Automatic Diagnosis and Screening of Personality Dimensions and Mental Health Problems

3

Yair Neuman

3.1 Introduction: Automatic Analysis of Personality

3.1.1 Personality and Diagnosis

The idea of personality [1–6] suggests that human beings, and other nonhuman organisms (e.g., [7]), exhibit unique patterns of thought, emotion, and behavior that may be used as a kind of a psychological "signature." These patterns are not unique, in the sense that a genetic signature is unique or a fingerprint is unique. For example, a person may be characterized as having an introvert type of personality—but being introvert does not uniquely characterize that particular individual, as there are many other introverts. Although each of us is unique in a very profound sense, the idea of personality suggests that we may be characterized by measuring a limited number of personality dimensions that are shared by other human beings. Measuring personality dimensions can be used for diagnosis or in screening, or in a clinical or a nonclinical context.

3.1.2 Diagnosis Versus Screening

The difference between *diagnosis* and *screening* should be clarified [8]. The aim of diagnosis is to confirm, or rule out, the hypothesis that a *specific individual* has a certain personality dimension. For example, a teenager is sent to a clinical psychologist, after being involved in too many accidents. Reasonably dismissing other explanations, the psychologist may hypothesize that the improbable cluster of

Y. Neuman (✉)
The Department of Cognitive and Brain Sciences and Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva, Israel
e-mail: yneuman@bgu.ac.il

F. Jotterand, M. Ienca (eds.), *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues*, Advances in Neuroethics,
https://doi.org/10.1007/978-3-030-74188-4_3

accidents is indicative of a self-harming behavior. The psychologist may therefore use a variety of tools to diagnose the teenager as having some kind of *accidental personality*. By diagnosing this specific individual, the psychologist is trying to conclude whether a particular personality, or personality dimension, validly characterizes the teenager. Diagnosis is highly important, as it organizes a cluster of behaviors, emotions, and cognitions into an interpretable pattern that may be used for prognosis, treatment, and prediction.

Unlike diagnosis, which is focused on the individual, screening is broadly used to determine which member of *a large group of individuals* has the attribute in question [8]. In other words, the group—rather than the individual—is the focus of analysis. For instance, a computer program might automatically screen a large group of Facebook users for signs of depression and suicidal thoughts, rank them according to their depressivity and risk level, and send the top-ranked individuals a recommendation for an in-depth personal diagnosis. Such automatic screening for depression has indeed been found to be effective (e.g., [9]), and one may use it in cases where personal clinical diagnosis is not easily available, at least in the initial phase of a process.

3.1.3 The Clinical Versus the Nonclinical Context

Personality dimensions may have both clinical and nonclinical aspects. The clinical context of diagnosis focuses on emotional, mental, and behavioral *disorders*. In other words, the aim of a clinical diagnosis is to determine the presence of a certain personality disorder. In contrast, diagnosis in the nonclinical context is broadly used to measure the personality dimensions of an individual, for use outside the clinical context. For example, one of the dimensions measured by the SWAP-200 personality test [6] is the *narcissistic personality disorder* [10]. When imagining a narcissist personality, we usually think about an individual with an exaggerated sense of self-importance, accompanied by a nonadaptive behavior. However, narcissism is a spectrum, ranging from a healthy form of self-love, to a pathological conflict over self-value. The typical narcissist has an exaggerated sense of self-importance—but narcissism is not necessarily pathological, and identifying and measuring the "soft" levels of narcissism may be used for practical purposes. Here are two examples where we may be interested in measuring nonpathological versus pathological narcissism.

Example 1 An intelligent targeted advertising engine might analyze the texts written by its "target" in social media to conclude that she scored highly on both extroversion and narcissism. In this case, the engine might send her a personalized advertisement for a rock-n-roll concert, but when designing the ad text, it might place strong emphasis on themes that resonate with an extrovert narcissist personality—for example, by appealing to her sense of self-importance (e.g., "The Greatest Rock Concert, for the Greatest People"). By first identifying the target's personality dimensions, then appealing to those dimensions at the unconscious level, the engine appears to have better chances of achieving its major aim of "seducing" the individual to click on the ad.

