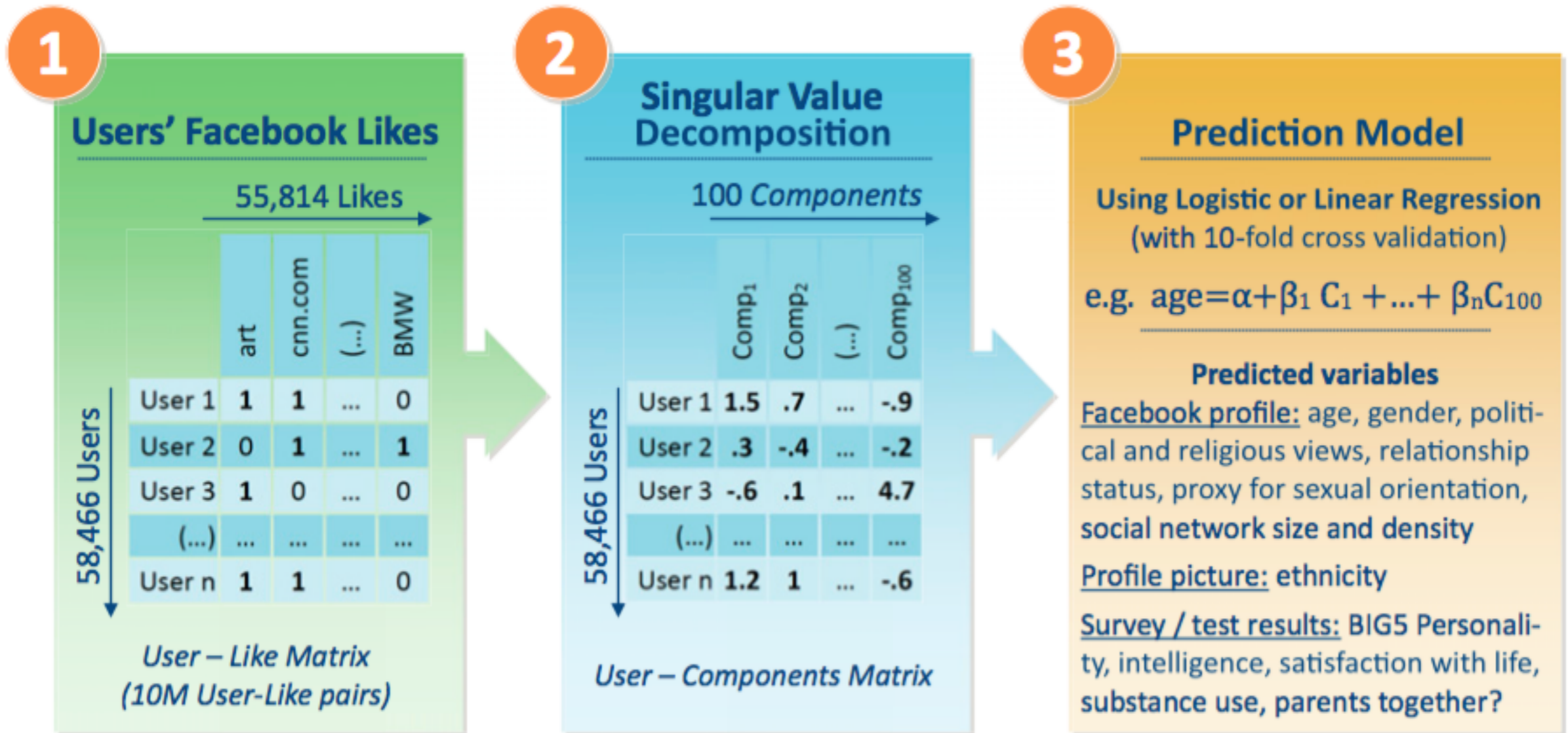
- MyPersonality dataset introduction
- Related works on Personality
  - Michal Kosinski and David Stillwell 's works.
  - BigFive in recommender system.
- What we are planing to do in MyPersonality
  - consider different forms of personality in RSs.

