Prediction of Human Personality Traits From Annotation Activities

Nizar Omheni[1], Omar Mazhoud[1], Anis Kalboussi[1] and Ahmed HadjKacem[2]

[1]*Higher Institute of Computer and Management, University of Kairouan, ReDCAD Laboratory, Khmais Alouini Street 3100, Kairouan, Tunisia*

[2]*Faculty of Economics and Management, University of Sfax, ReDCAD Laboratory, Road of the Airport Km4 3018, Sfax, Tunisia*

Keywords: Annotation, Personality, Big Five Personality Model.

Abstract: We show how reader's annotation activity captured during an active reading session relates to their personality, as measured by the standard Five Factor Model. For 120 volunteers having usually the habit of reading, we gather personality data and annotation practices. We examine correlations between readers personality and such features of their annotative activities such as the total number of annotation acts, average number of annotation acts, number of textual annotation acts, number of graphical annotation acts, number of referential annotation acts and number of compounding annotation acts. Our results show significant relationships between personality traits and such features of annotation practices. Then we show how multivariate regression allows prediction of the readers personalities traits given their annotation activities.

1 INTRODUCTION

Studying human activity and interaction with technology has grown dramatically over the last decade. Yet studying reading poses particular challenges. (Marshall,2010) reported the citation of Tzvetan Todorov, quoted by Nicholas Howe in Jonathan Boyarins compilation, the Ethnography of Reading: "Nothing is more commonplace than the reading experience, and yet nothing is more unknown. Reading is such a matter of course that at first glance it seems there is nothing to say about it". Although details of reading activity (moving eyes, writing annotation...) may tell us something about the reader. When people read and interact actively with their reading materials they do unselfconscious activities which can be keys features to their personalities.

For decades, the psychologists search to understand the human personality and to find a systematic way to measure it. After several researches they show a relation of dependence between human personality traits and different behaviors. (Ryckman,2010) reported the Allport's [1] definition of personality: "personality is the dynamic organization within the individual of those psychophysical systems that determine his characteristic behavior and thought". Thus, in Allports view human behavior is really controlled by internal forces known as the personality traits.

This paper attemps to bridge the gap between reading activity research and personality research through reader's annotation practices. Our core research question asks whether annotation activity can predict personality traits. If so, then there is an opportunity to use a natural human practice as a new source to better understand the reader personality.

Several works has shown the opportunity of predicting user personality using the information people reveal in their online social profile (Twitter, FaceBook) (Bachrach et al,2012) (Golbeck et al,2011). They refer to what people share, self-description, status updates, photos, tags, etc. We pretend that annotative activity is more spontaneously and natural practice and it can reveal something about human personality.

Personalization attracted increased attention in many areas. So the need to predict personality traits increases over time mainly when several research has shown the link between personality traits and success in human relationship and practices (Barrick and Mount,1991) (Eswaran et al,2011). By nature an introverted person is not interested to make so much relation with other while an extravert person do. Actually, certain developed recommendation systems con-

[1]Gordon Willard Allport (November 11, 1897 October 9, 1967) was an American psychologist. He was one of the first psychologists to focus on the study of the personality, and is often referred to as one of the founding figures of personality psychology.

Omheni ., Mazhoud O., Kalboussi A. and HadjKacem A..
Prediction of Human Personality Traits From Annotation Activities.
DOI: 10.5220/0004801302630269
In *Proceedings of the 10th International Conference on Web Information Systems and Technologies* (WEBIST-2014), pages 263-269
ISBN: 978-989-758-024-6

sider the personality traits as key feature of recommendation (Nunes et al,2008).

The paper is structured as follows: In Section 2, we present background on the Big Five personality index. Then we present our experimental setup and methods. In the third section we present the results on correlation between each annotative activity feature and personality factor. Next, we show how multivariate regression allows prediction of annotators personalities traits given their annotation activities. We conclude with a discussion of the possible implications that this work has for such domains of application.