Example 2 Let us imagine that we are asked to design an automated system for determining the risk factor of violent men who might pose a threat to their spouses, in order to take preventive steps and reduce the danger of homicide. A forensic psychologist might teach us that one of the dimensions worth examining is *pathological narcissism*—as a man who cannot see beyond his self-centered perspective is more dangerous than someone who cognitively, emotionally, and behaviorally understands that he is not the center of the universe. The engine that we may build should therefore run on data produced by violent men under inspection—such as the text messages they send to their (ex-)wives. By analyzing the texts, the engine should score the text for signs of pathological narcissism, and using machine learning (ML) algorithms, we may examine whether narcissism is a risk factor that distinguishes between dangerous husbands and those are merely "barking" with no real danger of "biting." In this context, of course, great emphasis should be placed on the selection and engineering of the appropriate features, as we are not simply seeking general signs of narcissism, but signs of pathological narcissism that may point to the risk of a potentially harmful husband. In this regard, the success of the automatic personality engine is measured by its ability to classify dangerous vs. non-dangerous husbands, based on their respective exhibited levels of pathological narcissism.

A similar idea may be applied to the context of depression and depressivity as a personality dimension. Depression, in its pathological form, is a risk factor for suicide. There are contexts in which we would like to screen for individuals who suffer from depression, as they may pose a threat to themselves and to their surroundings. To avoid a straw-man type of fallacious reasoning, I must emphasize that I am making no argument here about depression and dangerousness. The overwhelming majority of depressed individuals would not harm others, or themselves. However, in certain contexts, screening for individuals with pathological depressivity and suicidal intentions may save lives, and this is only one example in which computational personality analysis may contribute to the field of diagnosis [11, 12]. One manifest example is the one of the Germanwings Flight 9525. On March 2015, this Airbus plane crashed in the Alps, resulting in 150 casualties. This was not an accident: it was deliberately planned and executed by the co-pilot, Andreas Lubitz, who had been treated for suicidal tendencies. It was not only an act of suicide, but of murder (i.e., homicide-suicide), since by crashing the plane, Lubitz took the lives of many innocent people, who paid the price of the failure to screen out individuals such as himself, who suffered from suicidal tendencies, from serving as a pilot. As a rule, the aviation industry is highly sensitive to the safety of its passengers, as any mistake, improbable as it may be, may result in a humanitarian and economic catastrophe. The pilots are clearly one of the vulnerabilities of the system, as attested by the case of Germanwings Flight 9525. Had they been less zealous in protecting personal privacy, the German authorities could have used Lubitz's medical records—coupled with other sources of information—to prevent him from taking the lives of so many innocent people. In designing such an alarm system, ethical issues could have easily been resolved, and these should not be used to counter the necessity of verifying the mental health of pilots.

3.2 Computational Personality Analysis

In the previous sections, I explained the idea of personality, the difference between screening and diagnosis, and the expression of personality in clinical and non-clinical contexts. Traditionally, personality analysis is conducted by a human expert or by means of questionnaires that require the voluntary collaboration of the person being diagnosed, and the validity of her self-reported personality dimensions. However, when massive datasets are involved, the use of a human expert and manual data analysis is impractical. Moreover, in such contexts, it is usually extremely difficult, if not impossible, to gain the voluntary participation of the diagnosed subjects, or a valid measure of their personality dimensions. Here we have a clear answer to the questions "Why are automated approaches to personality analysis useful for mental health?" and "Why are they preferable to conventional (non-automated) approaches?" To clarify these points, let us examine an example.

It has been recently reported[1] that suicides among active-duty members of the U.S. Air Force surged to its highest level in over three decades. Given the ratio of mental health experts in the U.S. military to active-duty personnel, it is almost impossible to diagnose depression among the soldiers in a reasonable space of time. Moreover, using questionnaires may be the wrong strategy, as the soldiers may fail to answer them honestly, due to lack of self-awareness or fear of being "exposed" and dismissed from duty. In such contexts, automated approaches are the only solution, as they provide a valid means of diagnosing large numbers of people in a very short time, by using texts (written or spoken) that they *naturally* produce. Automated systems are therefore preferable whenever human experts cannot provide diagnosis, due to the constraints of number or time, or whenever the use of questionnaires or other conventional methods is less appropriate.

The same rationale and justification that apply to medical diagnosis also apply to the diagnosis of personality through automated tools. Google has recently demonstrated[2] how an automated system can identify skin diseases—a massive, quick, and valid diagnosis that can match the performance of human experts. Google justifies the use of such tools by the dearth of dermatologists, coupled with the relatively high number of individuals seeking diagnosis and treatment—the same justification as the one cited for using tools of computational personality analysis.

