Chapter 7
Mining Facebook Data for Personality Prediction: An Overview

Davide Marengo and Michele Settanni

Abstract Users' interaction with Facebook generates trails of digital footprints, consisting of activity logs, "Likes", and textual and visual data posted by users, which are extensively collected and mined for commercial purposes, and represent a precious data source for researchers. Recent studies have demonstrated that features obtained using these data show significant links with users' demographic, behavioral, and psychosocial characteristics. The existence of these links can be exploited for the development of predictive models allowing for the unobtrusive identification of online users' characteristics based on their recorded online behaviors. Here, we review the literature exploring use of different forms of digital footprints collected on Facebook, the most used social media platform, for the prediction of personality traits. Then, based on selected studies, meta-analytic calculations are performed to establish the overall accuracy of predictions based on the analyses of digital footprints collected on Facebook. Overall, the accuracy of personality predictions based on the mining of digital footprints extracted from Facebook appear to be moderate, and similar to that achievable using data collected on other social media platforms.

7.1 Introduction

Since its creation in 2004, Facebook has experienced a steady increase in active users, reaching a total of 2.41 billion monthly active users as of the second quarter of 2019 (Statista 2019). In spite of growing competition by other social media platforms—such as Instagram (which is also currently owned by Facebook), Twitter, and Snapchat—and mounting controversies concerning the handling of user privacy (e.g., see the recent Cambridge Analytica scandal, Cadwalladr and Graham-Harrison 2018), as of 2019, Facebook remains the most used social media platform worldwide. Every day, millions of Internet users from different cultural contexts express their thoughts, emotions, and beliefs by writing, posting, and sharing content on Facebook,

D. Marengo (✉) · M. Settanni
Department of Psychology, University of Turin, Turin, Italy
e-mail: davide.marengo@unito.it

H. Baumeister and C. Montag (eds.), *Digital Phenotyping and Mobile Sensing*, Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-31620-4_7

which can be commented and/or endorsed ("liked" in the Facebook jargon) by their network of Facebook friends, or the overall Facebook community if the user profile is public. This unceasing interactive process produces a massive dataset of user-generated data, also referred to as "digital footprints", "digital records", or "digital traces" (e.g., Settanni and Marengo 2015; Youyou et al. 2015; Farnadi et al. 2016) consisting of personal information, activity logs, texts, pictures, and videos, with potential connections to users' offline behavioral and psychosocial characteristics.

In the research area of psycho-informatics (Markowetz et al. 2014), an increasing number of studies have explored the feasibility of mining digital footprints from Facebook, as well as those collectable from other social media platforms, in efforts to infer individual psychosocial and behavioral characteristics, such as personality (e.g., Golbeck et al. 2011; Kosinski et al. 2013; Liu et al. 2016; Farnadi et al. 2016) psychological distress (e.g., Choudhury et al. 2013; Settanni and Marengo 2015); for an overview see Settanni et al. 2018) and engagement in offline risk behaviors (e.g., alcohol use, Curtis et al. 2018). One of the earliest and most enduring project in this field of study is the my Personality project (Kosinski et al. 2013), which has attracted over 6 million Facebook users who have donated their digital footprints and responded to online questionnaires on a wide variety of psychometric measures, including Big Five personality traits, satisfaction with life, and intelligence.

Studies in this field share a common research design, broadly consisting of four steps: (1) User digital footprints are collected and mined for the automated extraction of multiple features; (2) Information about user individual characteristics is collected by means of different approaches (e.g., online survey, ecological momentary assessment via mobile apps); (3) Datasets combining the features extracted from digital traces and users' information are mined to explore associations between features and users' characteristics, and to train models aimed at predicting individuals' traits, typically using a machine-learning approach; and (4) Competing trained models are compared based on their accuracy in predicting users' characteristics on new independent datasets, leading to the identification of the best performing model. Most of existing studies employing this approach have explored the feasibility of mining digital footprints for the prediction of personality traits, as defined by the Big Five model (McCrae and Costa 1987; McCrae and John 1992) and the Dark Triad model (Paulhus and Williams 2002). The focus on personality is largely due to the importance of personality in predicting many life-course aspects for individuals, including academic success (e.g., Komarraju et al. 2009), job performance (e.g., Neal et al. 2012), financial decision making (Lauriola and Levin 2001; Bibby and Ferguson 2011), health and health-related behaviors (e.g., Soldz and Vaillant 1999; Bogg and Roberts 2004, 2013), subjective well-being (e.g., Hayes and Joseph 2003), and online behaviors (e.g., Matz et al. 2017).

Due to the current dominant position of Facebook over all existing social media platforms, the majority of existing studies exploring digital footprints for the prediction of personality traits have focused on the use of Facebook data. In this chapter, we present an overview of these studies discussing differences in the types of examined digital footprints, and the various analytical approaches used for mining them. Next, we refer to existing meta-analytical results to discuss the accuracy of prediction based on digital footprints. Finally, we discuss ethical issues related to this field of study, in particular with respect to privacy violations. Popularity of social media platforms can drastically change over time, resulting in a significant decline in user activity (e.g., Myspace, Ribeiro and Faloutsos 2015), and eventually, this may apply to Facebook. Still, regardless of the social media platform under focus, current findings indicate the potential of the analysis of social media data for research in *psychoinformatics* is expected to increase in the future as technology improves, and new methodologies are developed for the analysis of digital footprints (Hinds and Joinson 2019).