2 BACKGROUND AND RELATED WORK

2.1 The Big Five Personality Model

The big five personality traits are the best accepted and most commonly used scientific measure of personality and have been extensively researched (Peabody and De Raad,2002). That personality is well described as five traits was discovered through the study of the adjectives from natural langage that people used to describe themselves and then analyzing the data with a statistical procedure known as factor analysis that is used to reduce lots of information down to its most important parts. In the following we cite a brief explanation of the five personality traits.

2.1.1 Openness to Experience

Openness includes traits like imagination, appreciation for art, depth of emotions, adventure, unusual ideas, intellectual curiosity, and willingness to experiment. People who score high in openness like usually to learn new things and enjoy new experiences.

2.1.2 Conscientiousness

Conscientiousness includes traits like orderliness, selfdiscipline, deliberateness, and striving the achievement. People that have a high degree of conscientiousness are planned, have the tendency to act dutifully, have the sense of responsibility and competence.

2.1.3 Extraversion

Extraversion includes traits like energy, positive emotions, surgency, assertiveness, sociability and talkativeness. Extraverts people get their energy from interacting with others, while introverts get their energy from within themselves.

2.1.4 Agreeableness

Agreeableness includes traits like trust in others, sincerity, altruism, compliance, modesty and sympathy. People that have high degree of agreeableness are friendly, cooperative, and compassionate, while people with low agreeableness may be more distant.

2.1.5 Neuroticism

Neuroticism relates to ones emotional stability and degree of negative emotions. This dimension measures the people degree of anxiety, angry, moodiness, and the sensitivity to stress. People that score high on neuroticism often experience emotional instability and negative emotions.

2.2 Related Work

In (Burger,2011) view the personality is a ”consistent behavior patterns and intrapersonal processes originating within the individual”. Trait psychologists assume that personality is relatively stable and predictable (Burger,2011). So, several research work has been done with personality traits as it influences human decision making process and interests. (Nunes et al,2008) pionneered the model and implement of personality traits in computers. Indeed, (Nunes et al,2008) propose to model the user's traits in a profile which they called User Psychological Profile - UPP. In order to fill in the profile UPP the authors utilised an online tool called the NEO-IPIP[2] inventory based on 300 items. Through user's answers to NEO-IPIP inventory the authors are able to predict the user personality. Through their experimentation (Nunes et al,2008) try to prove that Recommender Systems can be more efficient if they use the User Psychological Profile (UPP). Although the authors follow an explicit way to predict the user traits, the results presented in (Nunes,2008) are fruitful. (Tkalcic et al,2009) propose a personality-based approach that is based on the big five model for collaborative filtering Recommender Systems. In fact, the authors calculate the user personality scores by means of a questionnaire. Then they measure the user similarity, that is based on personality, that yields a list of close neighbours. This list is used after as a database to compile a personalized list of recommended items.

[2]The NEO-IPIP is a computer based Personality Inventory, able to measure people Personality Traits created by John Johnson (Johnson.).

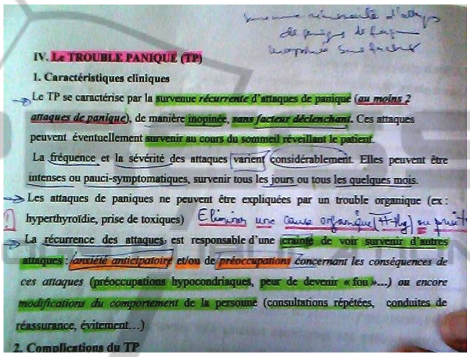
(Roshchina et al,2011) propose personality-based recommender system which the aim is to select for the user reviews that have been written by like-minded individuals. The user similarity is calculated based on the presonality traits according to the Big Five model. The authors predict the presonality traits based on linguistic cues collected from the user-generated text. (Bachrach et al,2012) and (Golbeck et al,2011) show the relationship between personality traits and various features of social media (FaceBook, Twitter). Their findings prove the possibility to predict accurately the user's personality through the publicly avaible information on their social network profile.

As it is mentioned above most of works has been done with the "Big Five" model of personality dimensions which has emerged as the most well-researched and regarded measures of personality in last decades. The best of our knowledge, our work is among the first to look at the relationship between annotation activity and personaltiy traits. Much works try to predict personality from what the user offer consciously (answers to a questionnaire, informations available on public social profile...). Despite the fact that the findings of these researches are fruitful we believe that predicting personality from annotative activity is more credible as the annotation is defined as "a basic and often unselfconscious way in which readers interact with texts" (Marshall,2010).