# MyPersonality Dataset

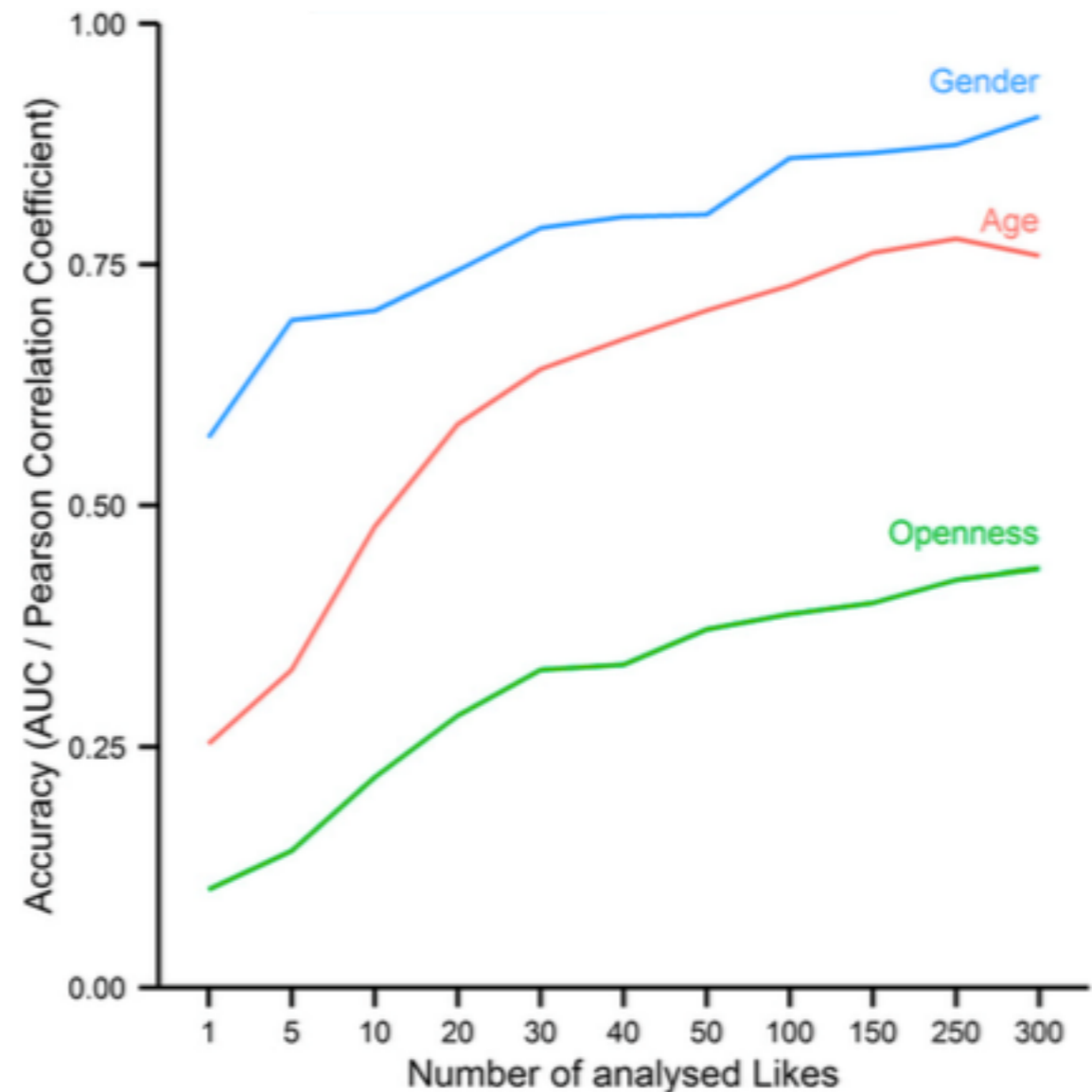
