Contextualizing Personality: Insights from Anthropology to Advance Personality Detection

Niek van de Pas, Ayoub Bagheri, Willy Sier

Abstract

In recent years, personality detection — the use of computational methods to automatically determine an individual's personality from various data sources — has seen widespread adoption across a variety of fields and has been successful in achieving coherence with traditional personality tests. This paper argues that despite their widespread use, conventional personality detection methods are limited in their ability to grasp human personality. Specifically, three limitations of conventional personality detection methods are discussed: (1) their limited ability to grasp the complexity of human personality due to their reliance on pre-structured methods; (2) their inability to grasp the impact of social and cultural context on human personality, and (3) their disregard of the performative nature of human personality in online environments. Drawing on insights from anthropology and social psychology, three solutions to these limitations are proposed: (1) embracing naturalistic inquiry to capture the complexity of human personality, (2) considering the contextual influences on personality expression through multimodal methods and ethnographic research; and (3) accounting for the systematic biases present in personality in online environments in how we present our results and draw conclusions. Integrating these solutions would allow researchers to develop a more comprehensive and accurate understanding of human personality in a wide variety of fields.

Introduction

The study of personality has recently come under great interest in the growing field of personality detection, which aims to computationally determine individuals' personality traits from a variety of sources (for an overview, see Fang et al., 2022; Phan & Rauthmann, 2021; Štajner & Yenikent, 2020) — as shown in Figure 1. In particular, the same rapid developments in the field of text mining and natural language processing that have allowed technologies like ChatGPT to enter the mainstream have led to a growing body of literature on text-based personality detection, a subset of personality detection that uses text data, such as social media posts and profiles (Aung & Myint, 2019; Howlader et al., 2018;

Ong et al., 2017), essays (Kazameini et al., 2020; Mohammad & Kiritchenko, 2021), and interview transcripts (Bounab et al., 2024). Personality detection methods have shown a significant degree of coherence with personality questionnaires, demonstrating their potential in supplementing — or even supplanting — the use of traditional questionnaires (e.g., Ren et al., 2021).

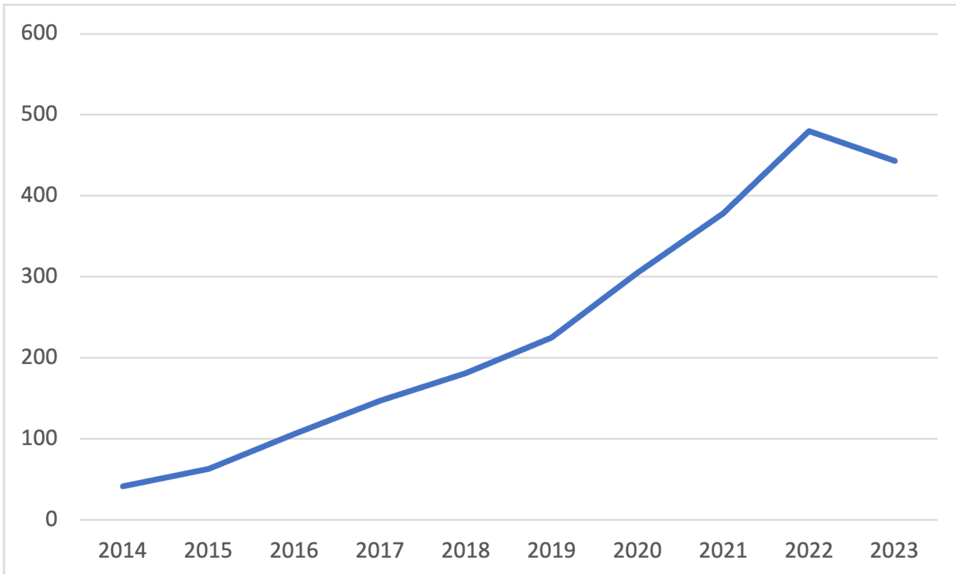

Figure 1. Google Scholar results for the phrase '"personality detection" OR "Personality computing"', by year (2014-2023). The figure shows an increasing trend, from 41 results in 2014 to 443 results in 2023

Personality detection has found applications in a wide variety of fields, including marketing and product recommendations (Chen et al., 2017; Jaimes Moreno et al., 2019; Roshchina et al., 2011; Tkalcic & Chen, 2015) and job candidate screening (Liem et al., 2018), with potential applications in many other fields, including health care counseling and forensics (Mehta et al., 2020) and voice assistants (Kazameini et al., 2020). The variety and significance of these applications raises the question: to what extent are these methods able to truly capture human personality? What are their limitations, and how can we address these? In this paper, then, three limitations of conventional personality detection approaches are discussed: (1) their limited ability to grasp the complexity of human personality due to their reliance on pre-structured methods; (2) their inability to grasp the impact of social and cultural context on human personality; and (3) their disregard of the performative nature of human personality in online environments. Drawing on anthropological understandings of 'identity', as well as on social-psychological theory on personality, this paper addresses these limitations by arguing for (1) embracing naturalistic inquiry to capture the complexity of human experience, (2) considering the contextual influences on personality expression through multimodal methods and the integration of ethnographic research, and (3) accounting for the systematic biases present in personality in online environments in how we present our results and draw conclusions.

Although the notion of 'identity' is distinct from that of 'personality', the ways in which anthropologists have conceptualized and studied identity can be of value to the development of personality detection methods, as scholars from both fields are interested in self-concept and self-expression. In particular, scholars of personality in social psychology have long recognized the role of external factors in structuring a person's personality, emphasizing the role played by cultural context (Adamopoulos & Kashima, 1999), interpersonal relationships (Veroff, 1983), and different (and potentially conflicting) identities (McAdams et al., 2021). Similarly, scholars of identity in cultural anthropology have developed a long tradition of understanding persons' identities holistically, considering the cultural context in which individuals live (Finke & Sökefeld, 2018) and their social relations (Ahmed, 2000) while also acknowledging the reality of contradictory identities (Sökefeld, 1999). Cultural practices, in this perspective, produce identity rather than solely shaping it (Hall, 2007).