---

[1] https://time.com/5780447/air-force-suicide-surge/.

[2] https://ai.googleblog.com/2019/09/using-deep-learning-to-inform.html.

3.2.1 The Relevance of Computational Personality Analysis in the Current Culture

The idea of personality analysis, and automatic personality analysis in particular, is particularly relevant for the modern and post-modern societies, where the individual has become the focus of interest, and where digital traces of individuals are evident everywhere. In the past, the idea of "personalized medicine" was irrelevant, as medicine was at an embryonic stage and the individual was not at the center. However, today we expect medicine—as an advanced practice supported by a wealth of data and technological tools—to address our particular signatures, uniqueness, and needs, in a manner that maximizes the effectiveness of diagnosis and treatment. The same is true for other fields as well—particularly that of personality analysis.

When the famous Edward Bernays launched a sophisticated pro-smoking campaign for women at 1929, he marketed it under the slogan "Torches of Freedom." Instead of seeing cigarettes as deadly poison, women were encouraged to perceive smoking as a feminist and emancipatory act. The campaign was extremely successful, and in retrospect, one may wonder whether the benefits of emancipation were worth the price paid by women who have joined the "cancer club." In any case, Bernays's campaign was extremely clever in its targeting of women instead of the public at large—but today, a more individualized campaign would probably have been called for. Although women share certain distinct needs and characteristics as a group, the unique attributes of each individual woman are such that—as with personalized medicine—they should be taken into account for maximum effectiveness.

In sum, although the identification of personality dimensions for any practical purpose is a long-established practice, it has become particularly important in modern-day, technologically oriented societies, where it is easier to identify such patterns by analyzing the digital traces left by everyone almost everywhere. For this reason, we use tools of AI—or more specifically, machine-learning tools—to automatically analyze the data and perform screening or diagnosis. Computational personality analysis, as its name suggests, is the field where methodologies and tools are developed for automated analysis of personality dimensions. In the following sections, this general approach is presented in a nutshell through a specific example, then elaborated further.

3.2.2 Computational Personality Analysis in a Nutshell

A project in computational personality analysis usually starts with a clear idea of (1) *Why* (i.e., Why do we need automated personality analysis?), (2) *Which* (i.e., Which personality dimensions are relevant for the task?), and (3) *How* (i.e., How are we going to measure the personality dimensions?).

When constructing a system for automatically measuring personality dimensions, we usually use a *supervised* form of learning, where the ML algorithms are trained on a tagged dataset—namely a training set of examples and their diagnosis/tag. For example, if we would like to teach the computer to measure depression among teenagers, we might provide it with personal text passages (e.g., diary entries) that they have written. After each text is read and scored by several experts according to clear criteria, it is given a “depression score,” on the assumption that the level of depression evident in the text reveals its author’s depressivity. If possible, we can even diagnose the people who wrote the texts, and score them on a depressivity scale, to validate the performance of our algorithm.

The training set is therefore composed of a set of documents, each scored according to the depression level, as measured by the human experts. In a simpler case, we do not score the text on a spectrum, and the expert may be required only to tag the text as either clinically depressed, or not. In instances where we would like to predict a continuous score, we use an ML model fit for regression; in categorical instances, we use a ML algorithm designed for classification.

There are various ML algorithms for regression and classification—from Naïve Bayes, to SVM, XGBoost, and Deep Neural Networks (DNN). The decision as to which ML algorithm to use is governed by the particular details of our task and data, and usually several ML algorithms are trained and tested on the dataset.

To teach the computer to diagnose depression using texts, we should provide it with a set of features that characterize each document. These features, sometimes called *variables* or *attributes*, are supposed to reveal whether or not the text is indicative of depressivity. There are various features that we can measure in each text, and in this case, too, the decision which features to analyze is determined by the particular characteristics of the project. For the diagnosis of depression, for example, the most intuitive attributes that we may want to measure are content categories, such as those we can measure using *Empath*.[3] Let us assume the computer is provided with the following text, which an expert has tagged as “depressed”:

> I’ sad and lonely. No one loves me and I feel abandoned and neglected. Life is hopeless and there is no hope, just despair.

