SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I11P106

Volume 13 Issue 11, 103-128, November 2026
© 2026 Seventh Sense Research Group®

Review Article

Evolution of Personality Trait Detection: From Machine Learning and Deep Learning to Multimodal and Foundation Models

M Yasmeen Banu[1], T. Sudalaimuthu[2], N. Muthuvairavan Pillai[3]

[1]*Department of CSE, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India.*
[2,3]*Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India.*
[3]*R.M.D Engineering College, Chennai, Tamil Nadu, India.*

[3]*Corresponding Author: muthuvikni@gmail.com*

Received: 15 September 2026 Revised: 16 October 2026 Accepted: 23 October 2026 Published: 27 Novembr 2026

Abstract *- Personality Trait Detection is an interdisciplinary field of study combining the fields of psychology and artificial intelligence, machine learning, NLP, computer vision, affective computing and HCI. Advancements in machine learning and deep learning have transformed the traditional questionnaire-based and handcrafted-feature approach to personality assessment into an automated, multimodal and data-driven prediction. This survey systematically looks at research mostly published between 2022 and 2026, with a focus on research evolution, data sources and modalities, approaches to learning, multimodal fusion, datasets, assessment practices, applications, and future challenges. There is evidence in the literature that the shift is towards the use of principally text-based and social media, and the involvement of linguistic, visual, auditory, physiological, behavioral, contextual and clinical data. Transformers, Large Language Models (LLMs), multimodal learning, generative AI, and personality-aware human-AI systems are examples of methodological advancements that have advanced from traditional machine learning and deep neural networks. However, a number of challenges need to be addressed for these advances to be realized, such as small or biased data sets, linguistic and cultural bias, limited psychological validation, limited generalization from subject to subject, limited longitudinal modelling, multimodal alignment and missing data, and privacy, fairness, explainability and responsible deployment. Future priorities identified by the survey include: Large-scale multimodal and longitudinal datasets, cross-cultural and multilingual modelling, psychologically-based AI, multimodal foundation models, personality-aware generative AI, explainable and privacy-preserving learning, adaptive personalization, and standardized benchmarking. In general, the future personality-aware AI system should be psychologically sound, being generalizable, comprehensible, equitable, privacy-respecting and reliable, not only accurate.*

Keywords *- Personality trait detection, Personality computing, Machine Learning, Deep Learning, Multimodal learning, Transformers, Large Language Models, Generative AI, Explainable AI, Human-AI Interaction, Trustworthy AI.*

1. Introduction

The personality is a relatively stable pattern of thoughts, feelings, attitudes and behaviors that guide an individual's views, choices, interactions and reactions in different situations. Historically, personality has been measured with self-reported questionnaires and psychometric measures due to the significant individual, interpersonal and professional outcomes associated with personality. Although they provide theoretically informed assessments, these approaches require active participation and could be affected by response bias, social desirability, and infrequent assessment. As a result of the quick development of digital technology, it is now possible to deduce personality from naturally occurring behavioral data. Written communication, social media, photos, videos, audio and digital footprints are all sources of behavioral cues for computational personality analysis [1]. The goal of automated personality trait detection is to deduce personality traits from digital and visible behavioral data. Some of its potential applications include recruitment and recommendation systems, marketing, education, healthcare, psychological testing, and human-computer interaction. Personality-aware algorithms can facilitate context-sensitive interaction, adaptive learning, candidate profiling, and personalized recommendations. For example, Kamble and Kulkarni [2] employed deep learning algorithms to automatically detect the Big Five Personality Traits during asynchronous video interviews. However, personality is multidimensional, and can assume different forms in different people, situations, cultures, and contexts, so personality detection is not as easy as a categorization

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

problem. Since computational predictions are not direct measures of psychological mechanisms, they should then be interpreted as behavioral judgments. Developed from handcrafted features and classic machine-learning approaches, computational personality detection has evolved to deep representation learning. In earlier attempts, classifiers such as support vector machine, logistic regression, decision tree and random forest were commonly used in conjunction with linguistic, behavioral or statistical features. The CNN, RNN, LSTM and other architectures have enabled models to be directly trained with complex representations from text, audio, pictures, and video. This shift not only expanded the range of data types that can be detected as personality but also reduced the need for manually engineered attributes.