This paper is structured as follows. First, the recent emergence of personality detection and the trait-psychological models that form the foundation of the discipline, as well as the methods commonly used in personality detection, are discussed. Secondly, the the range of currently employed personality detection methods are discussed. Subsequently, the way insights from anthropological approaches to identification can help us to understand personality naturalistically, contextually, and performatively, and how to address these challenges in the context of personality detection are discussed.

Finally, a note on terminology. In this paper, the term 'personality detection' is used rather than terms such as 'personality computing', since the focus is on the determining of personality traits from various data sources, and not, for instance, on the generation of appropriate text output for a chatbot given particular personality traits (e.g., Qian et al., 2017). In line with previous research, the 'OCEAN' or 'five-factor' model of personality is refered to as the 'Big 5'.

Human Personality and Trait Psychology

The study of human personality has a long history and is characterized by a wide variety of theoretical perspectives. One approach that has seen especially widespread application in both academic and non-academic fields is trait theory, which aims to capture one's personality in terms of a limited number of personality traits. These traits are understood to differ across individuals, to be relatively stable over time, and to influence behavior (Kalimeri et al., 2013). They are generally considered to exist as an entity that, while often measured through self-report, can be externally verified (McAdams et al., 2021). A variety of empirical models of personality traits exists; most popular are the five-factor model of personality, known also as the 'Big 5' or 'OCEAN' model (openness to experience, conscientiousness, extraversion, agreeableness, neuroticism) and the Myers-Briggs Type Indicator (MBTI).

Scholars of personality psychology differ on various fundamental questions, such as the importance of biological versus environmental factors (Specht et al., 2014) and the relative stability of personality traits (Asendorpf & Aken, 2003; Caspi et al., 2005; Roberts et al., 2006; Roberts & DelVecchio, 2000). In spite of this theoretical fragmentation, trait theory has seen widespread application in fields beyond personality psychology, and in both academic and nonacademic settings (Lloyd, 2012; Moyle & Hackston, 2018). In academic settings, trait-psychological models are used primarily for correlational research that investigates, e.g., the relationship between personality traits and, for instance, academic achievement (e.g., Komarraju et al., 2011; Wang et al., 2023) or mental health (Bucher et al., 2019). In nonacademic settings, on the other hand, trait theory is used primarily for recruitment and employee assessment purposes (Christiansen & Tett, 2013), though it has also found use in fields such as marketing and product recommendations (e.g. Chen et al., 2017; Jaimes Moreno et al., 2019; Roshchina et al., 2011; Tkalcic & Chen, 2015).

Traditionally, the assessment of an individual's personality traits has been accomplished using standardized questionnaires, such as the NEO–PI–3 (McCrae et al., 2005) and the Ten-Item Personality Inventory (Gosling et al., 2003). Over the years, these have become so commonplace that they have entered the collective consciousness as 'personality tests'. What all of these tests have in common is that they are time-consuming (and thus expensive) to administer (Yang et al., 2021), hindering their applicability. Additionally, traditional personality assessments are often subject to social desirability bias, where respondents portray themselves in a socially favorable light (Bäckström & Björklund, 2013; Sandal et al., 2005; although see Pelt, 2019 for a critical perspective). With "[t]he social sciences [having] entered the age of data science" (Schwartz, 2013, p. 1), however, has come the development of personality detection, also known as 'personality computing'. Harnessing the potential of computational natural language processing, personality detection methods allow researchers to automatically and computationally extract psychological traits from pre-existing (and often publicly accessible) data, such as social media posts and profiles (Aung & Myint, 2019; Howlader et al., 2018; Ong et al., 2017), essays (Kazameini et al., 2020; Mohammad & Kiritchenko, 2021), and interview transcripts (Bounab et al., 2024). Since personality detection methods have demonstrated strong coherence with traditional personality assessments (e.g., Ren et al., 2021), this has, to some extent, obviated the need for personality questionnaires.

The Landscape of Personality Detection Methods

The field of personality detection is characterized by a wide variety of methods, the development of which remains a source of scholarly attention (for a recent overview, see Perera & Costa, 2023). Personality detection originates from 'affective language processing', a subfield of computational linguistics that focuses on the computational analysis of subjective features of text. Early work in personality detection focused on classifying author personalities from creative texts, such as blog posts (Oberlander & Nowson, 2006) and essays (Mairesse et al., 2007): what is arguably the first work in the field of personality detection managed to correlate linguistic style with author personality in dairy entries, writing assignments, and journal abstracts (Pennebaker & King, 1999). Subsequently, scholars developed a variety of supervised machine-learning methods, many of which are still in use in one form or another (Fang et al., 2022). Supervised machine learning algorithms (or 'supervised learning') constitute a subset of machine learning that involves the 'training' of computer algorithms on human-annotated data ('labeled data'). During the training process, the algorithm learns to recognize particular (textual) patterns that map onto particular personality traits. As an example, an algorithm may be trained on social media posts that are labeled as particularly exemplifying extraversion or neuroticism. After training, this algorithm can detect these patterns on social media posts it has thus far not seen. These supervised methods have seen widespread success in achieving coherence with personality assessments (e.g. Evin et al., 2022). Their major downside, however, is their dependence on labeled training data, which can be time-consuming and expensive to obtain.

In contrast to supervised methods, Celli & Poesio (2014) have pioneered the use of unsupervised methods in personality detection. These aim to find naturally existing clusters or patterns in data and thus have the advantage that they do not require data that has been labeled by humans ahead of time, eliminating manual annotation effort. Downsides of unsupervised approaches, however, include the fact that they may be more difficult to interpret and evaluate, as well as their sensitivity to noise (Watson, 2023) and the fact that they require more data.