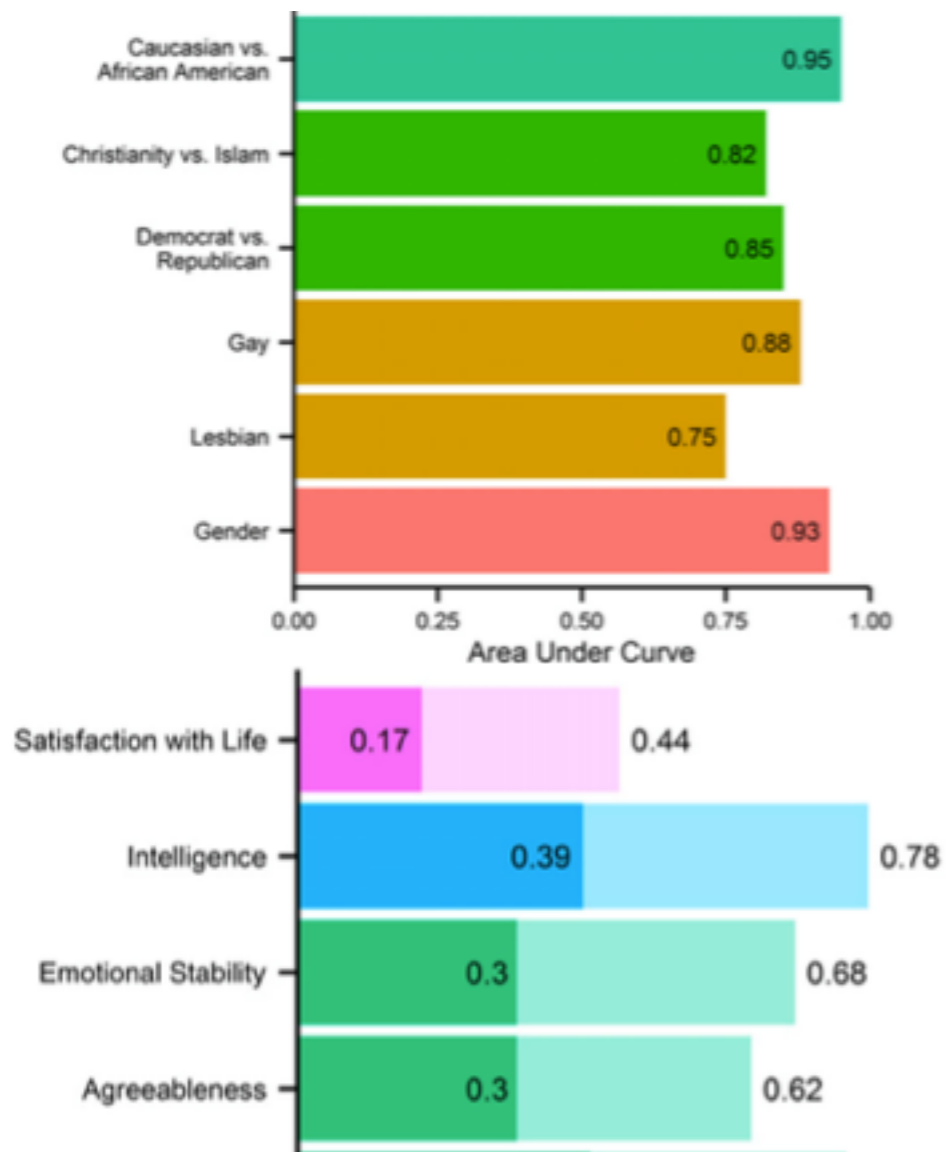