The next transformation is due to transformer-inspired architectures and multimodal learning. Language provides lexical and semantic information, speech provides vocal and prosodic information, and visual behavior provides postural, gestural, and facial information. Several new approaches have emerged over the past few years for integrating different modalities, such as early, late, attention-based and cross-modal fusion methods, which are more prevalent in recent studies. Sarikaya and Küçükmanısa [3] showed how the value of multimodal and culturally diverse resources is growing with the introduction of TB5I, a multimodal Turkish Big Five interview dataset that has synchronized video, audio and text data with personality annotations. Contextual representation learning has been further enhanced using transformer topologies and pretrained language models. Sarikaya and Küçşmanısa [3] showed that contextual language representations are suitable for personality prediction and compared a traditional regression model with a regression model based on BERT and TB5I. The era of LLM and generative AI has transformed personality research from just predicting personality to integrating it into human–AI interactions, adaptive communication, and personalized dialogue. Concerns like psychological validity, hallucinations, interpretability, bias, privacy, fairness, and responsible deployment are also raised by these discoveries. Likewise, new applications demonstrate a move towards personality-based decision-making. Albrecht et al. [4] explored the idea of a personalized digital representation and provided personalized job suggestions based on the Big Five qualities inferred from Instagram photos and metadata, while new multimodal AI research has focused on AI-generated digital twins.

In addition to the idea of an AI-generated digital twin, Ismail et al. [5] explored the use of Instagram photos and metadata to infer the Big Five qualities and offered personalized job suggestions. Although considerable progress has been made in methodology, there are still several problems to contend with regarding the reliability and generalizability of personality recognition. Existing datasets are often small and limited in diversity, and obtaining annotations of personae can be costly and context-dependent. Singh and Singh [6] pointed out that there should be more detailed personality characteristics and a more appropriate classification of information. Lack of regional and/or linguistic, cultural and demographic representation can further reduce the extent of cross-population generalization. Since personality can be inferred from digital footprints without an explicit statement of consent, it raises a number of questions around consent, openness, data ownership, and potential misuse, and privacy and ethical issues are therefore significant [1]. Moreover, multimodal systems and generative systems are plagued with other alignment issues, explainability, fairness, robustness, or privacy issues, and the high level of prediction accuracy does not guarantee psychological validity. An updated and comprehensive survey of the field is motivated by these problems. A thorough overview of machine learning and deep learning techniques for identifying personality traits was given by Naz et al. [55].

This study included methodologies, datasets, assessment metrics, applications and research problems. However, the research field has significantly increased due to the quick advent of multimodal datasets, Transformers, LLMs, multimodal fusion, personalized AI, and responsible AI requirements. This survey investigates personality trait detection as a blossoming research ecosystem to incorporate developments in data, modalities, algorithms, multimodal fusion, evaluation, applications, constraints and future directions. The key contributions of this survey are: (i) the chronological analysis of the research on personality detection from 2022–2026; (ii) the domain-wise classification of the studies from personality computing to NLP, multimodal AI, affective computing, psychological AI, or healthcare or human-computer interaction or responsible AI; (iii) analysis of the major data sources and modalities; (iv) the comparative analysis of Deep Learning, Transformers, LLMs, and emerging learning paradigms; (vi) the systematic analysis of datasets, personality annotations, and evaluation techniques; (vii) the analysis of application domains and the transition towards personality-aware AI. This article is organized as follows: The mental aspects of personality trait models are introduced in Section 2. In Section 3, the development and domain-wise categorization of research carried out during 2022-2026 is explored. While Section 5 addresses machine and deep learning techniques, such as Transformers and LLMs, Section 4 examines data sources and modalities.

The multimodal approaches to personality recognition and fusion are discussed in Section 6. Datasets and evaluation are explored in Section 7, while applications are included in Section 8. The current issues are presented in Section 9, open research issues and future directions are highlighted in Section 10, the general discussion in Section 11 and the survey is summarized in Section 12.

2. Personality Trait Models

2.1. The Big Five Personality Model

Personality Trait Models provide the theoretical basis for computational personality detection by defining it as measurable targets to be predicted. The structure of most people's definition of personality traits is: a fairly stable pattern of thoughts, emotions, and behaviors that distinguish people from one another. To make the automated prediction possible beyond the traditional self-report evaluation, these variables are translated into observable behavioral signals from text, speech, facial behavior, photos, social media and interaction patterns [1]. The Big Five personality model is the most widely used personality model in computer research due to its continuous dimensions that are easily turned into machine-learning prediction goals.