A third and final group of personality detection methods that has seen widespread application is termed 'multimodal' methods. Rather than focusing on a single form (or 'mode') of data, such as text, multimodal approaches integrate several forms of data, such as audio and video (Kindiroglu et al., 2017; Pianesi et al., 2008; Sidorov et al., 2014), audio, video, and text (Alam & Riccardi, 2014; Güçlütürk et al., 2016; Milić & Mladen, 2023), or a variety of smartphone data, such as anonymized call and SMS logs and Bluetooth and app usage (Chittaranjan et al., 2011). The major advantage of multimodal approaches is that they achieve higher accuracy than methods based solely on text.

It is clear, then, that over the past two decades, personality detection has been a source of major scholarly attention and technological development. The following section will explore the limitations of personality detection, and how insights from anthropology and social psychology could further enhance the field's ability to capture the complexity human personality.

Three Solutions for Advancing Personality Detection

The previous section has shown how personality detection methods based on trait-psychological models — and particularly the Big 5 — have achieved a significant degree of success in determining personality from written texts, such as social media posts, essays, and interview transcripts. The advantage of these automatic personality detection methods is that their alternative — the personality questionnaire — is time-consuming to administer, making computational personality detection more cost-efficient and scalable. The utility of personality detection methods has led to their rapid adoption, urging us to critically consider their conceptual foundations and methods. In this section, then, drawing on anthropological inquiry into 'identity', three limitations of conventional personality detection methods are discussed, aiming to illustrate the necessary steps for scholars in personality detection to develop a more comprehensive understanding of human personality.

Although the notion of 'identity' is distinct from that of 'personality', scholars in personality psychology and social psychology have recognized the interconnectedness between the two. Understood as a process of social interaction, rather than as something one 'possesses' (Buckingham, 2008), the notion of 'identity' in personality psychology "[shifts] attention to the outside social and cultural world" (McAdams et al., 2021, p. 3), and allows us to more fully contextualize the person and their personality (McAdams et al., 2021). Scholars such as Mark Snyder (Deaux & Snyder, 2018; Snyder, 2006) have previously argued for a closer integration of the disciplines of social psychology and personality psychology, claiming that in order to understand persons, we need to understand them as social beings, as people's personalities are shaped by the social situations they find themselves in. Incorporating this perspective into personality detection research may provide novel directions for future research and provide a more comprehensive and nuanced understanding of people's personalities. This raises the question: what insights might a reorientation towards identity offer for personality detection methods? In what follows, three potential directions for personality detection are discussed that may be addressed by incorporating these perspectives: personality detection without pre-structuring, contextual personality detection, and performance-sensitive personality detection.

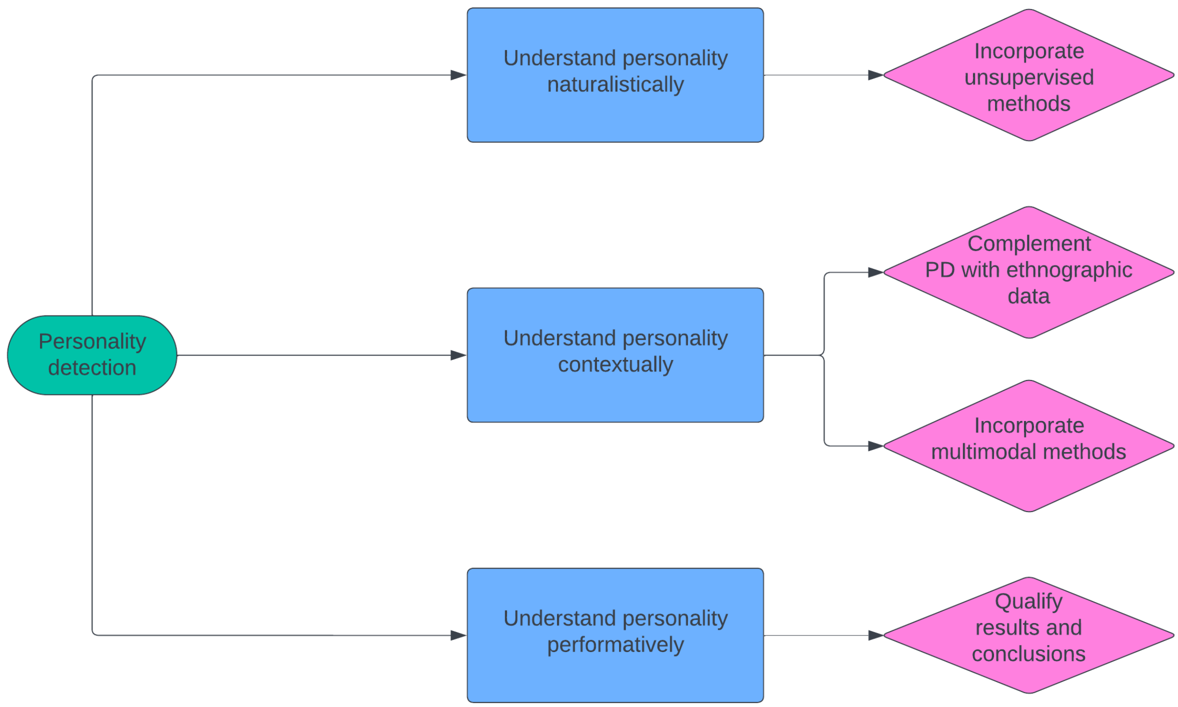

Figure 2. Proposed personality detection solutions

1) Aim to Understand Personality Naturalistically

A first important facet of the anthropological study of identity, and of anthropological methods in general, is its commitment to what Joost Beuving and Geert de Vries call 'naturalistic inquiry': "[the] study [of] social life as it presents itself to the members of a society under ordinary, everyday circumstances" (2015, p. 37). In contrast to positivist research designs, naturalistic designs aim to be unobtrusive and reactive, and are, in a sense, 'researcher-led'. This allows researchers to observe social phenomena in their natural context, providing more insight into the complex nature of human social behavior. Positivist research designs, on the other hand, typically employ a high degree of what Verschuren (2001) calls 'pre-structuring': the systematic recording of observations into predetermined categories by using, e.g., closed questions with pre-coded answers and observational scoring categories prior to doing the observing.