Due to the spontaneously and unselfconscious aspects of annotation we are interested to predict personality from this potential source of knowledge.

3 DATA COLLECTION

We consider group of 120 volunteers. The subjects selected were recruited with respect to certain criterias. Infact, the age of our volunteers should be between 18 and more and they should be academic people. In our sample we have the two sex (44 women and 76 men). Another criteria for selection , we asked if the volunteer has the habit of reading and does he annotated his documents frequently. If all these conditions exist the subject can be selected to our experimentation.
Each subject was instructed to answer a standard Five Factor Model questionnaire (the NEO-IPIP Inventory). Then, he obtained a feedback regarding his personality based on his responses. This step gives us the personality scores based on the Big Five Model for each volunteer. To associate personality scores to subjects annotative activities, we gather annotation practices for each people. Here, we collect documents annotated in a spontaneous and natural way. So we asked, first of all, if the subject had a document annotated previously (academic course, book,...). If not we asked him what topics interest him, then we give him an article with few pages to not weary him.
We are very careful to the comfortability of the volunteers during the experience to guarantee their spontaneous and natural reactions. Thus they are free to choose places and conditions to read and annotate the documents and they have enough time to do. The strategy followed give us fruitful results. Infact, the different subjects (who have not a document annotated previously) interact actively with the reading materials in view of the feel of comfortableness and the interest to the document read.


IV. Le TROUBLE PANIQUE (TP)

1. Caractéristiques cliniques

Le TP se caractérise par la survenue *récurrente* d'attaques de panique (*au moins 2 attaques de panique*), de manière inopinée, *sans facteur déclenchant*. Ces attaques peuvent éventuellement survenir au cours du sommeil réveillant le patient.

La fréquence et la sévérité des attaques varient considérablement. Elles peuvent être intenses ou pauci-symptomatiques, survenir tous les jours ou tous les quelques mois.

Les attaques de paniques ne peuvent être expliquées par un trouble organique (ex : hyperthyroïdie, prise de toxiques)

La récurrence des attaques est responsable d'une crainte de voir survenir d'autres attaques : *anxiété anticipatoire* et/ou de *préoccupations concernant les conséquences de ces attaques* (préoccupations hypocondriaques, peur de devenir « fou »...) *ou encore modifications du comportement* de la personne (consultations répétées, conduites de réassurance, évitement...)

2. Complications du TP

Figure 1: Annotation practices of a reader.

3.1 User Annotation Activity

(Marshall,2010) defines the annotation as "a basic and often unselfconscious way in which readers interact with texts". We mean by annotation the act to add a critical or explanatory notes to a written work, to highlight a passage, to write down, and so on marks the reader makes on a page during his reading activity. To fulfil our experimentation, we ask each subject to give us an annotated paper document. Then, we analyse the readers annotations to extract some features. We started by classifying annotations in the three general categories cited by (Agosti and Ferro,2003) : graphical annotation acts, textual annotation acts, and reference annotation act all depends to the materialization sign of the semantics of the annotation added to the annotated document. Then, for each reader, we collect a simple set of statistics about their annotative activity. These included the following:

1. Total Number of Annotation Act (TNAA)
2. Average Number of Annotation Act (number of annotation acts per a single annotated page)(ANAN)
3. Number of Graphical Annotation Act (NGAA)
4. Number of Textual Annotation Act (NTAA)

5. Number of Reference Annotation Act (NRAA)
6. Number of compounding Annotation Act (textual sign, graphic sign and reference sign of annotation act can be compounded together in order to express complex meanings of annotation)(NCAA).

This set of statistics tends to characterize quantitatively the reader's annotation practices. Next we run a Pearson correlation[3] analysis between subjects' personality scores and each of the features obtained from analyzing their annotative activities.