2.2. The Five Factor Model of Personality

The Big Five, or Five-Factor Model, is the five broad factors of personality: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. The Big Five is particularly well-suited for regression or multiple trait prediction due to its continuous trait scores used to describe a person as opposed to categorical personality systems. In recent computational studies, texts, social media data, interviews, and multimodal behavioral information were used in this methodology. For example, Sarikaya and Küçşmanısa [3] prepared a multimodal Turkish interview dataset annotated with the Big Five and compared BERT-based and traditional regression approaches. Complementary features of individual differences are captured by the five dimensions. Language choice, interests, subjects and online activities have been correlated with openness. It is also curious, creative, imaginative and open to new experiences. Task-related interaction patterns and behavioral regularity are all indicators of conscientiousness, which can be inferred from organization, self-control, accountability, perseverance and goal-directed behavior.

The importance of extraversion has a particular relevance to conversational, auditory, visual and social-interaction cues, which are linked to friendliness, assertiveness, activity, and interpersonal engagement. Hence, agreeableness is particularly relevant to its computational prediction due to its impact on collaboration, empathy, trust, politeness, and interpersonal orientation, which can be represented by communication style and social behaviors [6]. Neuroticism represents emotional instability, anxiety, stress vulnerability, and negative emotional dispositions. The algorithmic detection is harder because people can show different emotions at a given time, which may be different from how they are at all times. Therefore, those affective or psychological states of personality should be considered as ephemeral ones and should not be confused with the stable characteristics in the prediction of personality. For example, a pattern of emotional distress as detected by a model should not be assumed to represent a personality trait. This distinction is especially important when employing speech, facial expression, physiological signals, or information from social media to evaluate personality. Although in general, personality prediction is better understood not as classification into a single personality label, but as a multidimensional estimate problem, it is important to note that the Big Five provides a multidimensional representation of personality, within which the person can simultaneously exhibit varying levels across all five personality traits.

2.3. Stable vs. Unstable Personality Traits

In addition to the Big Five, there are other alternative frameworks like HEXACO and MBTI, or even clinical-oriented and categorical frameworks. The six dimensions of the HEXACO personality framework, Honesty–Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience, complement the traditional trait-based approach and might prove useful in contexts where interpersonal behavior, honesty or fairness are important. However, the accessibility of the growth of Big Five labelled datasets and benchmarks has been helping it sustain its dominance in machine learning and deep learning research. Instead of using continuous trait dimensions to characterize personality, the Myers-Briggs Type Indicator (MBTI) and related categorical methods use discrete types. This difference is directly reflected in the computational formulation: dimensional frameworks naturally produce regression or multi-trait prediction problems, whereas categorical frameworks tend to produce classification problems. While discretized outputs may be easier to comprehend, they may lose information on the continuous nature of individual differences. Also, it is necessary to differentiate the detection of personality traits from the recognition of psychological states and the assessment of personality disorders. Clinical models are concerned with clinically relevant patterns of cognition, emotion and behavior, while trait models reflect relatively stable individual differences. For example, although the signals can be observed at the same time, there is a distinction between the clinical behavioral datasets and normative personality information, with different purposes being served Aldosari et al. [8]. So, it should be clear whether computational investigations are used as a complement to clinical evaluation, to capture temporary psychological states, or to assess normal personality features. Eliminating the goals may lead to misinterpretation and misapplication.

2.4. Computational Personality Representation

In the end, the chosen psychological framework needs to be transformed into a computational target. Personality datasets tend to be continuous scores, categorical labels or multiple simultaneous outputs, and are used primarily for classification, regression and multi-trait prediction tasks. In classification-based prediction, people are classified into groups based on thresholds and/or pre-defined personality

types. Assessment of classification is very easy and can be carried out by using accuracy, precision, recall and related metrics. However, categorizing continuous personality ratings into discrete categories can lose information about individual variations and set up artificial boundaries.