7.2 Facebook Digital Footprints and Their Use for the Study of Personality

Facebook provides developers with access to the digital footprints of consenting Facebook users, which can be accessed via a specifically-devised application programming interface (API), the Facebook Graph API (Facebook for developers 2019a). Following the Cambridge Analytica scandal, Facebook has introduced stricter requirements for accessing user data (Facebook for developers 2019b). Starting from August 2018, in order to be granted extended login permissions (e.g., access to user posts, or Likes), Facebook requires all apps to undergo a review process requiring developers to explain in details how they plan to use and manage user data; further, developers are required to pass a business verification procedure, ultimately limiting data access to business companies. Non-business developers (e.g., researchers) can still access user data (excluding user posts) if they pass an individual verification procedure. Downloadable user data includes personal information (e.g., name, age, gender, hometown), as well as posts, likes, pictures, and videos shared by the users on their Facebook wall. Access to posts, likes, pictures, and videos, include the possibility to download the actual user-generated content, as well as attached metadata (e.g., day and time of posting, received likes and comments). In the following sections, we present a brief literature overview of the studies that analyzed the connection between different types of Facebook digital footprints and personality, and describe some of the approaches used to mine collected data for prediction purposes.

Demographics and activity statistics. Facebook provides researchers with a vast array of user demographic information, such as age, gender, geographical location, and information about activity on Facebook, such as number of friends, and frequency and time of online posting. Furthermore, based on the examination of users' feed data, it is possible to compute summary statistics of users' specific online behaviors, such as the number of times the user has updated his/her Facebook status, uploaded photos or videos, sent or accepted a friend request, the number of events he/she attended, or the number of times he/she has been tagged in a photo.

Studies have shown significant links between activity statistics and personality, and specifically, the Big 5 traits have been shown to be significantly associated with users' behaviors on social media. For example, individuals with high Extraversion have been characterized by higher levels of activity on social media, and have a greater number of friends (Kosinski et al. 2014) than introverted individuals. Individuals with high Conscientiousness appear to be cautious in managing their social media profiles; they post fewer pictures, and engage in less group activity on social media (Kosinski et al. 2014). Furthermore, individuals with high openness tend to have larger networks than individuals low on the trait (Quercia et al. 2012). Studies varies in the type of analyses employed to investigate the links between extracted features, and personality. Gosling and colleagues (2011) studied bivariate associations between count statistics of a wide range of user activity information, including the number of photos, number of wall posts, the total number of friends, and personality ratings provided by both the user itself and an external observer. Findings showed that both self-report and observer-rated Extraversion scores had positive associations with the number of user wall posts, uploaded photos, and the size of the friendship network, while self-report openness was positively related with the number of user online. In turn, Wald and colleagues (2012) extracted 31 features from users' profile information and wall post activity -including age, gender, relationship status, and the number of friends, photos, interests, and comments—to test them as attributes for personality prediction using a set of machine learning algorithms. When examining the predictive power of single features, number of friends emerged as the stronger predictor of individual differences in agreeableness. Trained machine-learning models combining all features showed good accuracy (75% of correctly classified individuals) in detecting individuals with high openness scores (over the 90th percentile), while prediction on other traits was less satisfactory (<0.65% of accuracy).

Facebook likes. Facebook gives its users the possibility to "like" Facebook pages created by groups, companies, public figures, or external websites. Likes represent a mechanism used by Facebook users to express their positive association with specific web pages, comments, photos, and offline activities among others (Youyou et al. 2015). By accessing user Likes data through the Graph API (*user_likes* authorization), the following information can be obtained: the name of each page liked by the user, the category each page was registered in by their creators (e.g., *musician/band*; *Media/News Company*; *Italian Restaurant*), and a timestamp indicating when each page was liked by the user.

Several links exist the frequency of "Like" behaviors and users' personality. For example, individuals scoring high on conscientiousness tend to express less "Likes" on Facebook (Kosinski et al. 2014), while individuals high on openness tend to "Like" more content found on social media (Bachrach et al. 2012). Likes can also be mined to obtain information about users' interests and preferences as regards brands, politics, music, etc. However, because of the massive number of existing Facebook pages (>42 million pages), when examined at the page-level, user Likes data usually generates very large sparse logical matrices (i.e., matrix in which each row represents a user and each column represents a specific page), even at small sample sizes. For this reason, examining user Likes data for personality prediction generally requires the implementation of some form of dimensionality reduction method on the predictors set (i.e., digital footprints, in this case Likes). For example, Kosinski and colleagues (2013) processed a large matrix consisting of an average of 170 liked pages per 58,466 Facebook users using singular value decomposition (SVD), retaining a smaller subset 100 SVD components for performing personality predictions using logistic regression. When applied to Facebook Likes, the emerging SVD components may be interpreted as reflecting latent users' interests and preferences emerging from the co-occurrence of Likes to similar pages, e.g., persons who "likes" the official Facebook page of Harry Potter and Lord of the Rings movies tend to score similarly on a SVD component reflecting an interest in fantasy novels or movies. Using this approach, Kosinski and colleagues found remarkable accuracy prediction; in particular concerning the 'Openness' trait, for which score predictions based on user's Likes was found to be roughly as informative as using self-report personality scores (Kosinski et al. 2013). Using SVD can improve performance albeit at the expense of the interpretability of results, since information about user specific preferences is lost in the process of producing SVD components, and interpretation of emerging component is not always straightforward. Further, because of the sparsity of Likes data, this approach is only viable using large datasets (Kosinski et al. 2016). Another analytical approach which may help prevent overfitting problems when performing personality prediction using large datasets is the least absolute shrinkage and selection operator algorithm, or LASSO regression (Tibshirani 1996). Using the LASSO approach, prediction is initially performed using all available Likes for the user, but only Likes that contribute significantly to the overall prediction are included in final model. Because of the large number of features used to perform prediction, interpretation of results is problematic and it is typically limited to the predictive accuracy of trained models over personality scores. Using this analytic approach, Youyou and colleagues (2015) demonstrated that score predictions of Big 5 traits derived from the analysis of Facebook-Likes can be more accurate than personal judgments of a user's friends, relatives, and even spouse. Furthermore, as shown by Torfarson and colleagues (2017) using the same analytical approach, prediction accuracy over personality seems to increase proportionally with the number of user Likes analyzed, with only 20 Likes needed to obtain personality scores as accurate as those provided by users' spouse.