4 PERSONALITY AND ANNOTATION FEATURES CORRELATION

We study the Pearson correlation between subjects' personality scores and each of the features obtained from analyzing their annotative activities. We report the correlation values in table I. Those that were statistically significant for $p < 0.05$ are bolded.

Table 1: Pearson correlation values between annotation features scores and presonality scores.

	Open.	Consc.	Extra.	Agree.	Neuro.
TNAA	-0,059	0,128	-0,138	0,089	**-0,287**
ANAA	0,003	0,080	**-0,210**	0,163	**-0,183**
NGAA	-0,067	0,040	-0,130	0,105	**-0,207**
NTAA	0,001	**0,182**	0,040	0,085	**-0,211**
NRAA	-0,075	0,045	-0,122	0,077	**-0,207**
NCAA	-0,059	-0,012	-0,147	0,014	**-0,219**

We found in our analysis fewer significant correlations, but we believe a larger sample size would produce much better results. However, the results we obtained even with a small sample show promise that the annotative activity can be useful for computing such personality traits. In fact, table I shows significant correlations for Neuroticism, Conscientiousness, and Extraversion traits. We need larger sample to verify the inference of the other traits from peoples annotations.

Next, we present the scatter plots for the most significant correlations between annotation practices features and personality traits. These plots presenting the relationship between annotative activity features and human traits, where horizontal axis represents the average personality trait scores and the vertical axis represnts the annotative activity feature values.

[3]Pearson's correlation $r \in [-1, 1]$ measures the linear relationship between two random variables.

4.1 Conscientiousness

As presented in table I Conscientiousness is positively related to the number of textual annotation act (fig.2). The rest of the correlation values are not considerated because of p-value > 0.05. But this is not a reason to reject definitively the rest of annotation features as a larger sample size may produces other significant correlations.
The considered correlation may indicates that conscientious people are interested to use textual annotation acts. Infact, conscientious individuals are prudent which means both wise and cautious, better organized and they avoid acting spontaneously and impulsively. Thus, it may be the case that people who have high degree of conscientiousness are interested to use textual annotation more than other annotation acts as it demands more reflexion, reasoning and cognitive effort.

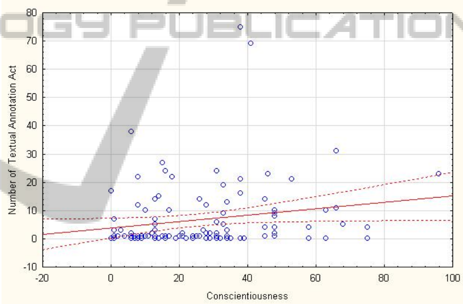


Figure 2: Scatter Plot showing Number of Textual Annotation Act against Conscientiousness scores.

4.2 Extraversion

According to results shown in table I, Extraversion is negatively correlated with the average number of annotation act (fig.3). The rest of the correlation values can be probably significant with a larger sample size. We can interpret the regression fit shown in figure 3 as follow: The fit is correlated negatively which is not surprising as extraversion is marked by pronounced engagement with the external world where extraverts tend to be energetic and talkative while introverts are more likely to be solitary and reserved. Thus, it may be the case that reading and annotation is an intimate activities, we do it in private, so people who are socially active are less willing to practice annotation.

4.3 Neuroticism

Table I shows that neuroticism is negatively correlated with all the features of annotation activity. Here,

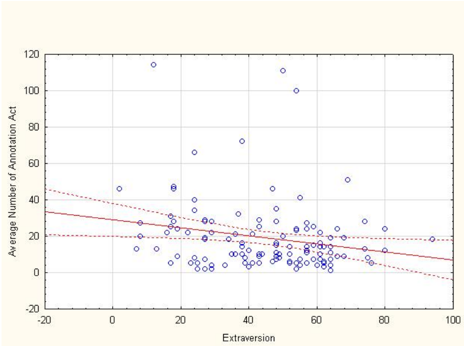


Figure 3: Scatter Plot showing Average Number of Annotation Act against Extraversion scores.

the chosen sample size is sufficient to have significant correlations for all the annotation features. The different correlation values are very significant which can show the sensitivity of annotation practices to the neuroticism trait.
In other hand, one possible explanation for these correlations is that more Neurotic people are emotionally reactive and they experience negative emotions for unusually long periods of time which can diminish the neurotics ability to think clearly and make decisions. Thus those who score high on Neuroticism are less eager to use annotation act as they can not actively and critically engaging with the content for a long periods of time.