Both of these approaches have their advantages. The aim of positivist, quantitative research is to generalize, and in order to generalize, one needs a basis for systematic comparison, making pre-structuring necessary. As such, it is impossible for those employing naturalistic inquiry to make generalizable claims. On the other hand, researchers employing naturalistic inquiry are more able to capture the complexity of social life. In the context of personality, this translates to a diminished ability for pre-structured research to capture personality in a manner that corresponds to an individual's lived experience, as an individual's personality is reduced to a pre-defined conceptual model, such as the Big 5. The advantage of such an approach is that it is able to systematically compare differences between individuals in a reliable and generalizable manner. The downside of such an approach is that relying on these pre-structured conceptual models may fail to capture richer insights, since pre-structured methods are not able to capture unforeseen phenomena. Such an approach may thus overlook the nuances and intricacies of individual personalities, as they are constrained within predefined conceptual models, highlighting the trade-off between systematic comparability and capturing the complexity of human experiences. Consciously weighing these approaches is crucial for advancing the field of personality detection and deepening our understanding of human behavior.

What distinguishes personality detection from traditional personality assessment methods is its reliance on naturalistic data, rather than on pre-structured data. Unlike questionnaire items with predetermined options, social media posts are spontaneously generated by individuals without researcher intervention. Current approaches to personality detection aim to reduce these naturalistic data to pre-structured trait scores (such as the Big 5), trading complexity and nuance for systematic comparability, by, for instance, predicting Big 5 personality traits from Facebook statuses (Liu et al., 2016). The value of such approaches notwithstanding, personality detection that prioritizes understanding personality without pre-structuring would be able to capture more of the dynamic and multifaceted nature of human experience, thus bridging the gap between systematic comparability and the fluid reality of human identity.

One approach would be the use of machine-learning methods not aimed at reducing naturalistic data to a predetermined trait model. Instead, these methods could identify patterns and clusters within the data itself. Methods such as topic modelling and clustering could unearth themes within the data itself without using pre-structuring. By embracing naturalistic inquiry in this manner, personality detection methods could offer a different perspective on human personality.

2) Aim to Understand Personality Contextually

A second distinguishing element of anthropological approaches to understanding identity, and a second potential avenue for personality detection research lies in anthropologists' understanding of identity as contextual. Anthropologists generally prefer the term *identification* over identity, highlighting how identities are positional, fragmented, not disparate, and always 'in progress' (1996), and how identity does not signify some eternally stable self (Brubaker & Cooper, 2000). Here, identification is a process that always happens in relation to other social beings, and is therefore heavily context-dependent. Trait-psychological models of personality start from the inverse assumption: they assume that one's personality remains stable across time and across different contexts (e.g. Bleidorn et al., 2021; Cobb-Clark & Schurer, 2012; Stein et al., 1986). Scholars in personality psychology, however, have long emphasized the context-dependent nature of personality, showing the ways personality changes after unemployment (Boyce et al., 2015), between different historical, cultural, developmental, organizational, and interpersonal contexts (Veroff, 1983), and across different everyday situations (Fleeson, 2001). Indeed, the recognition of the significance of context dates back as early as 1936 with Kurt Lewin's proposition that behavior is a function of both the individual and their environment (Lewin, 1936). This may also hold online, as the digital contexts of different online communities determines which behaviors are considered 'deviant' (Fichman & Sanfilippo, 2015).

In fact, the very notion of 'personality' as understood in trait-psychological models may be culturally contingent. As De Raad (1998) shows, the translation of trait terms such as 'agreeableness' to other languages may not be at all straightforward, and while the Big 5 has proven to be useful in WEIRD (Western, educated, industrialized, rich, and democratic) populations, its validity outside this context is by no means certain (Gurven et al., 2013; Laajaj et al., 2019). This implies that the cross-cultural application of trait-psychological frameworks is not unproblematic. For text-based approaches to personality detection, this problem is compounded by the issue of 'low-resource' languages — those for which little training data is available. With less training data available, the quality of machine-learning models suffers, diminishing these models' ability to accurately and comprehensively capture human personality. An approach to personality detection that takes into account the contextual and temporal dynamics that shape human personality may enhance our ability to adapt personality assessments to diverse contexts, thus providing more nuanced insights into the dynamics of human personality and fostering a more culturally inclusive approach to personality detection.

How can we address the problem of context in personality detection? This, of course, depends very much on what we mean by the word 'context'. If we aim to capture the dynamics present in small-scale behavioral and social contexts, it may be fruitful to use multimodal methods, which, as illustrated earlier, consider various forms of data. For instance, in one study, Pianesi et al. (2008) incorporated acoustic features into their model, by which they are able to capture the various verbal characteristics of different person-to-person interactions. Additionally, Kalimeri et al. (2013) used sociometric badges to capture body movements, speech features, interpersonal proximity line of sight, and face-to-face interactions, allowing them to capture what they term 'multimodal social context'. The incorporation of these socio-contextual factors into a text-based personality model may allow us to develop a greater understanding of social contextual factors.

A second path forward, and a way to incorporate the notion of 'context' more expansively, is the incorporation of ethnographic data and methods in personality detection. Ethnographic methods, which "[seek] to holistically understand and express the lived experiences of actors in a sociocultural context" (Paff, 2022, p. 8), excel at capturing the nuances and complexities of human behavior. While ethnographic and computational methods may seem at odds with one another, their incorporation is not entirely new, and, as Nelson (2021) has pointed out, both methods share a similar inductive logic of data gathering, data analysis, and theory development. Albris et al. (2021) have previously emphasized the potential value of creating quali-quantitative 'thick' datasets through the logbook method, where researchers record observations and reflections over time. Integrating these methods with or into personality detection would provide a way to understand personality traits within broader sociocultural frameworks and thus enrich our understanding of how personality manifests in different contexts.