This almost caricature-like example clearly represents depressivity—one need not be a certified psychologist to see that. After running the text through automatic analysis, Empath provides a list of content categories, and the extent in which they are expressed in the text. This reveals the following content categories, and the extent (i.e., frequency) to which they are expressed in the text:

Content category	Score
Shame	2
Negative emotion	1
Body	1
Love	1
Violence	1
Sadness	1
Contentment	1
Pain	1
Emotional	1
Nervousness	1
Cold	1

[3] http://empath.stanford.edu/.

We can see that the content categories identified by the computer can be theoretically associated with depressivity. The computer than learns that a "depressed text" (i.e., a depressed person), at least according to the above example, has a "signature"—a particular combination of content categories and their "weight" in the text—which may be optimally used to classify a text as "depressed" or "nondepressed."

When fed with enough examples and with the appropriate features, the machine learning algorithm learns a model that optimally classifies a text as "depressed" or "nondepressed." To test how well the machine has learned to identify depressed texts, we present it with another set of texts, which serves as the test set. The machine learning algorithm then uses the model that it built in the previous learning phase to identify *nontagged* texts as "depressed" or "nondepressed." These are new texts that the algorithm has not seen before, and therefore its ability to successfully classify the new texts is an indication to the extent in which it can validly classify/diagnose a text as "depressed." At this point, we measure the performance of the model through various diagnostic measures—such as precision, and recall—and by validating the results.

One important way of validation is through the *k-fold cross-validation procedure*. This procedure aims to address the problem of over-fitting our model to the data. In each run of the cross-validation, we divide the dataset into a training set and a test set: we train the model on one set and test it on the other, and the performance of the model is tested by averaging the results over several runs. If the model performs well, we may apply it in practice, and use it as a kind of a "digital psychologist."

How good is the performance of such computational personality analysis tools? In many cases, they provide a highly successful and efficient diagnosis. For example, [9] have provided 84% accuracy in diagnosing depression, and current studies provide much better results. In a recent study [12], we designed a computer algorithm for identifying a psychopathic signature in texts. The test set included 2333 texts—only 4% of which were texts with a distinctive psychopathic signature. Identifying such a text by chance has a very low probability ($p = 0.04$), but when applying our automated methodology, we were able to identify them with 67% precision, which is an enormous improvement over the base-rate of "psychopathic" texts in the dataset.

Having presented the idea of computational personality analysis in a nutshell, I shall now detail and elaborate it in the next section.

3.3 Computational Personality Analysis Further Detailed

For automatically measuring personality dimensions and disorders, we need data, which may come in various forms and modalities.

It is generally assumed that the language we use is a window onto our personality. If someone says: "I'm depressed and lonely," then given the appropriate context of interpretation, we may hypothesize that he is trying to convey his despair, and when the incidence of words such as *depressed*, *lonely*, *helplessness*, etc. is measured automatically, we may score the depressivity level of the text as indicating the depressivity level of its author. However, the person may be joking, or being ironic, or simply citing something that he heard from someone else—which is why we must also take into account contextual knowledge to gain a valid conclusion about the depressivity level as expressed in the text, and whether it truly represents the depressivity level of the author. In any event, there is a wealth of evidence that the language that we use is an enormously rich mine of information for personality analysis. Other sources of information may also be identified and used if possible. For example, when analyzing depression among individuals, we may analyze their medical records and the visual images they upload to social media. In a past unpublished study, we analyzed the images uploaded to Instagram by young people—mostly women involved in self-harming behavior. It was clear from the images that these young people were depressed and self-harming: dark images, with signs of loneliness, blood, and cuts, were everywhere. An automatic image analysis algorithm could have easily classified them. Therefore, when using the term *text*, I may use it in the most generic sense to include visual images, facial expressions, body posture, and so on.

It is important to emphasize that when analyzing texts, we use a corpus of *personal* texts produced by the individuals. Why is it so important to use such personal texts? The reason is that a scientific report is probably not a good source for diagnosis, but personal texts—of the sort published in social media, diaries, stories, conversations in informal settings—are all better candidates, because in principle, at least, they reveal the individuals' *inner life*. Thus, if an accountant is preparing a financial statement for a company, we should not expect his inner life or personality dimensions to be expressed in the statement. However, if she keeps a journal, corresponds with others on Facebook, or writes a personal essay, then a personality signature should be evident. Gaining access to a personal text is a necessary step. In addition, and for the first phase of building a personality analysis system, each text is *labeled/tagged*. There are various ways in which we might do so, for example, by asking the subject who completed a personality questionnaire, by interviewing the subject, by scoring the text according to well-defined protocol and criteria.