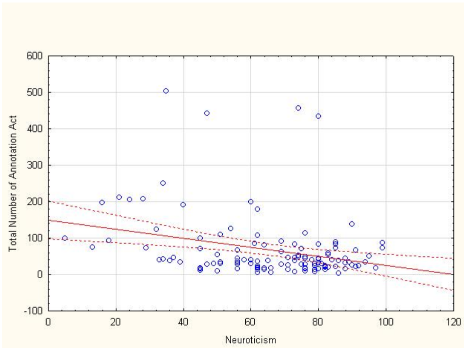


Figure 4: Scatter Plot showing Total Number of Annotation Act against Neuroticism scores.

4.4 Openness and Agreeableness

Unfortunately, the correlation values related to the Openness and Agreeableness traits are very low. But we can not reject definetly the hypothese of prediction of these traits from annotation activity. We may obtain significant values if we increase the sample size. As an exemple, the p-value of the regression fit of the Average Number of Annotation Act against the Agreeableness traits is $p = 0.076$. This value can be ameliorated with a larger sample size.

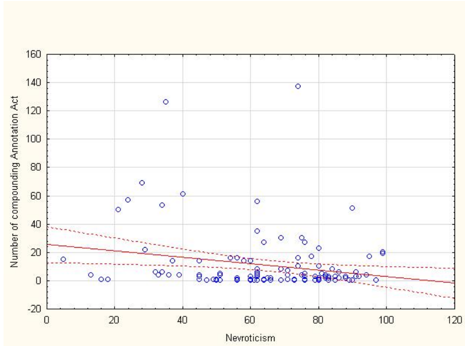


Figure 5: Scatter Plot showing Number of compounding Annotation Act against Neuroticism scores.

5 PREDICTING PERSONALITY

Previously we examined the correlations between each of Big Five personality dimension and annotative activity features. Now, we are interested to make predictions about a subject's personality based on multiple annotation features. First of all, we used the multivariate linear regression to predict each personality trait using the annotation features. Next, we used the coefficient of multiple determination R^2 to measure the strength of fit. Also, we measure the F-test to verify the statistical significance of the collective influence that have the annotation features on the personality traits. Thus, larger values of the F-test statistic provide stronger evidence against H_0[4]. To reject H_0 the value of the F-test should exceeds a critical value calculated as follow:

$$F = \frac{R^2/k}{(1-R^2)/[n-(K+1)]}$$

Where k is the number of explanatory variables in our model which corresponds to the number of annotation activity features(K=6)and n represents the sample size (n=120). So the $F_{observed}$ is compared against a critical F with 6 degree of freedom in the numerator and n-7 degrees of freedom for error in denominator. In our case $F_{critical}$=2.18 for alpha level[5] of 0.05. Results shown in table II indicate that the null hypothesis is rejected for two cases. In fact, Neuroticism and Conscientiousness can be predicted with reasonable accuracy using features of annotation activity, whereas other traits are more difficult to be predicted using annotations. Prediction regarding Conscientiousness is reasonably accurate, with R^2 value of 0.12, $F_{observed}$

[4] The null hypothesis states that there is no relationship between annotation activity features and personality traits.

[5] The alpha level is defined as the probability of what is called a Type I error in statistics. That is the probability of rejecting H_0 when in fact it was true.

value of 2.52 which exceeds the $F_{critical}$ value and P-value of 0.03 which is lower than the α value where P-value is the probability the F-test statistic is larger than the observed F-value. For Neuroticism we obtained the model with the best fit, with an R^2 value of 0.14, $F_{observed}$ value of 3.11 and P-value of 0.01, indicating quite accurate a prediction. The model for Extraversion has a lower fit and the model for Agreeableness is even less accurate. It seems that Openness is the hardest trait to predict using annotation activity features.

Table 2: Predicting personality traits using annotation activity features through multivariate linear regression.