3) Aim to Account for the Performativity of Personality

A third and final insight from the anthropological study of identity concerns the relationship between a person's personality and the data that are collected about the person. Social scholars have long emphasized the crucial role of 'performativity' in social life. Introduced by gender scholar Judith Butler (1990), performativity refers to the ways in which both verbal and nonverbal communication serve to define one's identity, and has its roots in the works of literary theorist Kenneth Burke (1945) and, especially, of sociologist Erving Goffman (1959). Goffman referred to 'frontstage' and 'backstage' behavior to analyze how behavioral norms are internalized and 'performed' when others are watching, implying that our behaviors are not a direct reflection of our internal workings, but rather, that our behaviors are mediated by what we presume others' expectations to be (Goffman, 1959).

The notion of performativity is especially relevant in the context of personality detection using social media data. Psychologists have frequently highlighted the nature of social media feeds as a 'highlight reel' of private life, referring to the fact that individuals tend to showcase only or primarily the positive aspects of their personal lives (Faelens et al., 2021; Steers et al., 2014). It is important to take this into account when interpreting social media data, which cannot be unproblematically understood as a neutral reflection of one's internal personality. Personality detection methods frequently derive personality traits from the text individuals post on social media, suggesting a direct relationship between the text one produces and the underlying personality of the individual. For instance, Štajner & Yenikent (2020) claim that Facebook "[likely] contains more personal statements and is thus more suited for automatic text-based personality detection." However, the concept of a 'personal statement' is open to interpretation. Similarly, Howlader et al. (2018) claim that "social media personalities of users mirror their true personalities". (p. 340), and Ong et al. (2017) aim to "[extract] the personality trait of the user" (p. 65) from social media data.

It remains an open question, however, to what extent such a direct relationship between text and individual personality can be reliably inferred, and while social media data can offer valuable insights into individuals' personalities, it is wise to take caution in drawing conclusions on users' personalities from these data. Research suggests that personality traits differ significantly between on- and offline contexts, as well as from social medium to social medium. Taber & Whittaker (2018) find individuals' personality traits on Facebook to be less neurotic, open to experience, and agreeable than offline personality, while personality traits on Snapchat are more extraverted as compared to offline personality. Additionally, 'Finsta' accounts — secondary Instagram accounts with fewer followers where users post more candid or unfiltered content — were more socially undesirable, less conscientious, and less agreeable, possibly due to differing audience perceptions (Taber & Whittaker, 2020). Research has also suggested that these differences are gendered: women's perceived higher agreeableness and extraversion is more pronounced on social media than offline, while women's perceived higher neuroticism than men is less strong on social media than offline (Bunker et al., 2021).

Taken together, these findings suggest that we would do well to recognize the performative nature of online interactions and the curated nature of social media profiles. Rather than viewing social media data as a direct reflection of one's internal personality, we should approach it as a way to "analyze how users performatively and strategically express their identities" (Xi et al., 2022, p. 1437). Additionally, measuring and accounting for the systematic biases that are present in personality expression online is an essential step that would increase the validity of our conclusions,

Conclusion

Personality detection methods have recently come to see widespread use in a variety of fields, and have achieved a high degree of coherence with traditional personality questionnaires. This paper has outlined three limitations of conventional personality detection methods in grasping human personality: their limited ability to grasp the complexity of human personality due to their reliance on pre-structured methods; their inability to grasp the impact of social and cultural context on human personality, and their disregard of the performative nature of human personality in online environments. Drawing on insights from anthropology and social psychology, three solutions to these limitations are proposed. Each of these also raises particular methodological challenges that provide fertile ground for future research.

Firstly, while positivist research designs aim for systematic comparison and generalization through pre-structuring, they may overlook the complexity of individual personalities. Embracing naturalistic inquiry and developing models that go beyond pre-structuring would allow researchers to better capture the complexity of human personality. However, challenges may arise in interpreting the unstructured data and ensuring reliability.

Secondly, researchers should consider the impact of social and cultural context in personality expression, as has been emphasized by scholars in personality psychology. This can be achieved by incorporating multimodal and ethnographic methods that take into account different kinds of contextual clues. One challenge is the variety of multimodal and ethnographic methods available, and the decision-making process involved in selecting the most appropriate methods for a particular research context.

Acknowledging and addressing the systematic biases inherent in personality expression in online environments is crucial for accurately presenting results and drawing conclusions. Research has consistently shown differences in personality expression between online and offline contexts, with platforms like social media often amplifying extraverted traits. By accounting for the 'highlight reel' effect and considering how it influences personality expression, personality detection methods can better capture the true essence of an individual's personality across various contexts, ensuring more robust and reliable results.

Incorporating these insights into personality detection can help us develop more comprehensive, nuanced, and inclusive understandings of human personality for personality detection. Although addressing the limitations of current personality detection methods is a challenging task, it is essential for developing a more comprehensive understanding of human personality, which is especially crucial considering the growth of social media and the increasing accessibility of machine learning technology. It is imperative that we continue to refine our methods and deepen our understanding of human personality, ultimately enriching not only the field of personality detection but our general comprehension of human nature as well.

Bibliography

Adamopoulos, J., & Kashima, Y. (1999). *Social psychology and cultural context*. Sage.

Ahmed, S. (2000). *Strange encounters: Embodied others in postcoloniality*. Routledge.