# Kosinski & Stillwell 's works

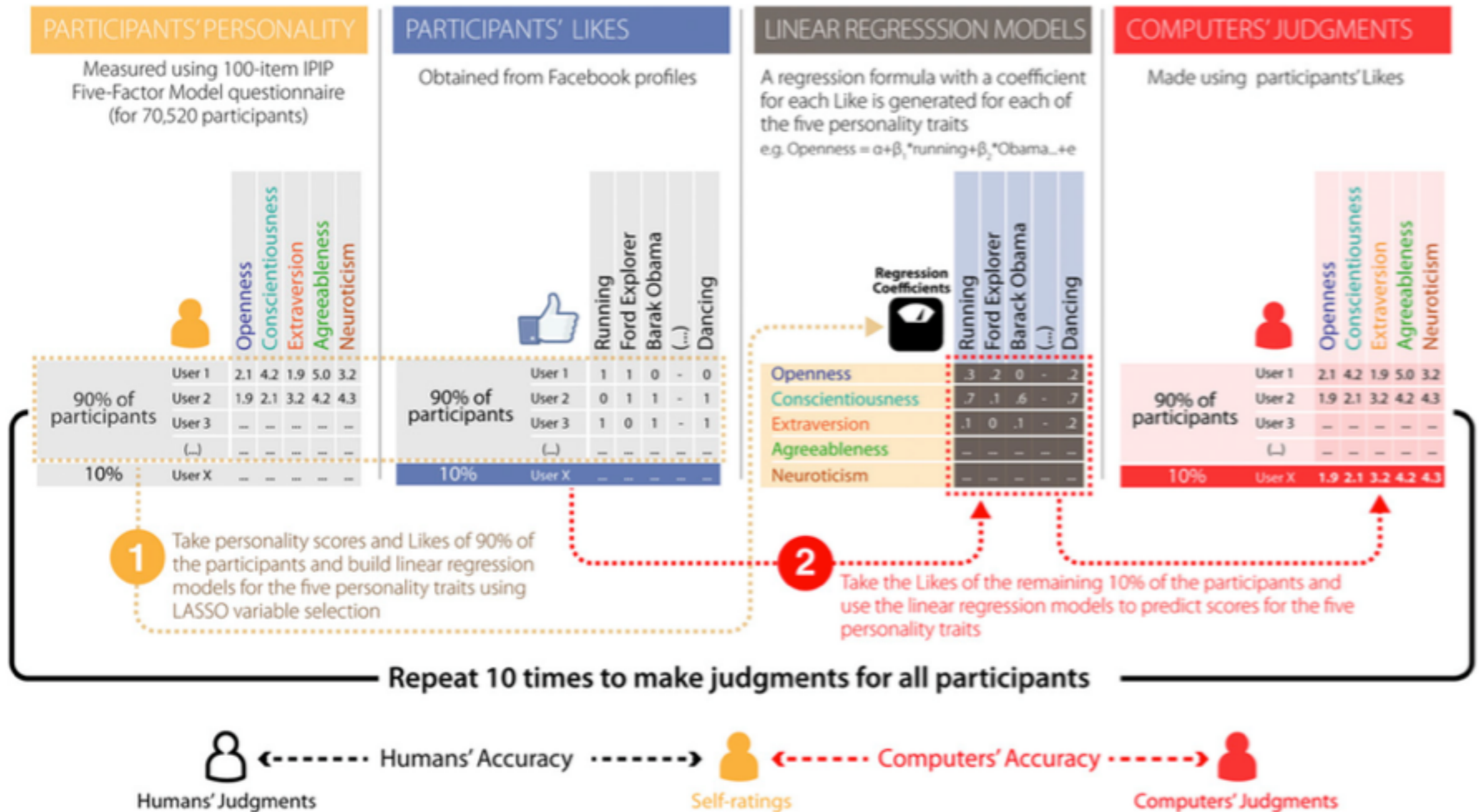
- *Private traits and attributes are predictable from digital records of human behavior (PNAS,2013)* 
- **brief:** the study extracted features out of FB-likes data(using SVD) and use the features to predict the users' traits (with linearModel/logisticModel).
- *Computer-based personality judgments are more accurate than those made by humans (PNAS,2014)* 
- **brief:** the study compares the accuracy of personality judgment between computer models and humans.



- Users' FB Likes - user-like sparse matrix (0.3%).
- SVD (math), 100 components as users' features.
- Use the 100-D features from #2 to train predictor — linear model for regression & logistic for classification.