Trait	R^2	F test	P-value
Openness	0.03	0.57	0.76
Conscientiousness	0.12	2.52	0.03
Extraversion	0.07	1.32	0.25
Agreeableness	0.05	1.03	0.41
Neuroticism	0.14	3.11	0.01

6 DISCUSSION

In this study we show that Neuroticism and Conscientiousness traits are correlated with annotation activity features. We expect a larger sample size can be helpful to verify the correlation of the other human traits to annotation practices.

Our findings are based on pen-and-paper approach which is qualified by its relative ease with which the reader may interact with a document in an intuitive and familiar manner.

Recent researchs endeavor to replace the "pen-and paper" pardigm for the annotating needs. Different systems and tools of annotation are developed such as: iAnnotate (Plimmer et al,2010), u-Annotate (Chatti et al,2006), YAWAS[6], iMarkup (2013), etc. Such tools enable readers to annotate their digital documents with free form annotations similarly to "pen-and paper" case. iAnnotate for example is an annotation tool for android system and it enables users to add annotations with the pencil, highlighter, and note tools.

Recently we intend talking about new products for reading such as the tablet. With the aid of such devices, the user may interact easily with a digital document and enter her annotations as he do in the case of "pen-and-paper". Thus, our findings is promising and original to be applied in such digital areas.

[6]Yawas is a free web annotation tool for Firefox and Chrome built on top of Google Bookmarks. Yawas enables you to highlight text on any webpage and save it in your Google Bookmarks account.

We believe it's the occasion to develop a system to attract those that have a curiosity to use annotation platforms and practices - from end users including scholars, scientists, journalists, public servants...etc. We expect our system enables readers to interact actively with their reading materials via annotation practices. Our goal is to use the traced annotations on the digital document to infer the annotator personality traits. Thus the expected system should contain free form annotation tool easy to be used, be able to infer user personality traits from captured annotation activity and refines the user traits profile by reference to new captured annotations. Based on the modelled profile we expect our system be able to offer services to users such as friendship recommendation based on similarity of users personality, customizing user interface using such predefined personas, sharing annotated documents...etc.

To achieve the personalization process we need to know certain user's features. Several research works prove that prediction of personality traits reveals a lot about a user's features. These findings was applied in several domains such as recommender systems (Nunes et al,2008) (Roshchina et al,2011) and the results is interesting.

The annotation as an unselfconscious practice constitutes a credible source of knowledge. In (Kalboussi et al,2013) the annotation activity is used to invoke the appropriate Web services to users. This proves that annotation is rich enough to be used differently.

In this paper we use annotation to infer such reader traits. That is a promising work and represents a new tendency to model user personality from human behaviour.

In other hand, let's be objective, there are some limitations to this work. The most important issue is the sample size as we expect more significant results with a larger sample. This limitation may be due to the dependency to "pen-and-paper" approach which prevents us to benefit from the population of readers over the web. In addition, people are not interested to participate in our experimentation unless they are motivated. We expect resolving these issues in future works.

Finally, our work is the first step to study the relation of reader annotation practices to human personality traits. So much perspectives can be subjects of future works such as studying factors which are likely to influence annotating behaviour such as familiarity with annotation tools and interest in the content topic. Also our research can be extended to study the possibility to predict human traits through social annotations. These avenues and others are very interesting and represent an opened future directions which needs more investigations.

7 CONCLUSION

In this paper, we have shown that such users' personalities traits can be predicted from their annotation practices. With this ability of prediction many opportunities are opened which suggests future directions in variety of areas such as user modeling, recommender systems, user interface design and so on areas relative to personalization research domain. Furthermore, this work bridges the gap between the reading and the personality research domains and it remains an open research questions to see whether personality can also be predicted using other potentially features of reading activity as well as the influence of such environmental factors on human annotation behaviour.

REFERENCES

Agosti,M. Ferro,N., 2003. *Annotations: Enriching a Digital Library*. In: Proceedings of the 7th European Conference (ECDL) Trondheim, Norway. Springer Pages 88-100.