Alam, F., & Riccardi, G. (2014). Predicting personality traits using multimodal information. *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*, 15–18. https://doi.org/10.1145/2659522.2659531

Albris, K., Otto, E. I., Astrupgaard, S. L., Gregersen, E. M., Jørgensen, L. S., Jørgensen, O., Sandbye, C. R., & Schønning, S. (2021). A view from anthropology: Should anthropologists fear the data machines? *Big Data & Society*, *8*(2), 20539517211043655. https://doi.org/10.1177/20539517211043655

Asendorpf, J. B., & Aken, M. A. G. van. (2003). Personality-relationship transaction in adolescence: Core versus surface personality characteristics. *Journal* *of Personality*, *71*(4), 629–666. https://doi.org/10.1111/1467-6494.7104005

Aung, Z. M. M., & Myint, P. H. (2019). Personality prediction based on content of facebook users:a literature review. *2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 34–38. https://doi.org/10.1109/SNPD.2019.8935692

Bäckström, M., & Björklund, F. (2013). Social desirability in personality inventories: Symptoms, diagnosis and prescribed cure. *Scandinavian Journal of Psychology*, *54*(2), 152–159. https://doi.org/10.1111/sjop.12015

Beuving, J., & Vries, G. de. (2015). *Doing qualitative research: The craft of naturalistic inquiry*. Amsterdam University Press. http://www.jstor.org/stable/j.ctt130h8g7

Bleidorn, W., Hopwood, C. J., Back, M. D., Denissen, J. J. A., Hennecke, M., Hill, P. L., Jokela, M., Kandler, C., Lucas, R. E., Luhmann, M., Orth, U., Roberts, B. W., Wagner, J., Wrzus, C., & Zimmermann, J. (2021). Personality trait stability and change. *Personality Science*, *2*, 1–20. https://doi.org/10.5964/ps.6009

Bounab, Y., Oussalah, M., Arhab, N., & Bekhouche, S. (2024). Towards job screening and personality traits estimation from video transcriptions. *Expert Systems with Applications*, *238*, 122016. https://doi.org/https://doi.org/10.1016/j.eswa.2023.122016

Boyce, C. J., Wood, A. M., Daly, M., & Sedikides, C. (2015). Personality change following unemployment. *Journal of Applied Psychology*, *100*, 991–1011. https://doi.org/10.1037/a0038647

Brubaker, R., & Cooper, F. (2000). Beyond "identity." *Theory and Society*, *29*(1), 1–47. https://doi.org/10.1023/A:1007068714468

Bucher, M. A., Suzuki, T., & Samuel, D. B. (2019). A meta-analytic review of personality traits and their associations with mental health treatment outcomes. *Clinical Psychology Review*, *70*, 51–63. https://doi.org/https://doi.org/10.1016/j.cpr.2019.04.002

Buckingham, D. (2008). Introducing identity. In D. Buckingham (Ed.), *Youth, identity, and digital media* (pp. 1–24). The MIT Press. https://doi.org/10.1162/dmal.9780262524834.001

Bunker, C. J., Saysavanh, S. E., & Kwan, V. S. Y. (2021). Are gender differences in the big five the same on social media as offline? *Computers in Human Behavior Reports*, *3*, 100085. https://doi.org/https://doi.org/10.1016/j.chbr.2021.100085

Burke, K. (1945). *A grammar of motives*. Prentice-Hall.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453–484.

Celli, F., & Poesio, M. (2014). *PR2: A language independent unsupervised tool for personality recognition from text*. https://arxiv.org/abs/1402.2796

Chen, H., Yin, H., Li, X., Wang, M., Chen, W., & Chen, T. (2017). People opinion topic model: Opinion based user clustering in social networks. *Proceedings of the 26th International Conference on World Wide Web Companion*, 1353–1359.

Chittaranjan, G., Blom, J., & Gatica-Perez, D. (2011). Who's who with big-five: Analyzing and classifying personality traits with smartphones. *2011 15th Annual International Symposium on Wearable Computers*, 29–36. https://doi.org/10.1109/ISWC.2011.29

Christiansen, Neil., & Tett, Robert. (2013). *Handbook of personality at work*. Brunner-Routledge.

Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, *115*(1), 11–15. https://doi.org/https://doi.org/10.1016/j.econlet.2011.11.015

De Raad, B. (1998). Five big, big five issues. *European Psychologist*, *3*(2), 113–124.

Deaux, K., & Snyder, M. (2018). *The Oxford Handbook of Personality and Social Psychology*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190224837.001.0001

du Gay, P. (1996). Introduction: Who needs 'identity'? In S. Hall & P. du Gay (Eds.), *(Eds.). (1996). Questions of cultural identity*. Sage.

Evin, M., Hidalgo-Munoz, A., Béquet, A. J., Moreau, F., Tattegrain, H., Berthelon, C., Fort, A., & Jallais, C. (2022). Personality trait prediction by machine learning using physiological data and driving behavior. *Machine Learning with Applications*, *9*, 100353. https://doi.org/https://doi.org/10.1016/j.mlwa.2022.100353

Faelens, L., Hoorelbeke, K., Cambier, R., van Put, J., Van de Putte, E., De Raedt, R., & Koster, E. H. W. (2021). The relationship between instagram use and indicators of mental health: A systematic review. *Computers in Human Behavior Reports*, *4*, 100121. https://doi.org/https://doi.org/10.1016/j.chbr.2021.100121

Fang, Q., Giachanou, A., Bagheri, A., Boeschoten, L., Kesteren, E.-J. van, Kamalabad, M. S., & Oberski, D. L. (2022). *On text-based personality computing: Challenges and future directions*. https://arxiv.org/abs/2212.06711

Fichman, P., & Sanfilippo, M. R. (2015). The bad boys and girls of cyberspace: How gender and context impact perception of and reaction to trolling. *Social Science Computer Review*, *33*(2), 163–180. https://doi.org/10.1177/0894439314533169

Finke, P., & Sökefeld, M. (2018). Identity in anthropology. In *The international encyclopedia of anthropology* (pp. 1–13). John Wiley & Sons, Ltd. https://doi.org/https://doi.org/10.1002/9781118924396.wbiea2142

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*(6), 1011–1027. https://doi.org/10.1037/0022-3514.80.6.1011

Goffman, E. (1959). *The presentation of self in everyday life*. Bantam.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/https://doi.org/10.1016/S0092-6566(03)00046-1

Güçlütürk, Y., Güçlü, U., Gerven, M. A. J. van, & Lier, R. van. (2016). Deep impression: Audiovisual deep residual networks for multimodal apparent personality trait recognition. In G. Hua & H. Jégou (Eds.), *Computer vision – ECCV 2016 workshops* (pp. 349–358). Springer International Publishing.