The model calculates continuous trait scores in regression-based prediction. This works particularly well for the Big Five, since the dimensions of personality are on a continuum. Sarikaya and Küçükmanısa [3], for example, provided a comparison of context transformer representations based on BERT with the regression models Random Forest, Lasso, Ridge, ElasticNet and Decision Tree on Big Five prediction to illustrate the increased adoption of contextual transformer representations for continuous personality assessment. Because of the multidimensionality of the personality, multi-trait prediction requires multiple outputs along with a shared format to estimate multiple traits. Such formulations can leverage relationships between traits and minimize repetitive modelling, but may also have variations in terms of their prediction, reliability in annotation, or modality dependence. The significance of comparing the single-trait and multi-trait combinations to evaluate personality prediction. It is important to note that there are important differences between research and categorical labels when making comparisons, because different prediction problems might be solved by models that seem to do the same thing. In personality detection investigations, all elements of the personality framework, the annotation process, the target representation, the setting of the prediction, and the evaluation metrics must be publicly reported.

3. Research on the Identification of Personality Traits in Development

The evolution of personality trait detection studies between 2022 and 2026 demonstrates the transition from predictive personality assessment to the multimodal, contextual, personalized, generative and reliable AI. During this period, text and questionnaire-based data were replaced by social media activity, audio, visual and video behavior, physiological signals, clinical data, interaction logs and multimodal data. At the same time, the computer personality modelling field was enriched by the progress of deep learning, attention processes, Transformers, Large Language Models (LLMs), multimodal LLMs, generative AI and adaptive human-AI interaction. Moreover, fields closely related to the field of emotion recognition, mental health assessment, social interaction, cognitive behavior, engagement, psychological state modelling and human-AI adaptation have also contributed to the evolution. Much of this research did not directly predict personality traits, but did offer techniques and data sources applicable to personality models. This has led to greater progress in the literature between 2022 and 2026 in the direction of using computational models to predict specific human traits, rather than using a single trait to predict another.

3.1. 2022 Research Trends

Research from 2022 was instrumental in paving the way for the future of personality-aware AI. The use of digital behavioral footprints on top of classic questionnaires was a great step forward. Mehta et al. [1] called for psychological validity, standardized measurement, privacy and informed consent, and discussed the potential for social media activity, smartphone use, wearable technology, sensor and text activity, photos, and other behavioral signals to predict personality. Kamble and Kulkarni [2] demonstrated how automated Big Five recognition could be beneficial for e-recruitment through the use of resumes, psychometric tests and asynchronous video interviews. Multimodal modelling was another important technique. Bagwari et al. [9] had combined visual information, MBTI data, EEG/ECG signals and consumer preferences while Bennett et al. [10] had added speech, writing, facial expressions and games to Social AI. The use of the audiovisual model was further explained by Kathan et al. [11], who explained that the advantages of customizing the audiovisual model were further illustrated. Affective and therapeutic studies related to these included naturalistic non-verbal behavior Terhürne et al. [13], longitudinal physiological and behavioral sensing Zhong et al. [12] and face and bodily behavior [14]. The investigations showed the need for tailored models and multiple illustrations of behavior.

Further, studies on personality-related behavioral variables and human-centered AI have increased. Holzinger et al. [15] proposed the Personas for AI concept, while Brinkschulte et al. [16] explored the area of social intelligence and Wienrich et al. [17] demonstrated the relevance of the individual. Multimodal studies on sarcasm, emotion, conversational systems and chatbots [18-21] further cemented the correlation between affective behavior and personality-driven interaction. In the field of education, logs, eye-tracking, facial behavior, screen recordings, and self-reports have all been demonstrated to have potential. In the context of education, both behavior and multimodal learning analytics have been found to have promise: logs, eye-tracking, facial behavior, screen recordings and self-report [22, 23]. At the same time, methodological and ethical limitations came to light. The importance of privacy, consent, profiling and possible misuse was highlighted in the personality studies, while the issues of fairness and ethical automated decision making were raised in the recruitment studies [1, 24, 25]. The same dataset dependency, small sample sizes, subjective annotations, ecological validity, and sociocultural differences are observed in related affective and mental health-related research [26-28]. Also, recently, co-adaptation and personalization between humans and AI have gained relevance [11, 29, 30]. In all respects, 2022 marked the beginning of the exploration of digital behavioral,

multimodal, and application-based personality tests, and the importance of responsible and customized AI.