Bachrach,Y. Kosinski,M. Graepel,T. Kohli,P. Stillwell,D.,2012. *Personality and Patterns of Facebook Usage*. In: Web Science12 Proceedings of the 3rd Annual ACM Web Science Conference. ACM New York, NY, USA, Pages 24-32.

Burger, J. M., 2011. *Personality*. Editor Wadsworth, USA.

Barrick, M. R. Mount, M.K., 1991. *The big five personality dimensions and job performance: a meta-analysis*. Personnel psychology, Volume 44, Issue 1, Pages 126.

Chatti, M. A. Sodhi,T. Specht, M. Klamma,R. Klemke,R., 2006. *u-Annotate: An Application for User Driven Freeform Digital Ink Annotation of E-Learning Content*. In: Proceedings of the Sixth IEEE International Conference on Advanced Learning Technologies (ICALT), Washington, DC, USA. Pages 1039-1043.

Eswaran,S. Aminul Islam,Md. Dayang Hasliza,M.Y., 2011. *A Study of the Relationship between the Big Five Personality Dimensions and Job Involvement in a Foreign Based Financial Institution in Penang*. In: International Business Research. Vol. 4, No. 4.

Johnson, J.A. *The IPIP-NEO: International Personality Item Pool Representation of the NEOPI-R*, viewed 8 June 2013, http://www.personal.psu.edu/ j5j/IPIP/ipipneo300.htm

Golbeck, J. Robles,C. Edmondson, M. Turner,K., 2011. *Predicting Personality from Twitter*. In: Privacy, Security, Risk and Trust PASSAT, 2011 IEEE Third International Conference on Privacy, Security, Risk, and Trust, and 2011 IEEE Third International Confernece on Social Computing.

Kalboussi, A. Mazhoud, O. Hadj Kacem,A.,2013. *Annotative Activity as a Potential Source of Web Service Invocation*. In: Proceedings of the 9th International Conference on Web Information Systems and Technologies (WEBIST), Aachen, Germany. SciTePress Pages 288-292.

Nunes, M.A. Cerri,S.A. Blanc,N., 2008. *Improving Recommendations by Using Personality Traits in User Profiles*. In: International Conferences on Knowledge Management and New Media Technology, Graz, Austria.

Nunes, M.A. Cerri,S.A. Blanc,N., 2008. *Towards User Psychological Profile*. In: IHC '08 Proceedings of the VIII Brazilian Symposium on Human Factors in Computing Systems. ACM Pages 196-203.

Nunes,M.A.,2008. *Recommender Systems based on Personality Traits*. PhD Thesis, University of MONTPELLIER 2.

Marshall, C., 2010. *Reading and Writing the Electronic Book.*. Editor Gary Marchionini, University of North Carolina, Chapel Hill.

Plimmer, B. Hsiao-Heng Chang, S. Doshi, M. Laycock,L. Seneviratne,N., 2010. *iAnnotate: Exploring Multi-User Ink Annotation in Web Browsers*. In: Proceedings of the 11th Australasian User Interface Conference (AUIC). ACM Pages 52-60.

Peabody, D. De Raad, B., 2002. *The Substantive Nature of Psycholexical Personality Factors: A Comparison Across Languages*. Journal of Personality and Social Psychology, vol 83, issu 10; Pages 983-997.

Roshchina, A. Cardiff, J. Rosso, P., 2011. *User Profile Construction in the TWIN Personality-based Recommender System.*. In: Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), IJCNLP. Pages 7379.

Ryckman,R. M., 2008. *Theories of Personality*. Editor Thomson Wadsworth USA.

Schmitt, D. Allik, J. McCrae, R. Benet-Martinez,V., 2007. *The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations*. Journal of Cross-Cultural Psychology, vol 38, Pages 173-212.

Tkalcic, M. Kunaver, M. Tasic, J. Košir,A., 2009. *Personality based user similarity measure for a collaborative recommender system*. In: Proceedings of the 5th Workshop on Emotion in Human-Computer Interaction-Real world challenges, Cambridge, UK. Pages 30-37.

The iMarkup Annotation Tool: a Commercial tool, viewed 13 September 2013, http://www.bplogix.com/support/imarkup-client.aspx