Gurven, M., Rueden, C. von, Massenkoff, M., Kaplan, H., & Lero Vie, M. (2013). How universal is the big five? Testing the five-factor model of personality variation among forager–farmers in the bolivian amazon. *Journal of Personality and Social Psychology*, *104*(2), 354–370. https://doi.org/10.1037/a0030841

Hall, S. (2007). 2 "CULTURAL IDENTITY AND DIASPORA." In *A reader* (pp. 131–138). Edinburgh University Press. https://doi.org/doi:10.1515/9781474470674-024

Howlader, P., Pal, K. K., Cuzzocrea, A., & Kumar, S. D. M. (2018). Predicting facebook-users' personality based on status and linguistic features via flexible regression analysis techniques. *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, 339–345. https://doi.org/10.1145/3167132.3167166

Jaimes Moreno, D. R., Carlos Gomez, J., Almanza-Ojeda, D.-L., & Ibarra-Manzano, M.-A. (2019). Prediction of personality traits in twitter users with latent features. *2019 International Conference on Electronics, Communications and Computers (CONIELECOMP)*, 176–181. https://doi.org/10.1109/CONIELECOMP.2019.8673242

Kalimeri, K., Lepri, B., & Pianesi, F. (2013). Going beyond traits: Multimodal classification of personality states in the wild. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, 27–34. https://doi.org/10.1145/2522848.2522878

Kazameini, A., Fatehi, S., Mehta, Y., Eetemadi, S., & Cambria, E. (2020). Personality trait detection using bagged SVM over BERT word embedding ensembles. *CoRR*, *abs/2010.01309*. https://arxiv.org/abs/2010.01309

Kindiroglu, A. A., Akarun, L., & Aran, O. (2017). Multi-domain and multi-task prediction of extraversion and leadership from meeting videos. *EURASIP Journal* *on Image and Video Processing*, *2017*(1), 77. https://doi.org/10.1186/s13640-017-0224-z

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The big five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, *51*(4), 472–477. https://doi.org/https://doi.org/10.1016/j.paid.2011.04.019

Laajaj, R., Macours, K., Pinzon Hernandez, D. A., Arias, O., Gosling, S. D., Potter, J., Rubio-Codina, M., & Vakis, R. (2019). Challenges to capture the big five personality traits in non-WEIRD populations. *Science Advances*, *5*(7), eaaw5226. https://doi.org/10.1126/sciadv.aaw5226

Lewin, K. (1936). *Principles of topological psychology*. McGraw-Hill.

Liem, C. C. S., Langer, M., Demetriou, A., Hiemstra, A. M. F., Achmadnoer Sukma Wicaksana, S., Born, M. Ph., & König, C. J. (2018). Psychology meets machine learning: Interdisciplinary perspectives on algorithmic job candidate screening. In H. Jair Escalante, S. Escalera, I. Guyon, X. Baró, & Y. Güçlütürk (Eds.), *Explainable and interpretable models in computer vision and machine learning* (pp. 197–253). Springer. https://doi.org/10.1007/978-3-319-98131-4_9

Liu, Y., Wang, J., & Jiang, Y. (2016). PT-LDA: A latent variable model to predict personality traits of social network users. *Neurocomputing*, *210*, 155–163. https://doi.org/https://doi.org/10.1016/j.neucom.2015.10.144

Lloyd, J. B. (2012). The myers-briggs type indicator® and mainstream psychology: Analysis and evaluation of an unresolved hostility. *Journal of Beliefs & Values*, *33*(1), 23–34. https://doi.org/10.1080/13617672.2012.650028

Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research*, *2007*(30).

McAdams, D. P., Trzesniewski, K., Lilgendahl, J., Benet-Martinez, V., & Robins, R. W. (2021). Self and identity in personality psychology. *Personality Science*, *2*, 1–20. https://doi.org/10.5964/ps.6035

McCrae, R. R., Paul T. Costa, Jr., & Martin, T. A. (2005). The NEO–PI–3: A more readable revised NEO personality inventory. *Journal of Personality Assessment*, *84*(3), 261–270. https://doi.org/10.1207/s15327752jpa8403/_05

Mehta, Y., Majumder, N., Gelbukh, A., & Cambria, E. (2020). Recent trends in deep learning based personality detection. *Artificial Intelligence Review*, *53*(4), 2313–2339. https://doi.org/10.1007/s10462-019-09770-z

Milić, V., & Mladen, V. (2023). Trait analysis based on multimodal prediction and optimization of the output parameters: A survey. *Serbian Journal of Electrical Engineering*, *20*(2), 229–242.

Mohammad, S., & Kiritchenko, S. (2021). Using nuances of emotion to identify personality. *Proceedings of the International AAAI Conference on Web and*

Social Media, *7*(2), 27–30. https://doi.org/10.1609/icwsm.v7i2.14468

Moyle, P., & Hackston, J. (2018). Personality assessment for employee development: Ivory tower or real world? *Journal of Personality Assessment*, *100*(5), 507–517. https://doi.org/10.1080/00223891.2018.1481078

Nelson, L. K. (2021). Leveraging the alignment between machine learning and intersectionality: Using word embeddings to measure intersectional experiences of the nineteenth century u.s. south. *Poetics*, *88*, 101539. https://doi.org/https://doi.org/10.1016/j.poetic.2021.101539

Oberlander, J., & Nowson, S. (2006). Whose thumb is it anyway? Classifying author personality from weblog text. *Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions*, 627–634. https://aclanthology.org/P06-2081

Ong, V., Rahmanto, A. D. S., Williem, W., & Suhartono, D. (2017). *Exploring personality prediction from text on social media: A literature review*. *9*, 65–70.