3.2. 2023 Research Trends

The area of research in 2023 further included generative AI, trustworthy AI, mental-health AI, modelling of multimodal human states and cognitive-behavioral analysis. Yan et al. [31] and Javed et al. [32] demonstrated the increased number of projects utilizing visual, auditory, linguistic, physiological, neuroimaging, behavioral and clinical data for human-state and cognitive-health assessment, respectively. Wiedbusch et al. [33] used multimodal behavioral, physiological, cognitive, affective and contextual representations for dynamic engagement analysis, whereas Jabeen et al. [34] and Bhatt et al. [35] emphasized multimodal heterogeneous representations. The advent of generative AI and LLMs was a significant methodological advancement. Abd-alrazaq et al. [36] highlighted the opportunities and challenges of personalizing with LLMs, while Sohail et al. [37] noted the rapid advancement of ChatGPT research and the associated issues of bias, trust, and reliability. Kowsalya and Divya [38] connected generative language modelling with story generation and proposed some applications in content generation and personality-conditioned communication. Further, reliable AI was introduced to the forefront. Khairnar et al. [39] studied domain shift, security and generalization, while Xu et al. [40] studied adversarial robustness, federated learning, uncertainty and explainability. The overall literature for 2023 has focused on the shift from individual prediction to a holistic approach to human modelling, personalization, reliance, and secure AI deployment.

3.3. 2024 Research Trends

In 2024, the methodological maturity is reflected in multimodal personality detection, attention processes, Transformers and Contextual Modelling, as well as in the explainability and adaptive interaction. Singh and Singh [6] highlighted the continued progress of AI-driven personality prediction using text and social media information. Ghassemi et al. [41] addressed the subject dependency issue using multimodal (auditory, visual, and verbal) information, placing an emphasis on generalization to unseen subjects. Attention-based multimodal modelling incorporating text, audio, and visual data was further introduced by Sreevidya et al. [42]. The integration of multimodal became more systematic. Zhu et al. [43] analyzed early, late, model-based and hybrid fusion methods, while Chen et al. [44] examined multimodal AI in situation awareness and healthcare. In related studies of affective computing, verbal, physiological, facial, body-gesture, speech, and questionnaire information were all used [45, 46]. These findings reinforced the notion of using complementary behavioral and contextual cues to best convey personality-related traits. Additional developments in explainability and heterogeneous data integration were made possible by AI in healthcare and psychology [7, 47, 48]. At the same time, the study of Transformer and multimodal foundation models expanded the possibilities for cross-modal and contextual reasoning [49, 50]. In human–computer interaction, adaptive systems that adapt to the unique characteristics of the users are increasingly common [46, 51]. Studies on explainability and adversarial robustness [52-54] also highlighted the need for trustworthy multimodal personality systems. In 2024, therefore, attention-based, subject-independent, contextual, explainable and adaptable modelling was introduced instead of multimodal prediction.

3.4. 2025 Research Trends

NLP, Transformers, multimodal psychological testing, LLMs, LMMs, generative AI, agentic AI, and trustworthy AI were all increasingly included in personality research in 2025. In Naz et al. [55] comprehensive survey of machine and deep learning models for personality trait detection, a wide range of methods, including conventional models, deep learning, Transformers, datasets, assessment, applications, and current challenges, were explored. This illustrates how the field is becoming more established as a separate field of study. The multimodal behavioral/psychological evaluation continued to increase. Hong and Xia [56] studied explainability, privacy, validation, fusion, and missing data handling by looking at behavioral, physiological, voice, visual and textual data. Essahraui et al. [57] and Shrestha and He [58] both showed the growing importance of human-behavior analysis by multiple methods. Al-Azani and El-Alfy [59] and Chang and Chen [60] highlighted several factors, including synchronization, lack of modalities, annotation quality, noise, and interpretability. The rapid growth of LLMs, LMMs, generative AI, and agentic AI was the main conceptual development of 2025.

Li et al. [61] demonstrated an increasingly powerful multimodal system, Zhao et al. [62] presented an autonomous system with a strong generative capability, El Saddik et al. [63] introduced an increasingly powerful multimodal system, Sun et al. [64] introduced an increasingly powerful autonomous system with a powerful generative capability, and Ramos et al. [65] introduced an increasingly powerful autonomous system with a powerful generative capability. Not only could these advances enable personality data to be used as a prediction target, but they also enabled the use of personality data as a method of conditioning interaction and adapting intelligent agents. The significance of explainable and privacy-preserving AI has also increased. The combination of federated multimodal learning with SHAP-based explanation was demonstrated by Jahan et al. [66] in the field of healthcare and personalized-AI research [67-69], which illustrated individualized and multimodal modelling. Education and immersive AI further contributed to personalization [70-74]. In responsible-AI research, the importance of responsible ethics, such as bias, accountability, openness, privacy, reproducibility and human

oversight, was highlighted [75-78]. This development means that instead of static personality prediction, 2025 will witness the emergence of dynamic, multimodal, contextual, generative and agentic human modelling.