At this point, and for each individual, we should have a personal text and personality tags and/or the specific score she has gained on each of the required personality dimensions. The text is then pre-processed, cleaned, and expanded upon using a variety of Natural Language Processing (NLP) tools, to prepare it for the main analysis. For example, we may be interested in analyzing only certain parts of speech of the text—such as nouns, verbs, or adjectives—in which case, we would use a Part-of-Speech Tagger. Next, various textual *features* are extracted from the text—such as the degree in which the person uses various words or word categories. For example, we may use LIWC [13], or Empath, [14] to measure the prevalence of positive vs. negative sentiment in the text—since a high level of negative emotion expressed in the text may be an important indicator of depressivity. Next, an ML algorithm is trained and tested to find the optimal model that can best "predict" (i.e., classify) the individuals' respective personality labels/scores. What do we mean by an optimal model? An ML algorithm is basically a sophisticated *optimization* engine. Given the tag of the text and the list of personality features and their score, it builds a classification model that assigns weights to the various features, so the classification performance is maximized. For example, the ML algorithm may show us that some features that we believed to be valuable in fact contribute nothing to the performance, and therefore can be ignored. The selection of attributes or features is therefore an important phase in constructing a successful model. Moreover, the algorithm can calculate the weights—or “importance”—that should be attributed to each feature. Different features may have a different predictive value, and the ML algorithm knows how to identify it. A computational personality project is ultimately judged by its success. The impetus for any given project is the specific of task that we would like to perform—such as choosing the best CEOs among many candidates, identifying depressed individuals, screening for lone-wolf perpetrators.

If the ML algorithm has produced good results according to some relevant standards, we can use the system. Deciding what performance is good enough must be clarified within a wider context of decision-making. For example, diagnosing PTSD through the use of human experts is costly. Let us assume that only 1% of people suffering from PTSD are diagnosed in time: if an automated system improves this diagnosis rate by 1%, should it be considered effective enough to be adopted? The answer depends on the wider context.

This, then, in a nutshell, is the essence of automatic personality analysis. In some contexts—the automatic profiling of shooters [15]; the measurement of disorders [16]; the screening of suicide ideation [17]; and the measurement of the “Big Five” personality dimensions (neuroticism, extraversion, consciousnesses, openness, agreeableness)—this general approach seems to work quite well.

In conclusion, here is an example, from a nonclinical context.

Targeted advertising is a type of online advertising that targets audiences with certain traits, based on the product or person the advertiser is promoting. By way of example, let us assume that we are developing a targeted advertisement engine that promotes music concerts. In a past study, in the context of computational personality analysis, we found a link between the lyrics of various music genres and certain personality types [18]. This finding can be used for automatically and optimally targeting advertisements, by analyzing texts written by individuals and deciding whether they are of the extrovert “Rock-n-Roll” type of person, or the introvert “Mellow” kind of personality. For a targeted advertising engine seeking to improve its performance, it may be highly informative to know whether a given individual is an extrovert or an introvert: if they are an extrovert, the engine might decide to send them an advertisement for a rock concert; if they are of the introvert type, they would get an advertisement for a mellow jazz show. In this case, determining the correct approach to the individual based on their particular personality type is justifiable.

3.4 A Critical Perspective

What are the problems in applying computational personality analysis? First, we should be careful when choosing a personality theory and personality dimensions. For example, the Big Five is a dogma with many theoretical and empirical problems

[19]. (See the paper by [20] for one possible criticism.) Therefore, although it is the main theory used in automatic personality analysis, one should critically decide whether and when to use it. Given the problem-oriented perspective that I have presented, the personality theory and personality dimensions that we choose should be carefully selected by their *clear relevance* to the challenge that we aim to address. The fact that the Big Five model is simple and easy to understand does not mean that it is relevant everywhere. For example, it is highly questionable whether it is of any relevance in the analysis of suicidal intentions. Clearly, one may find a statistical correlation between depressivity and neuroticism, since both involve negative emotion. However, the real challenge is not to identify statistical correlations of their p-values, but to construct methodologies that are meaningful in real-world challenges, by using the powerful tools of ML. From algorithmic finance to the automatic identification of lone-wolf perpetrators, one finds almost the same methodological criticisms and the same calls for a meaningful, relevant, and reality-based approach to the design of intelligent systems. In my experience of academic and non-academic/commercial projects to do with automatic personality analysis—including those in which we measured the Big Five—I must admit that, in hindsight, the Big Five model has no significant value for most real-world applications that I have encountered.