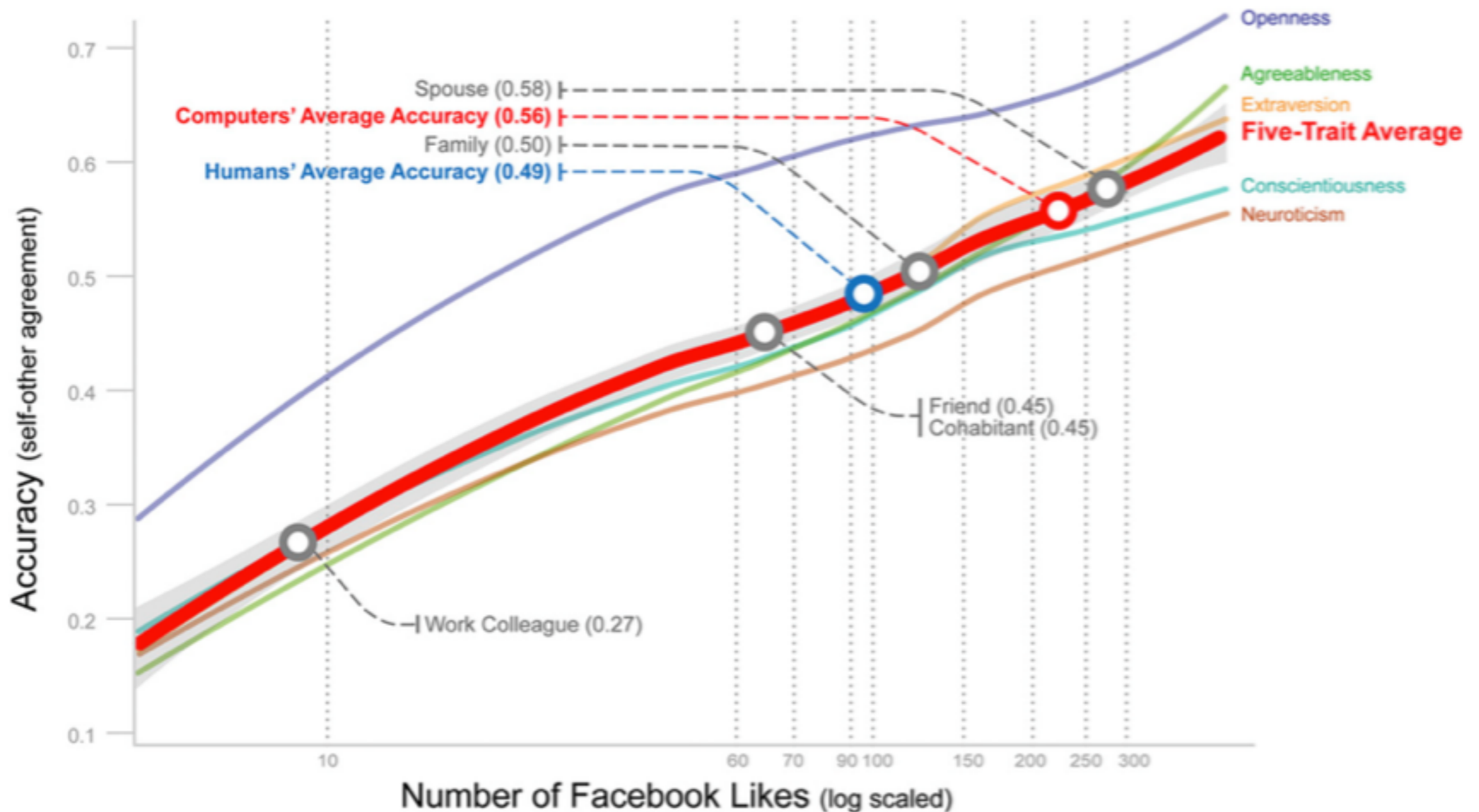


- Classification (top-left) & regression (bottom-left) results, AUC. The transparent bars indicate the questionnaire's baseline accuracy (test-retest).
- Predictive Power of Likes (right).



- Using 90% of the users (with BigFive and FB-Likes) to train a Linear Model (Lasso), predict the BigFive of users in the rest 10%. The BigFive of these 10% users is called computer-based personality.
- repeat 10 times to get all computer-based personality for all users.





- Human-based personality is judged by couple (friends, spouse, colleague) — circle on the red line.
- Accuracy (correlation) made by computer-based (red line) surpass human-based (blue circle) with 100 likes, and the accuracy of computer-based is monotony.