An alternative approach that can help face the problem of the sparsity of Facebook Likes consists in analyzing collected data at the category-level, as opposed to the page-level. In doing this, Facebook pages that are registered in Facebook under the same category (e.g., "Retail company" category: *Amazon.com, ebay,it*, *Macy's*; Musician/Band category: *Frank Sinatra, Adele, Kraftwerk*) are counted in a single category indicator. Using this approach, Baik and colleagues (2016) examined Likes data by recoding 8.355 distinct Likes pages into 183 categories. This procedure allowed them to perform linear regression analyses with no variable selection, while also preserving interpretability of results. Using this approached they were able to predict users' extraversion with average accuracy ($r = 42$), and provide some insight on the association between the personality trait and specific user interests (e.g., extroverts were more likely to show interest in hotels, sports, and shopping/retail, whereas introvert users showed interested in musicians, bands, and games/toy categories).

Texts. The *user_posts* Graph API authorization allows access to user posting activity, including personal status updates and comments, and the number of Likes received on users' posts.

Studies have shown significant links between the Big Five personality traits and features extracted from Facebook texts (e.g., Hall and Caton 2017; Schwartz et al. 2013). For example, Extraversion has been shown to be positively associated with the frequency of use of words about family and friends, and positive emotions (Schwartz et al. 2013) and to be negatively associated with use of words indicating cognitive processes (insight words, Hall and Caton 2017; words indicating tentativeness, causation, inhibition, Schwartz et al. 2013). Coherent with findings about depression and language use (Eichstaedt et al. 2018). Neuroticism has been linked with increased use of words indicating use of 1st person singular pronouns, negative emotions, and coarse language (Schwartz et al. 2013); in turn the Agreeableness trait has been linked to increased positive emotion words (Hall and Caton 2017; Schwartz et al. 2013). Furthermore, Garcia and Sikstrom (2014) explored associations between the Dark Triad personality traits, i.e., Machiavellianism, Narcissism, and Psychopathy, and textual features extracted from Facebook texts. Findings showed that Psychopathy was the personality trait most easily predictable from the semantic content of status updates. Results also showed that individuals with high levels on the Psychopathy and Narcissism traits posted more negative words in their Facebook posts, and published more "atypical" content when compared to individuals with low scores on these traits.

For the purpose of personality prediction, most studies have extracted features from texts using two text analyses approaches: the more traditional *closed-vocabulary* analysis and the recently emerging *open-vocabulary* analysis. Closed-vocabulary analysis has a long history in psychological science and can be viewed as theory-driven approach that consists of scoring language data according to predetermined semantic categories. One of the most popular instruments to apply this kind of approach is the Linguistic Inquiry and Word Count (LIWC) software, which has been developed over the past 20 years to measure multiple dimensions by computing the relative frequency of word categories (Pennebaker et al. 2015; Tausczik and Pennebaker 2010). LIWC allows the scoring of text documents based on a set of predetermined categories ranging from parts of speech (e.g., use of pronouns, numbers, punctuation), emotional expression (e.g., positive or negative emotions, anger, sadness), cognitive processes (e.g., insight, discrepancy), to social processes (e.g., friends, family), and personal concerns (e.g., body, death, money, occupation).

In contrast, the more recent open-vocabulary analysis employs data driven analytic approaches to explore the distribution of topics, words, and phrases naturally occurring in analyzed texts; thus producing results that are not limited by predetermined categories (Schwartz and Ungar 2015). For this reason, the open vocabulary approach is particularly suited for the analyses of alternative forms of communication, such as use of abbreviations (e.g., OMG, IMHO, NSFW), as well as pictorial symbols (e.g., emoticons, emoji), which represent a significant component of modern computer-mediated communication. Some of the most used approaches for performing open vocabulary analysis on Facebook texts are Latent Semantic Analysis (Dumais 2004) and Latent Dirichlet Allocation (Blei et al. 2003) algorithms. Both approaches have been used to infer the semantic and topical content of Facebook texts in studies exploring the link between personality and language use (e.g., Schwartz et al. 2013; Garcia and Sikström 2014). Despite the advantages linked to their use, these approaches generally require large datasets and emerging results are typically harder to interpret than those provided by closed-vocabulary analyses. Table 7.1 and Fig. 7.1 provide an illustrative example on how results from closed- and open-vocabulary analyses are usually presented when examining their association with personality scores. Presented results are based on the analyses collected on a sample of 296 adult Facebook users from Italy (Female = 67%, Age: M = 28.44, SD = 7.38, previously unpublished results), and show the correlation between Neuroticism (Ten Item Personality Inventory, Gosling et al. 2003) and LIWC closed-vocabulary features (Table 7.1) and LDA open-vocabulary features (Fig. 7.1) computed on users' status updates. Topics emerging from LDA analyses are depicted using word clouds in which words that are more strongly related to the topic are depicted with larger font size and darker tones. Findings using LIWC and LDA topics are coherent in showing that Neuroticism correlates with increased negative emotionality in Facebook posts, and appear to be negatively related to the expression of positive emotions. However, each approach can provide different insights on the specific forma or semantic language features associated with the examined trait. Based on presented results, it is worthy to note that, similarly to what it is usually observed in the literature, effect-size of correlations between personality and both closed- and open-vocabulary features, is generally quite low.