In the context of identifying suicidal intentions, for example, one may prefer the modern psychodynamic approach to personality [21], with its focus on the conflicts and defense mechanisms [22] that constitute the human personality. However, the psychodynamic approach is also fraught with difficulties, as it was designed for the clinical context. In addition, it is very difficult to translate the theory's ideas into measurable features. For example, *splitting*—seeing the world in binary terms of good and bad, black and white—is a primitive defense mechanism that some people use in order to cope with their anxieties. You can see it in action when you hear zealous ideologists—be they Islamic fundamentalists, zealous vegetarians, or fanatical BDS supporters—when presenting their worldview.

In some problem-oriented contexts, it may be important to identify the most zealous individuals—those who see the world in black and white. For example, suppose that we are interested in a new European apocalyptic sect similar to the Order of the Solar Temple, whose members committed mass suicide. Specifically, we would like to know how zealous are its members? In a case of this sort, measuring the degree of splitting within the texts (written or spoken) produced by the sect members is very important, and although it proved to be a challenge, we have shown that it is feasible to measure splitting in a text, and its relevance for the forensic context [23].

In sum, choosing the right approach and the right features is crucial. Now, by using specific examples, let me underline the problem of conducting an automatic personality analysis without due regard to the pragmatic aspect.

Lone-wolf perpetrators are a pressing issue for law enforcement agencies in the United States and in Europe. In a study conducted by [24], the researchers used "A unique dataset of 119 lone-actor terrorists and a matched sample of group-based terrorists" and compared the prevalence of mental illness (*Yes/No*) among lone-wolf terrorists and group terrorists. They found a significant difference between the two groups in this regard: among the lone-wolfs, the prevalence of mental illness was 32%, while among "group-based" terrorists it was 3%. The authors concluded that "…mental health professionals may have a role in *preventing* lone-actor terrorist attacks" (my emphasis) and that "…screening processes can be carried out by security agencies on patients that present similar antecedents and behaviors in medical evaluations." These scientifically invalid—and ethically dangerous—conclusions seem to ignore the simple lessons of reasoning, since the question is not whether there is a difference between lone-wolf and group-based terrorists, but whether mental illness is a significant risk factor and a relevant feature for intervention and prevention.

To address this question, one must ask what is the probability of someone engaging in acts of terrorism *given* their mental illness. The answer is almost nil. It might be inferred from the above study that people who suffer from mental illness pose a danger to society—but such an inference is scientifically invalid, ethically dangerous, and pragmatically irrelevant. Therefore, in the context of personality analysis, which is problem-oriented, one should clearly examine whether:

1. The findings are scientifically valid.
2. Pragmatically meaningful and usable.
3. Whether the implications are ethically justified.

3.5 Summary and Conclusions

The above critique is imperative for the reflective scientist, data engineer, or practitioner. However, critical reflections must not mask the achievements and future potential of computational personality analysis. Automatic personality analysis can have enormous benefits in improving our understanding of people in contexts ranging from screening for mental health problems, to the effective recruitment of human resources in companies. This field is still in its infancy, and there are several challenges to be addressed:

1. Most of the approaches to the automatic analysis of personality rely on *low-level features* (such as words), or their simple categorization. However, the complexity of human personality cannot be easily encompassed by low-level features alone. There is a need for more sophisticated methods that use deep syntactic-semantic analysis and infer personality dimensions through higher and more abstract features, that are extracted from the text.

Almost all the studies in the field rely on a tagged corpus, where texts are produced by individuals who are tagged according to their personality dimensions. In some cases, such corpora are *extremely difficult to obtain*—and even when they are, their artificial nature means that they lack ecological validity. In addition, their "shelf life" is limited, due to the contextual, dynamic, and changing nature of language.

2. Personality is a dynamic phenomenon that “lives” in time, and sometimes the most important information is identified by analyzing the behavior of personality dimensions along the timeline. When trying to identify whether the mental state of a teenager is moving toward a tipping-point of despair, for example, we must take the trajectory of the mental state into account.