3.5. 2026 Research Trends

The literature from 2026 reveals the highest conceptual development in the field towards systems that are intelligent and have a personality awareness. In direct personality study, the use of transformer-based and multimodal approaches for the prediction of the Big Five remained in use. Ismail et al. [5] explored the ability of inferring personality based on Instagram images and metadata for giving job advice, and Sarikaya and Küçükmanısa [3] analyzed the prediction of the Big Five personality traits from interview data. The studies indicate a growing trend towards combining personality recognition with multimodal and naturalistic behavioral data. Tailored and adaptive AI, using personality data, was a key development. The multimodal and collaborative human-AI processes were explored by Chen et al. [79]; personalized digital twins with text, voice and photorealistic avatars were explored by Albrecht et al. [4]; and some on personalization in conversational AI were explored by Alharbi et al. [80] and Chandra et al. [81]. Jo and Lee [82] also illustrated some more applications of personality-aware adaptation in AR-based therapy. These results suggest a conceptual shift in which personality is an input and output of the intelligent system adaptation. The foundations of multimodal emotional and psychological studies have continued. Chittepu et al. [83], Muhammad et al. [84] and Gasz et al. [85] studied combinations of language, speech, facial behavior, gestures, posture, gait and physiological signs. Mental-health and psychological AI [86-89] was then used to integrate the clinical, psychometric, voice, facial, textual, and physiological information. Such studies emphasize the need to distinguish between qualities of the person and transient affective and psychological dispositions. The significance of reliable AI has increased. Privacy, fairness, security, explainability, accountability, governance, federated learning, and differential privacy were highlighted by Ray [90] and Foster [91]. The importance of individualized modelling of heterogeneous data was further demonstrated in broader personalized healthcare and multimodal AI research [8, 92-94] that also tackled the challenges of limited labels and generalization. Overall, 2026 is the year of personality-aware AI, as personality, behavior, emotion, context and interaction are all increasingly considered in combination. The research question thus has become one of how to reliably, ethically, long-term and adaptively integrate personality data into intelligent systems as opposed to being able to predict it.

3.6. Overall Evolution from 2022 to 2026

The evolution of contextual and Transformer-based modelling, personality-aware adaptive AI, multimodal learning, foundation and generative models, and finally digital behavioral evidence, was the first five-year cycle, as given in Table 1. The first five years were digital behavioral evidence, multimodal learning, foundation and generative models, contextual and Transformer-based modelling, and personality-aware adaptive AI. 2022 focused on more behavioral data and the display of multimodal prediction, 2023 was about enriching the methodological framework with multimodal human-state analysis, generative AI and trustworthy AI, 2024 was about more systematic approaches using attention, Transformer, subject-independency and explainability, and 2025 for LLM, LMM, generative AI, agentic AI, and privacy-preserving learning, and 2026 for personality as part of adaptive and personalized intelligent systems.

Table 1. Overall evolution from 2022 to 2026

Year	Dominant Research Dimension	Data/Modalities	Major Methodological Development	Major Transition
2022	Digital behavioral and multimodal personality modelling	Text, social media, audio, video, EEG/ECG, smartphones, wearables	ML/DL, multimodal fusion, personalization	Conventional assessment → digital and multimodal prediction
2023	Multimodal human-state, clinical and generative AI	Text, image, video, speech, physiological, behavioral, clinical	Multimodal DL, emerging generative AI, XAI, federated learning	Isolated modalities → integrated human modelling
2024	Multimodal personality, contextual AI and adaptive HCI	Audio, visual, verbal, physiological, behavioral	Attention, Transformers, fusion, subject-independent learning, XAI	Multimodal prediction → contextual and generalizable modelling
2025	LLM/LMM, psychological assessment and agentic AI	Multimodal behavioral, clinical, social-media and immersive data	Transformers, LLMs, LMMs, generative/agentic AI, federated learning	Prediction → dynamic and intelligent human modelling
2026	Personality-aware, adaptive and responsible AI	Interview, social media, speech, video, physiological, clinical, XR	Multimodal LLMs, digital twins, self-supervised/federated learning, XAI	Personality prediction → personality-aware intelligent systems

4. Data Sources and Modalities

Data Modalities and Sources. As the number and diversity of behavioral data have increased, so has the detection of personality traits. Research in the past five years has adopted social media data, voice, face, and video behavior, physiological signals, psychological assessment, interaction logs, mobile sensing, and multimodal data, but most of the computational research that has been conducted has focused on questionnaires and textual content. These are expressed in a variety of ways through language, through sound, through sight, through body and actions, and they are free. This makes the field transition from the previously mostly unimodal behavioral analysis to multimodal and context-aware representation learning between 2022 and 2026.