Table 7.1 Personality and Facebook language: closed-vocabulary (LIWC) correlates of neuroticism ($n = 296$, Correlations significant at $p < 0.05$)

LIWC category	r
I. Standard linguistic dimensions	
Negations	0.15
II. Psychological processes	
Affective processes	
Optimism	−0.15
Negative emotions	0.20
Anger	0.12
Sadness	0.28
Cognitive processes	
Possibility	0.18
Certainty	−0.13
Social processes	
Family	−0.15
Personal concerns	
Money and financial issues	−0.14
Body states, symptoms	0.12

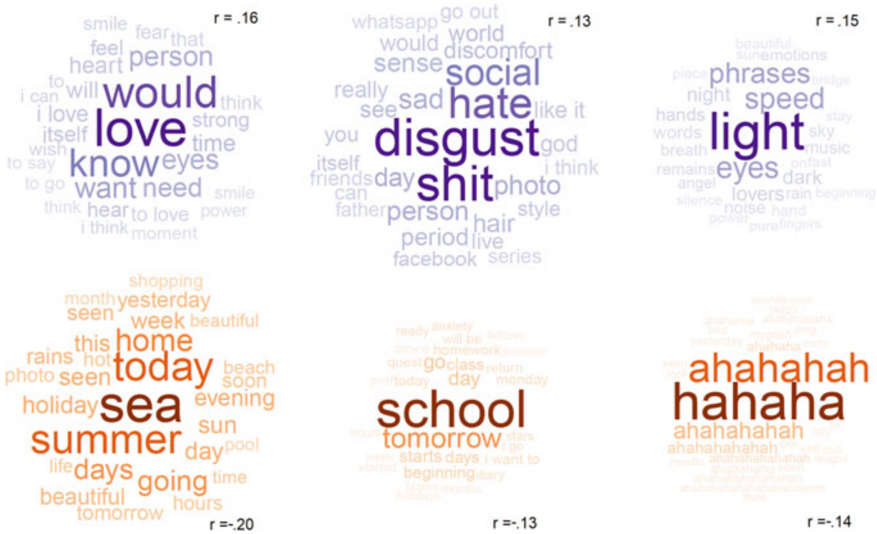

Fig. 7.1 Personality and Facebook language: open-vocabulary (LDA) correlates of neuroticism (n = 296, Correlations significant at $p < 0.05$)

However, as shown by Schwartz and colleagues (2013), when all extracted features are combined using predictive models, predictive accuracy is expected to improve significantly. Further, when compared to closed-vocabulary, open-vocabulary analysis has proven to yield additional insights and more in depth information about the behavioral characteristics of personality types beyond those that can be obtained by a traditional closed-vocabulary approach, resulting in increased prediction accuracy over personality (Schwartz et al. 2013).

Visual data. Sharing of pictures and videos is an increasingly frequent online activity and in the last few years new social media platforms have flourished focusing on the sharing of pictures and videos (e.g., Instagram, Snapchat) among users. This new phenomenon has also witnessed an increase on Facebook, as users are constantly provided with new features allowing for the proliferation of pictures and videos in their overall generated content, indicating a progressive shift from textual based interactions, to an increase in sharing of videos and photographs. Given the relative novelty of this phenomenon, only a few studies have explored the possible relations between visual data collected from Facebook, and personality, usually focusing on the analysis of user profile pictures (e.g., Celli et al. 2014; Torfason et al. 2016; Segalin et al. 2017). Access to the *default*, *user_photos* and *user_videos* Graph API authorizations allow researchers to inspect and download users' profile pictures, uploaded pictures, and videos. Celli and colleagues processed user profile pictures using a *bag-of-visual-words* approach, extracting 4096 non-interpretable visual features to perform predictions over personality scores. Prediction accuracy over personality traits was generally poor. However, they were able to show that extrovert and emotionally stable users post more frequently pictures in which they smile, and in which other people are included. On the other hand, introverts tend to appear alone, while Neurotics tend to post images without humans, and their pictures often feature close-up faces. Torfarson and colleagues, on the other hand, extracted information about specific facial features (e.g., eyeglasses, smiling, wearing lipstick) based on models trained on the CelebA dataset (Liu et al. 2015); based solely on features extracted from users' profile pictures, they observed a correlation of 0.18 between observed and predicted scores for the Big 5 personality traits. Finally, Segalin and colleagues (2017) extracted a large set of variables representing aesthetics-based features of images (i.e., color, composition, textual properties, and content), byte-level features, and both visual word and concept features extracted using other approaches (e.g., pyramid histogram of visual words; convolutional neural networks). In the study extroverts and agreeable individuals showed an increased inclination to post warm colored pictures and to exhibit many faces in their profile pictures; in turn, individuals high on neuroticism were more likely to post pictures of indoor places. As in the aforementioned studies, prediction accuracy was limited (~60% of correctly classified individuals), but still higher than that obtained in the study when using human raters. Overall, existing findings indicate that the accuracy of predicting personality achieved using visual data is still relatively limited compared to that achieved using other types of digital footprints (Azucar et al. 2018).