In conclusion, most ML approaches to computational personality analysis adopt a “ready-to-wear” approach, whereby ML classifiers are trained, validated, and tested on a tagged corpus. However, as with any ready-to-wear approach, this approach is limited in its ability to provide the “client” with the best fit. The promise of computational personality analysis is huge [25], and addressing the challenge of building such a system in vivo requires reflectivity and sensitivity to various issues, such as the ones discussed above.

References

1. Corr PJ, Matthews G, editors. The Cambridge handbook of personality psychology. Cambridge, UK: Cambridge University Press; 2009.
2. Snyder M, Deaux K. The Oxford handbook of personality and social psychology. New York: Oxford University press; 2012.
3. Funder DC. The personality puzzle. 5th ed. New York: WW Norton & Co.; 2013.
4. Neuman Y. Personality from a cognitive-biological perspective. Phys Life Rev. 2014;11:650–86.
5. Neuman Y. Shakespeare for the intelligence agent toward understanding real personalities. Lanham, MD: Rowman & Littlefield; 2016.
6. Westen D, Shedler J, Bradley B, Defife JA. An empirically derived taxonomy for personality diagnosis: bridging science and practice in conceptualizing personality. Am J Psychiatr. 2012;169:273–84.
7. Freeman HD, Gosling SD. Personality in nonhuman primates: a review and evaluation of past research. Am J Primatol. 2010;72(8):653–71.
8. Streiner DL. Diagnosing tests: using and misusing diagnostic and screening tests. J Pers Assess. 2003;81:209–19.
9. Neuman Y, Cohen Y, Assaf D, Kedma G. Proactive screening for depression through metaphorical and automatic text analysis. Artif Intell Med. 2012;56:19–25.
10. Russ E, Shedler J, Bradley R, Westen D. Refining the construct of narcissistic personality disorder: diagnostic criteria and subtypes. Am J Psychiatr. 2008;165:1473–81.
11. Neuman Y. Artificial intelligence in public health surveillance and research. In: Luxton DD, editor. Artificial intelligence in behavioral and mental health care. Amsterdam: Elsevier/Academic Press; 2016. p. 231–54.
12. Neuman Y, Cohen Y, Neuman Y. How to (better) find a perpetrator in a haystack. J Big Data. 2019; https://doi.org/10.1186/s40537-019-0172-9.
13. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol. 2009;29:24–54.
14. Fast E, Chen B, Bernstein MS. Empath: understanding topic signals in large-scale text. In: Proceedings of the 2016 CHI conference on human factors in computing systems. ACM; 2016. p. 4647–57.
15. Neuman Y, Assaf D, Cohen Y, Knoll J. Profiling school shooters: automatic text-based analysis. Front Psych. 2005; https://doi.org/10.3389/fpsyt.2015.00086.
16. Neuman Y, Cohen Y. A vectorial semantics approach to personality assessment. Sci Rep. 2014; https://doi.org/10.1038/srep04761.

17. Fernandes A, Dutta R, Velupillai S, Sanyal J, Stewart R, Chandran D. Identifying suicide ideation and suicidal attempts in a psychiatric clinical research database using natural language processing. Sci Rep. 2018; https://doi.org/10.1038/s41598-018-25773-2.
18. Neuman Y, Perlovsky L, Cohen Y, Livshits D. The personality of music genres. Psychol Music. 2016;44:1044–57.
19. Block J. A contrarian view of the five-factor approach to personality description. Psychol Bull. 1995;117:187–215.
20. Molenaar PC, Campbell CG. The new person-specific paradigm in psychology. Curr Dir Psychol Sci. 2009;18:112–7.
21. Psychodynamic diagnostic manual (PDM). Silver Spring, MD: Alliance of Psychoanalytic Organizations; 2006.
22. Bowins B. Psychological defense mechanisms: a new perspective. Am J Psychoanal. 2004;64:1–26.
23. Neuman Y, Assaf D, Cohen Y, Knoll JL. Metonymy and mass murder: diagnosing splitting through automatic text analysis. In: Arntfield M, Danesi M, editors. The criminal humanities: an introduction. New York: Peter Lang; 2016. p. 61–73.
24. Corner E, Gill P. A false dichotomy? Mental illness and lone-actor terrorism. Law Hum Behav. 2015;39:23–34.
25. Neuman Y. Computational personality analysis: introduction, practical applications and novel directions. New York: Springer; 2016.