4.1. Text and Linguistic Data

Text: Linguistic and semantic information; Speech: Acoustic and prosodic information; Visual and video data: Non-verbal and temporal behavior; Physiological data: Complementary information about (internal) behavior; Social media and interaction data: Naturally occurring behavior; Psychosocial assessments: Theoretically supported reference measures.

Table 2 gives a summary of the primary data sources that were used in the present personality study. The teacher recognizes the importance of language and text. The teacher appreciates the role of language and text. Oral and written communication may reveal communication style, semantic preferences, emotional expression, opinions, and recurrent behavioral trends; hence, text remains an important resource for personality detection. Some common sources include social media posts, interview transcriptions, articles, discussions, reviews, and questionnaires [1, 12]. Text analysis has progressed in method and evolved from characteristics of the text itself (lexical, syntactic) to representation learning of the text based on context. Naz et al. [55] explained how deep learning and Transformer-based techniques are moving forward, while Singh and Singh [6] highlighted the necessity of labelled datasets, feature representation, and many personality traits. Transformer-based studies [3] have enabled the contextual modelling of personality-related language patterns, such as predicting personality types from interview data, based on the Big Five personality traits. Human behavior data and social media. Data and social media of human behavior

4.2. Social-Media and Digital Behavioral Data

Social media sites provide rich, naturally occurring behavioral traces through posts, comments, photos, metadata, engagement and interactions [1, 6]. These data can record social behavior and self-expression, which help to go beyond controlled evaluation when studying personality. Some examples include the use of Instagram photos and information for recommending careers and inferring personality by Ismail et al. [5]. However, social media behavior is influenced by the platform, the audience, the culture, and the context; therefore, these data need to be viewed as evidence of behavior and not as simple personality traits.

4.3. Speech and Audio Data

Audio and Speech Information Pitch, intensity, rhythm, speaking tempo, pauses, prosody and voice quality are all elements of speech that carry information which is absent from transcripts. These attributes can complement textual information and provide additional evidence of emotive behavior and communication [10, 11]. Further, multimodal studies have combined the physiological and behavioral indicators with the auditory information [31, 97, 98].

The personality modelling functions are of particular importance for speaker features, such as speech, language, accent, and context, since the auditory features may convey temporary emotional/situational states or relatively stable individual dispositions. The ultimate in digital video, visual, and facial information. Visual modalities are nonverbal in nature and are conveyed through facial expressions, gaze, appearance, posture, gestures, and visual self-presentation, and can be captured by video for the temporal dynamics.

4.4. Facial, Visual, and Video Data

A multimodal approach has been demonstrated to be complementary and useful in combination with language and audio [14, 20, 41, 42, 101]. The integration of verbal, visual, and auditory information, for example, as done by Ghassemi et al. [41] and Sreevidya et al. [42] and the extension of visual personality analysis to naturally occurring images, as performed by social-media research, such as by [5], can be cited as examples. Thus, video modelling can be used to model dynamic behavioral patterns such as gaze, gestures, posture and interaction style. This is the point where medical professionals can offer assistance. This is where medical practitioners can help.

4.5. Physiological and Bio Signal Data

Physiological modalities provide more information than behavior can provide, which can be observed from the outside. EEG, ECG, EMG, PPG, heart-rate variability, and breathing are examples of common signals. It has been applied in models of personality and psychological processes [9, 12, 31, 35, 33, 45] and has been the subject of multimodal research combining physiological, behavioral and self-report measures.

Physiological and behavioral integration has been further extended by more recent research [84, 85]. These signals are, however, greatly influenced by affective, physical and environmental factors; therefore, they should be seen more as supplementary data, and not as personality indicators.