7.3 Establishing the Accuracy of Personality Predictions

Studies exploring the predictive power of Facebook digital footprints over personality vary significantly in the methods employed to assess such associations. Some studies implement simple bivariate analyses (e.g., zero-order correlation, independent samples t-test) examining the strength of associations between personality scores and numerical variables representing features extracted from digital footprints (e.g., Gosling et al. 2011; Panicheva et al. 2016), while other studies investigate the accuracy of personality predictions based of the mining of large set of features using predictive models (e.g., Kosinski et al. 2013; Celli et al. 2014). Other studies present both kinds of analyses (e.g., Schwartz et al. 2013; Farnadi et al. 2016).

Another important difference emerging from these studies relates to the use of validation techniques to avoid overfitting problems and support generalizability of results. Amongst the most used validation methods are the *holdout* method, the *k-fold* validation method, and the *leave-one-out* method. When using the holdout method, collected data is randomly split in two datasets of unequal size, a larger training set and a smaller test set, consisting of mutually independent observations. Analyses are first performed on the training set, resulting in a set of parameters or coefficients describing the association between features extracted from digital footprints, and personality scores. Then, the developed models are applied on the smaller test set using the estimated parameters or coefficients: Accuracy of personality predictions informs about the expected performance of the model on new, unseen observations. The *k*-fold validation method is similar, in that it also involves randomly splitting the dataset in a training set and a test set, but the process is repeated *k* times, resulting in *k* sets of results which are averaged to produce a single estimation of accuracy. Finally, in the leave-one-out method, the split is repeated for as many observations present in the dataset, i.e., the dataset is iteratively split so that only one observation is used to the test accuracy of predictions, while the remaining observations are used to estimate model parameters or coefficients. In general, *using k-fold validation and leave-one-out methods* is preferable over the *simple holdout* method (Kohavi 1995). As regards the number of folds, authors have suggested either using k = 5 and k = 10 folds when performing cross-validation over k = 2 folds, as using a larger number of folds is expected to decrease bias in estimating prediction errors (Rodriguez et al. 2010); however, it should be noted that as increasing the number of folds is only feasible using large datasets, as the large sample condition needs to be achieved in each of the fold (Wong 2015).

Most of the studies examining the predictive power of digital footprints collected from Facebook used a k-fold method to validate personality predictions (e.g., Golbeck et al. 2011; Bachrach et al. 2012; Quercia et al. 2012; Kosinski et al. 2013; Farnadi et al. 2016, 2018; Baik et al. 2016; Thilakaratne et al. 2016) followed by the holdout method (e.g., Celli et al. 2014; Schwartz et al. 2013). It is worthy to note that some authors did not employ a cross-validation technique in their studies, but they analyzed data from their whole samples (e.g., Gosling et al. 2011; Garcia and Sikström 2014). Caution should be used in interpreting their findings due to the limitations cited above, i.e. overfitting and lack of generalizability.

7.4 Accuracy of Personality Predictions Based on Facebook Data

Two recent meta-analytic studies have examined the literature aiming at estimating the overall prediction accuracy of digital footprints collected on a wide range of social media platforms (e.g., Facebook, Twitter, Sina Weibo, and Instagram) over users' individual characteristics. Settanni and colleagues (2018) identified 38 papers investigating associations between digital footprints and a set of individual characteristics, including personality (e.g., traits from the Big Five and the Dark Triad personality models), psychological well-being (e.g., satisfaction with life, depression), and intelligence. Overall, based on findings from a subset of 18 studies, the estimated overall accuracy in predicting personality, computed as a meta-analytical correlation was moderate (r = 0.34). Further, the results of the meta-analysis showed that the overall accuracy in predicting personality was lower than the accuracy in predicting psychological well-being (r = 0.37), but higher than what was computed for the prediction of intelligence (r = 0.29). In turn, the meta-analysis by Azucar and colleagues (2018) examined literature focusing on studies presenting associations between digital footprints and traits from the Big Five model—i.e., Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Based on prediction results presented in 16 independent studies, Extraversion appears to be associated with the highest overall prediction accuracy (r = 0.40), followed by Openness (r = 0.39), Conscientiousness (r = 0.35), Neuroticism (r = 0.33), and Agreeableness (r = 0.29). In the aforementioned meta-analyses, a set of meta-regression was conducted to recognize factors influencing prediction accuracy, finding that the accuracy improves when models include information about users' demographics and more than one type of digital footprints. In both meta-analyses, examined studies referred to a conceptualization of personality as a multi-dimensional construct; hence, reported results should not be interpreted as indicating accuracy in predicting discrete personality types.