Paff, S. (2022). Anthropology by data science. *Annals of Anthropological Practice*, *46*(1), 7–18. https://doi.org/https://doi.org/10.1111/napa.12169

Pelt, D. (2019). *Social desirability in self-report personality questionnaires for personnel selection: Friend or foe?* [PhD thesis]. http://hdl.handle.net/1765/118659

Pennebaker, J., & King, L. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*(6), 1296–1312. https://doi.org/10.1037//0022-3514.77.6.1296

Perera, H., & Costa, L. (2023). Personality classification of text through machine learning and deep learning: A review. *International Journal For Research In Advanced Computer Science And Engineering*, *9*(4), 6–12. https://doi.org/10.53555/cse.v9i4.2266

Phan, L. V., & Rauthmann, J. F. (2021). Personality computing: New frontiers in personality assessment. *Social and Personality Psychology Compass*, *15*(7), e12624. https://doi.org/https://doi.org/10.1111/spc3.12624

Pianesi, F., Mana, N., Cappelletti, A., Lepri, B., & Zancanaro, M. (2008). Multimodal recognition of personality traits in social interactions. *Proceedings of the 10th International Conference on Multimodal Interfaces*, 53–60. https://doi.org/10.1145/1452392.1452404

Qian, Q., Huang, M., Zhao, H., Xu, J., & Zhu, X. (2017). *Assigning personality/identity to a chatting machine for coherent conversation generation*. https://arxiv.org/abs/1706.02861

Ren, Z., Shen, Q., Diao, X., & Xu, H. (2021). A sentiment-aware deep learning approach for personality detection from text. *Information Processing & Management*, *58*(3), 102532. https://doi.org/https://doi.org/10.1016/j.ipm.2021.102532

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*(1), 3–25. https://doi.org/10.1037/0033-2909.126.1.3

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychol Bull*, *132*(1), 1–25.

Roshchina, A., Cardiff, J., & Rosso, P. (2011). A comparative evaluation of personality estimation algorithms for the twin recommender system. *Proceedings of the 3rd International Workshop on Search and Mining User-Generated Contents*, 11–18. https://doi.org/10.1145/2065023.2065028

Sandal, G. M., Musson, D., Helmreich, Robert. L., & Gravdal, L. (2005). Social desirability bias in personality testing: Implications for astronaut selection. *Acta Astronautica*, *57*(2), 634–641. https://doi.org/https://doi.org/10.1016/j.actaastro.2005.03.011

Schwartz, J. C. A. K., H. Andrew AND Eichstaedt. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLOS ONE*, *8*(9), 1–16. https://doi.org/10.1371/journal.pone.0073791

Sidorov, M., Ultes, S., & Schmitt, A. (2014). Automatic recognition of personality traits: A multimodal approach. *Proceedings of the 2014 Workshop on Mapping Personality Traits Challenge and Workshop*, 11–15. https://doi.org/10.1145/2668024.2668028

Snyder, M. (2006). Building bridges between personality and social psychology: Understanding the ties that bind persons and situations. In P. A. M. V. Lange (Ed.), *Bridging social psychology* (pp. 187–192). Taylor & Francis.

Sökefeld, M. (1999). Debating self, identity, and culture in anthropology. *Current Anthropology*, *40*(4), 417–448. https://doi.org/10.1086/200042

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., Luhmann, M., Orth, U., Reitz, A. K., & Zimmermann, J. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality*, *28*(3), 216–230. https://doi.org/10.1002/per.1966

Štajner, S., & Yenikent, S. (2020). A survey of automatic personality detection from texts. *Proceedings of the 28th International Conference on Computational Linguistics*, 6284–6295. https://doi.org/10.18653/v1/2020.coling-main.553

Steers, M.-L. N., Wickham, R. E., & Acitelli, L. K. (2014). Seeing everyone else's highlight reels: How facebook usage is linked to depressive symptoms. *Journal of Social and Clinical Psychology*, *33*(8), 701–731. https://doi.org/10.1521/jscp.2014.33.8.701

Stein, J. A., Newcomb, M. D., & Bentler, P. M. (1986). Stability and change in personality: A longitudinal study from early adolescence to young adulthood. *Journal of Research in Personality*, *20*(3), 276–291. https://doi.org/https://doi.org/10.1016/0092-6566(86)90135-2

Taber, L., & Whittaker, S. (2018). Personality depends on the medium: Differences in self-perception on snapchat, facebook and offline. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3173574.3174181

Taber, L., & Whittaker, S. (2020). "On finsta, i can say 'hail satan'": Being authentic but disagreeable on instagram. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376182

Tkalcic, M., & Chen, L. (2015). Personality and recommender systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender systems handbook* (pp. 715–739). Springer US. https://doi.org/10.1007/978-1-4899-7637-6_21

Veroff, J. (1983). Contextual determinants of personality. *Personality and Social Psychology Bulletin*, *9*(3), 331–343. https://doi.org/10.1177/0146167283093002

Verschuren, P. J. M. (2001). Holism versus reductionism in modern social science research. *Quality and Quantity*, *35*(4), 389–405. https://doi.org/10.1023/A:1012242620544

Wang, H., Liu, Y., Wang, Z., & Wang, T. (2023). The influences of the big five personality traits on academic achievements: Chain mediating effect based on major identity and self-efficacy. *Frontiers in Psychology*, *14*, 1065554. https://doi.org/10.3389/fpsyg.2023.1065554

Watson, D. S. (2023). On the philosophy of unsupervised learning. *Philosophy & Technology*, *36*(2), 28. https://doi.org/10.1007/s13347-023-00635-6

Xi, Y., Chen, A., & Zhang, W. (2022). The expression of cultural identities in hong kong's anti-extradition law amendment bill movement: A semantic network analysis of tweets. *Social Science Computer Review*, *40*(6), 1436–1455. https://doi.org/10.1177/08944393211012267

Yang, F., Quan, X., Yang, Y., & Yu, J. (2021). Multi-document transformer for personality detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(16), 14221–14229. https://doi.org/10.1609/aaai.v35i16.17673