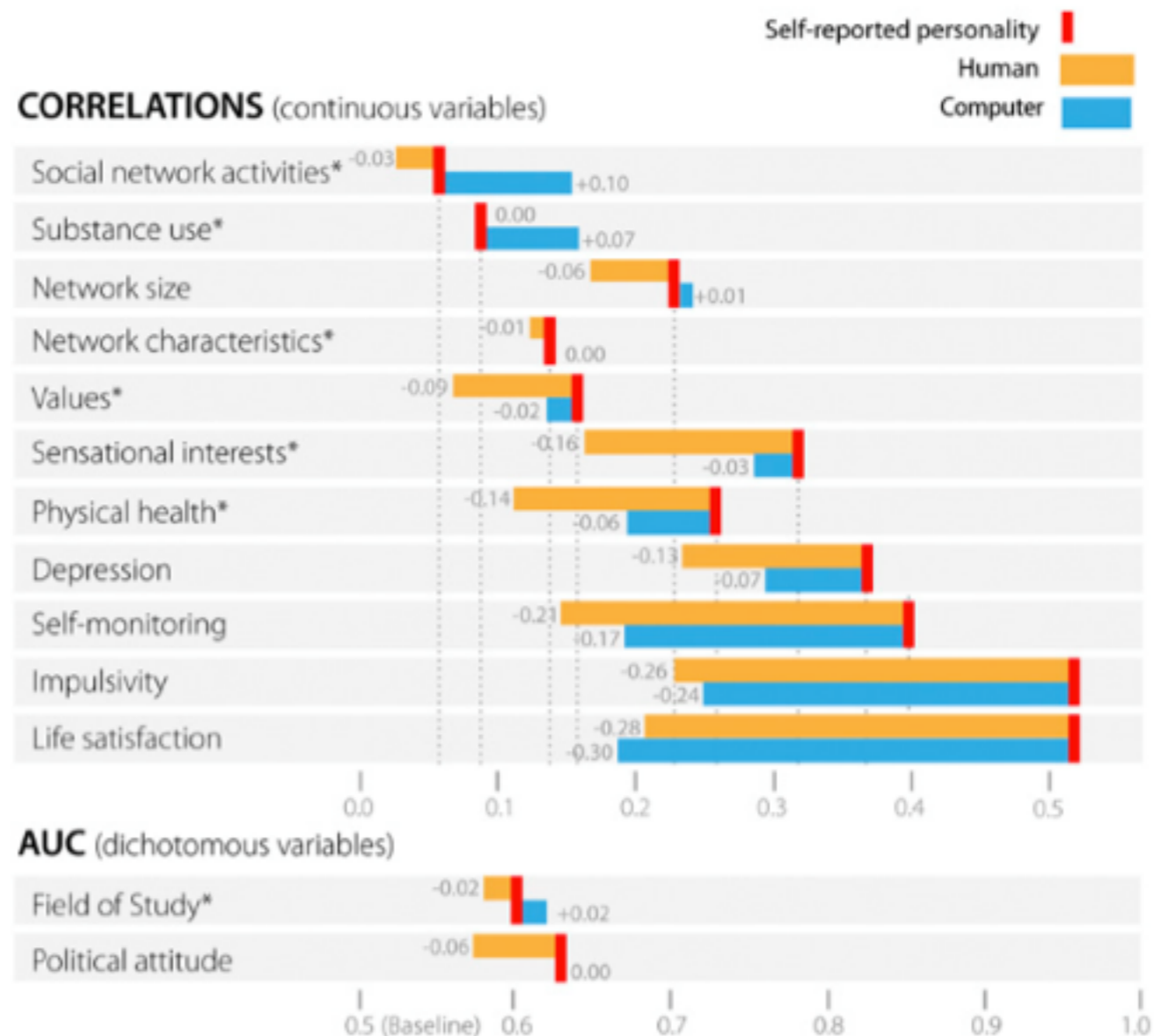
- **Interjudge Agreement**

- two judges that agree with each other are more likely to be accurate than those that do not. (table on the right)


	human	computer
r-inter	0.62	0.38

- **External Validity**

- how well a judgement predicts external criteria, such as real-life behavior (figure on the right).
- *the fact (in the figure) computer beats human and even self-reported in “Social network activities” somewhat motivate our work.*



# BigFive in RSs

- *M Braunhofer, M Elahi, and F Ricci, User Personality and the New User Problem in a Context-Aware Point of Interest Recommender System, ICTT, 2015.* 
- brief: the paper integrate BigFive (Openness, Agreeableness, Extraversion, Conscientiousness, and Neuroticism) into a matrix factorization framework.