Given the relevance of Facebook in the social media world, we re-estimated overall prediction accuracy reported in the two cited meta-analyses, including only those studies presenting predictions based on Facebook data. Very similar results emerged, with r values ranging from 0.33 (Neuroticism) to 0.43 (Extraversion), and an overall correlation equal to 0.34, 95% CI [0.24–0.44] indicating that the use of data extracted from differing social media platforms is not expected to have a significant effect on the accuracy of personality predictions. As noted by Kosinski and colleagues (2013) this correlation size corresponds to the "personality coefficient" (a Pearson correlation ranging from 0.30 to 0.40; Meyer et al. 2001; Roberts et al. 2007), which is the upper limit of correlations between behaviors and personality traits reported in past psychological studies. Among the examined traits, Extraversion appears to be most easily predictable based on the examination of digital footprint from Facebook, a finding which is compatible with findings emerging from studies exploring other sources of digital footprints (e.g., smartphone usage, Stachl et al. 2017). Still, upon inspecting these results, it is important to note that the average prediction accuracy reported by existing studies is still quite limited. For this reason, reliability of individual personality predictions obtained by mining digital footprint is still quite low, in particular when compared with that obtainable with traditional self-report assessments, limiting their use of predicted scores for assessment purposes.

7.5 Conclusions

In this chapter, we presented an overview of studies examining the feasibility of inferring individual differences in personality of Facebook users based on the analyses of their digital footprints (e.g., user demographics, texts, Facebook Likes, and pictures). Published studies vary significantly in employed analytical approaches, both in terms of methods used for extracting features from raw social media data, performing predictions, and validating results. Overall, the accuracy of personality predictions based on the mining of digital footprints extracted from Facebook appears to be moderate, and similar to that achievable using data collected on other social media platforms. Given the relatively recency of this area of research, and the rapid evolution of data mining techniques, we expect accuracy of personality prediction to improve in the near future. The ability to use digital footprints for the unobtrusive assessment of personality traits can represent a rapid, cost-effective alternative to surveys to reach large online populations, an approach which can be beneficial for academic, health-related, and commercial purposes pursued in the online environment (e.g., improve the efficacy of online health-related messages and interventions, enhance online recommender systems, improve user experiences, enhance efficacy of advertising by tailoring online message to personality attributes, including political messages, Matz et al. 2017).

7.5.1 Future Directions and Ethical Concerns

Meta-analyses conducted to determine the predictive power of social media data on psychological characteristics in general, and in particular on personality traits, revealed that collecting information from multiple sources (e.g. from pictures and text) and different social media platforms, permit to achieve greater predictive power. This means that in the near future, the aggregation of features extracted from different types of data and the inclusion of data from different social media platforms, or from other sources, such as wearables (e.g., iwatch, runkeeper, etc.) or mobiles, will probably lead to relevant improvements in the predictive power of these kinds of models. Furthermore, the progressive expansion and complexification of individuals' online activities will support the creation of an increasing number of datasets, easily available for the development of predictive models. It is foreseeable that the development of new and more efficient approaches to data collection, integration, and analysis (e.g., using deep learning algorithms) will contribute to making predictions more accurate and reliable, extending their reach well beyond the field of personality traits, towards the prediction of more specific characteristics, behaviors, and even biological features (Montag et al. 2017; Sariyska et al. 2018). The acquisition of these new capabilities will raise important ethical issues that cannot be underestimated.

Social media data may be used in ways that surpass what users intend, or understand, when they give consent to their collection. Apparently innocuous data points may be and have been used to reveal information that users might expect to stay private. These predictions can have negative consequences for social media users: First, predicted traits can be used to make decisions relative to single Facebook users, without their explicit consent to disclose such characteristics (e.g., in hiring procedures); Second, as recently highlighted by Matz and colleagues (2017), psychological targeting procedures might be developed and aimed at manipulating the behavior of large groups of people, without the individuals being aware of it.

Given the possibility of using raw data to infer relevant individual characteristics, the need is emerging for a more careful consideration of ethical challenges related to the use of data extracted from Facebook or other social media. It is worthy to note that, while research in medicine and psychology is routinely subjected to IRB ethical approvals, the same does not apply for computer science. Same as in clinical disciplines, computers scientists should also develop and apply ethical restrictions when they do research in this field, and research projects should adhere to the principle of beneficence: the good of research participants should be taken in high consideration and influence the assessment of risks versus benefits when planning a research. As recently noted in a Nature editorial (2018) on the Cambridge Analytica scandal, the fact that data are there should not be a sufficient reason to exploit these in order to conduct research.

References

Azucar D, Marengo D, Settanni M (2018) Predicting the Big 5 personality traits from digital footprints on social media: a meta-analysis. Pers Individ Dif 124:150–159. https://doi.org/10.1016/j.paid.2017.12.018

Bachrach Y, Kosinski M, Graepel T et al (2012) Personality and patterns of facebook usage. In: Proceedings of the 4th annual ACM web science conference. ACM, New York, NY, USA, pp 24–32

Baik J, Lee K, Lee S et al (2016) Predicting personality traits related to consumer behavior using SNS analysis. New Rev Hypermedia Multimed 22(3):189–206. https://doi.org/10.1080/13614568.2016.1152313

Bibby PA, Ferguson E (2011) The ability to process emotional information predicts loss aversion. Pers Individ Dif 51(3):263–266. https://doi.org/10.1016/j.paid.2010.05.001

Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3(Jan):993–1022

Bogg T, Roberts BW (2004) Conscientiousness and health-related behaviors: a meta-analysis of the leading behavioral contributors to mortality. Psychol Bull 130(6):887–919. https://doi.org/10.1037/0033-2909.130.6.887

Bogg T, Roberts BW (2013) The case for conscientiousness: evidence and implications for a personality trait marker of health and longevity. Ann Behav Med 45(3):278–288. https://doi.org/10.1007/s12160-012-9454-6