- a matrix factorization framework in this paper.

$$\hat{r}_{ui} = \mu + b_u + b_i + q_i^T \left( p_u + \sum_{a \in A(u)} y_a \right)$$

- Integrate BigFive personality — e.g., if user  $u$  is high on Extraversion and low on the other four, the equation will be:

$$\hat{r}_{ui} = \mu + b_u + b_i + q_i^T (p_u + y_{ope\_low} + y_{con\_low} + y_{ext\_high} + y_{age\_low} + y_{neu\_low})$$

# Planing to do in MyPersonality

- Users' personalities provide a great deal of information about users' interests. Therefore, the performance of a recommender system could be improved when taking into account users' personalities. In our project, we plan to find the answer to two questions:
  - Q1: How to use the user personality in a recommender system?
  - Q2: Which one is better for the human-based, computer-based personality.

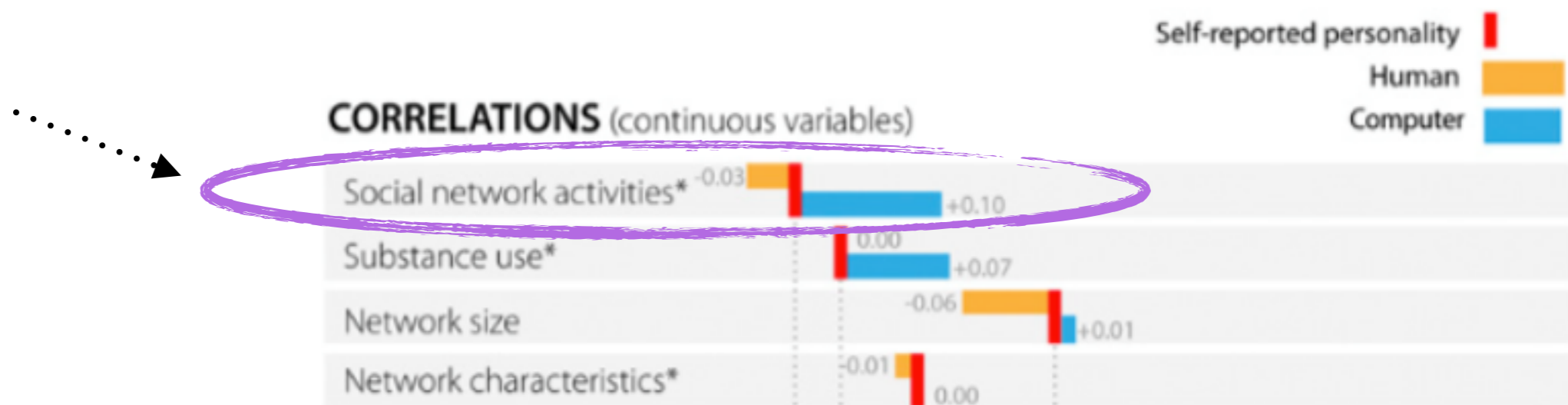
# 4 Q1

- Luckily, M Braunhofer had bring us a feasible plan, but the equation treat the BigFive personality as tag. Yet those personality in MyPersonality dataset is continued. i.e., we need to mod the framework for MyPersonality.

$$\hat{r}_{ui} = \mu + b_u + b_i + q_i^T \left( p_u + \frac{1}{Norm} \sum_{a \in A_u} y_a \right)$$

# 4 Q2

- Computer-based vs human-base, with decent among of likes data, computer actually beats human (Kosinski and Stillwell). But their 'External Validity' experiments shows that in 'socail-net activity', computer personality is even better than self-reported one, which gives us window to consider why and how computer-based outperform self-reported personality in RSs.



Questions



Thank You