Cadwalladr C, Graham-Harrison E (2018) Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian

Celli F, Bruni E, Lepri B (2014) Automatic personality and interaction style recognition from facebook profile pictures. In: Proceedings of the 22nd ACM international conference on multimedia. ACM, New York, NY, USA, pp 1101–1104

Choudhury MD, Gamon M, Counts S, Horvitz EJ (2013) Predicting Depression via Social Media. In: ICWSM

Curtis B, Giorgi S, Buffone AEK et al (2018) Can Twitter be used to predict county excessive alcohol consumption rates? PLoS ONE 13(4):e0194290. https://doi.org/10.1371/journal.pone.0194290

Dumais ST (2004) Latent semantic analysis. ARIST 38(1):188–230. https://doi.org/10.1002/aris.1440380105

Eichstaedt JC, Smith RJ, Merchant RM et al (2018) Facebook language predicts depression in medical records. Proc Natl Acad Sci U S A 115(44):11203–11208. https://doi.org/10.1073/pnas.1802331115

Facebook for developers (2019a) Graph API. https://developers.facebook.com/docs/graph-api. Accessed 30 Jul 2019

Facebook for developers (2019b) App Review. https://developers.facebook.com/docs/apps/review/. Accessed 30 Jul 2019

Farnadi G, Sitaraman G, Sushmita S et al (2016) Computational personality recognition in social media. User Model User-Adap Inter 26(2):109–142. https://doi.org/10.1007/s11257-016-9171-0

Farnadi G, Tang J, De Cock M, Moens M-F (2018) User profiling through deep multimodal fusion. In: Proceedings of the eleventh ACM international conference on web search and data mining. ACM, New York, NY, USA, pp 171–179
Garcia D, Sikström S (2014) The dark side of Facebook: semantic representations of status updates predict the dark triad of personality. Pers Individ Dif 67:92–96. https://doi.org/10.1016/j.paid.2013.10.001
Golbeck J, Robles C, Turner K (2011) Predicting personality with social media. In: CHI '11 extended abstracts on human factors in computing systems. ACM, New York, NY, USA, pp 253–262
Gosling SD, Augustine AA, Vazire S et al (2011) Manifestations of personality in online social networks: self-reported facebook-related behaviors and observable profile information. Cyberpsychol Behav Soc Netw 14(9):483–488. https://doi.org/10.1089/cyber.2010.0087
Gosling SD, Rentfrow PJ, Swann WB (2003) A very brief measure of the big-five personality domains. J Res Pers 37(6):504–528. https://doi.org/10.1016/S0092-6566(03)00046-1
Hall M, Caton S (2017) Am I who I say I am? Unobtrusive self-representation and personality recognition on Facebook. PLoS ONE 12(9):e0184417. https://doi.org/10.1371/journal.pone.0184417
Hayes N, Joseph S (2003) Big 5 correlates of three measures of subjective well-being. Pers Individ Dif 34(4):723–727. https://doi.org/10.1016/S0191-8869(02)00057-0
Hinds J, Joinson A (2019) Human and computer personality prediction from digital footprints. Curr Dir Psychol Sci 28(2):204–211. https://doi.org/10.1177/0963721419827849
Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of international joint conference on AI. pp 1137–1145
Komarraju M, Karau SJ, Schmeck RR (2009) Role of the Big Five personality traits in predicting college students' academic motivation and achievement. Learn Individ Differ 19(1):47–52. https://doi.org/10.1016/j.lindif.2008.07.001
Kosinski M, Bachrach Y, Kohli P et al (2014) Manifestations of user personality in website choice and behaviour on online social networks. Mach Learn 95(3):357–380. https://doi.org/10.1007/s10994-013-5415-y
Kosinski M, Stillwell D, Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci U S A 110(15):5802–5805. https://doi.org/10.1073/pnas.1218772110
Kosinski M, Wang Y, Lakkaraju H, Leskovec J (2016) Mining big data to extract patterns and predict real-life outcomes. Psychol Methods 21(4):493–506. https://doi.org/10.1037/met0000105
Lauriola M, Levin IP (2001) Personality traits and risky decision-making in a controlled experimental task: an exploratory study. Pers Individ Dif 31(2):215–226. https://doi.org/10.1016/S0191-8869(00)00130-6
Liu Z, Luo P, Wang X, Tang X (2015) Deep learning face attributes in the wild. pp 3730–3738
Liu L, Preotiuc-Pietro D, Samani ZR, Moghaddam ME, Ungar LH (2016) Analyzing personality through social media profile picture choice. In: ICWSM, pp 211–220
Markowetz A, Błaszkiewicz K, Montag C et al (2014) Psycho-informatics: big data shaping modern psychometrics. Med Hypotheses 82(4):405–411. https://doi.org/10.1016/j.mehy.2013.11.030
Matz SC, Kosinski M, Nave G, Stillwell DJ (2017) Psychological targeting as an effective approach to digital mass persuasion. Proc Natl Acad Sci U S A 114(48):12714–12719. https://doi.org/10.1073/pnas.1710966114
McCrae RR, Costa PT (1987) Validation of the five-factor model of personality across instruments and observers. J Pers Soc Psychol 52(1):81–90. https://doi.org/10.1037/0022-3514.52.1.81
McCrae RR, John OP (1992) An introduction to the five-factor model and its applications. J Pers 60(2):175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x
Meyer GJ, Finn SE, Eyde LD et al (2001) Psychological testing and psychological assessment: a review of evidence and issues. Am Psychol 56(2):128
Montag C, Markowetz A, Blaszkiewicz K et al (2017) Facebook usage on smartphones and gray matter volume of the nucleus accumbens. Behav Brain Res 329:221–228. https://doi.org/10.1016/j.bbr.2017.04.035

Nature (2018) Cambridge Analytica controversy must spur researchers to update data ethics. Nature 555:559–560. https://doi.org/10.1038/d41586018038564

Neal A, Yeo G, Koy A, Xiao T (2012) Predicting the form and direction of work role performance from the Big 5 model of personality traits. J Organ Behav 33(2):175–192. https://doi.org/10.1002/job.742

Panicheva P, Ledovaya Y, Bogolyubova O (2016) Lexical, morphological and semantic correlates of the dark triad personality traits in Russian Facebook texts. In: 2016 IEEE artificial intelligence and natural language conference (AINL). pp 1–8

Paulhus DL, Williams KM (2002) The dark triad of personality: narcissism, machiavellianism, and psychopathy. J Res Pers 36(6):556–563. https://doi.org/10.1016/S0092-6566(02)00505-6

Pennebaker JW, Boyd RL, Jordan K, Blackburn K (2015) The development and psychometric properties of LIWC2015

Quercia D, Lambiotte R, Stillwell D et al (2012) The personality of popular Facebook users. In: Proceedings of the ACM 2012 conference on computer supported cooperative work. ACM, New York, NY, USA, pp 955–964

Ribeiro B, Faloutsos C (2015) Modeling website popularity competition in the attention-activity marketplace. In: Proceedings of the eighth ACM international conference on web search and data mining. ACM, New York, NY, USA, pp 389–398

Roberts BW, Kuncel NR, Shiner R, Caspi A, Goldberg LR (2007) The power of personality: the comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. Perspect Psychol Sci 2(4):313–345

Rodriguez JD, Perez A, Lozano JA (2010) Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans Pattern Anal Mach Intell 32(3):569–575. https://doi.org/10.1109/TPAMI.2009.187

Sariyska R, Rathner E-M, Baumeister H, Montag C (2018) Feasibility of linking molecular genetic markers to real-world social network size tracked on smartphones. Front Neurosci 12. https://doi.org/10.3389/fnins.2018.00945

Schwartz HA, Eichstaedt JC, Kern ML et al (2013) Personality, gender, and age in the language of social media: the open-vocabulary approach. PLoS ONE 8(9):e73791. https://doi.org/10.1371/journal.pone.0073791

Schwartz HA, Ungar LH (2015) Data-driven content analysis of social media: a systematic overview of automated methods. Ann Am Acad Pol Soc Sci 659(1):78–94. https://doi.org/10.1177/0002716215569197

Segalin C, Celli F, Polonio L et al (2017) What your facebook profile picture reveals about your personality. In: Proceedings of the 25th ACM international conference on multimedia. ACM, New York, NY, USA, pp 460–468

Settanni M, Azucar D, Marengo D (2018) Predicting individual characteristics from digital traces on social media: a meta-analysis. Cyberpsychol Behav Soc Netw 21(4):217–228. https://doi.org/10.1089/cyber.2017.0384

Settanni M, Marengo D (2015) Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts. Front Psychol 6. https://doi.org/10.3389/fpsyg.2015.01045

Soldz S, Vaillant GE (1999) The Big Five personality traits and the life course: a 45-year longitudinal study. J Res Pers 33(2):208–232. https://doi.org/10.1006/jrpe.1999.2243

Stachl C, Hilbert S, Au J-Q et al (2017) Personality traits predict smartphone usage. Eur J Pers 31(6):701–722. https://doi.org/10.1002/per.2113

Statista (2019) Number of monthly active Facebook users worldwide as of 2nd quarter 2019 (in millions). https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/. Accessed 30 Jul 2019

Tausczik YR, Pennebaker JW (2010) The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol 29(1):24–54. https://doi.org/10.1177/0261927X09351676

Thilakaratne M, Weerasinghe R, Perera S (2016) Knowledge-driven approach to predict personality traits by leveraging social media data. In: 2016 IEEE/WIC/ACM international conference on web intelligence (WI). pp 288–295

Tibshirani R (1996) Regression shrinkage and selection via the lasso. J Royal Stat Soc Ser B (Methodol) 58(1):267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Torfason R, Agustsson E, Rothe R, Timofte R (2016) From face images and attributes to attributes. In: Lai S-H, Lepetit V, Nishino K, Sato Y (eds) Computer vision—ACCV 2016. Springer International Publishing, pp 313–329

Torfason R, Agustsson E, Rothe R, Timofte R (2017) From face images and attributes to attributes. In: Lai S-H, Lepetit V, Nishino K, Sato Y (eds) Computer vision—ACCV 2016. Springer International Publishing, pp 313–329

Wald R, Khoshgoftaar T, Sumner C (2012) Machine prediction of personality from Facebook profiles. In: 2012 IEEE 13th international conference on information reuse integration (IRI). pp 109–115

Wong T-T (2015) Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognit 48(9):2839–2846. https://doi.org/10.1016/j.patcog.2015.03.009

Youyou W, Kosinski M, Stillwell D (2015) Computer-based personality judgments are more accurate than those made by humans. Proc Natl Acad Sci U S A 112(4):1036–1040. https://doi.org/10.1073/pnas.1418680112