4.6. Psychological and Clinical Data

Clinical and Psychological Information. An essential reference layer for computational personality modelling is provided by psychological data. Psychometric questionnaires, self-reports, psychological evaluations, cognitive tests, clinical data, and behavioral observations are some of these. Psychological and clinical AI research combines verbal, behavioral, neuroimaging and clinical data [7, 32, 47, 56] while other work uses self-reported personality information in conjunction with behavioral and environmental data [33, 35, 38]. These sources are particularly important for the validation of computational predictions against measurements for which there are theories. However, personality assessment, identification of the psychological state, and clinical disorder evaluation are different objectives and should not be confused.

4.7. Longitudinal, Interaction, and Immersive Data

Immersion, Interaction, and Longitudinal Data. Recent research has begun to consider the use of wearable technology, behavioral logs, using a smartphone to sense behavior, repeated observations, and interactive environments to capture temporal variance in individual behavior [1, 12, 67]. Personality-related behavior can be studied as well during an interaction in immersive environments such as VR and AR. Zheng et al. [72] focused on psychological evaluation in virtual environments, while Jo and Lee [82] combined gait, smart-glasses data and multimodal feedback with Personality awareness for rehabilitation. The developments make the transition from static assessment to an adaptive, context-dependent approach to behavioral modelling smoother.

4.8. Multimodal Datasets and Data Integration

Integrates data from multiple datasets across different modalities and formats. With the amalgamation of multiple data sources, multimodal datasets are growing increasingly important. These sets of data can include text, audio, face video, physiological signals, behavioral observations, social media data, questionnaires, and contextual metadata. Early multimodal research included self-reports, games, physiological sensing, voice, text, and facial behavior [10-12, 23] while later work widened the scope of modalities and integration methodologies [34, 41-43]. Besides increasing the number of modalities, the multimodal data is essential for creating complementary and psychologically meaningful representations. Therefore, data quality, temporal alignment, data completeness and validity of annotations are important factors [3, 56, 59, 60]. The problems motivate the multimodal modelling and fusion techniques introduced in Section 6.

This has resulted in a change in the data landscape from predominantly textual, questionnaire-based inputs to multimodal, longitudinal, interactive and immersive data representations. The text captures the semantic and linguistic behavior, the audio data the acoustic characteristics, visual data and video data the non-verbal and temporal behavior, physiologic data provide complementary data on the internal status, the social media provide naturalistic traces of behavior, and the psychological tests provide reference measures based on theory. The methodological issue of how to represent and combine these different sources in order to improve personality inference without sacrificing its psychological validity and generalizability is now the central question, and it is not simply one of selecting the most informative individual modality. This change is the stepping stone to understanding the deep learning and machine learning architectures used for this data.

5. Machine Learning and Deep Learning Approaches

Methods of deep learning and machine learning, Automatic personality trait detection has come a long way from the simple machine learning based on features to the deep representation learning, Transformer architectures, Large Language Models (LLMs), and Generative and Agentic AI. This is being driven by the increasing richness and diversity of data related to personality, including text, social media, audio, facial behaviors, video, physiological signs and multimodal interactions. While deep learning lessens reliance on manual feature engineering, traditional machine learning is still helpful for relatively small and structured datasets. LLMs and generative systems are adding the notions of personality, personalization, and adaptive interaction into conversational representation, while transformers provide for contextual and transferable representations. Thus, the methodological trajectory might be summed up as follows: The flow of handcrafted features, deep learning, traditional ML, transformers, LLMs, generative and agentic AI.

5.1. Traditional Machine Learning

Conventional Machine Learning: Traditional machine learning is a crucial aspect of computational personality detection. The linguistic, behavioral, auditory, visual and psychometric features have been studied using Support Vector Machines (SVM), Random Forest (RF), Logistic Regression, Decision Trees, Naïve Bayes and regularized regression methods. The Ridge, Lasso and ElasticNet are useful when personality is measured as continuous trait scores; SVM is useful in high-dimensional feature spaces; and RF can capture nonlinear correlations. Recent research indicates that conventional machine learning is still relevant. Although Kamble and Kulkarni [2] demonstrated their implementation for interview-based personality recognition, the traditional and ensemble approaches were pointed out as important aspects in personality-trait detection by Naz et al. [55]. Sarikaya and Küçükmanısa [3] compared Transformer-based methods with RF, Logistic Regression, Ridge, Lasso, ElasticNet and Decision Tree models for Big Five prediction.

Their main advantages are that they can be easily interpreted, cost-effectively calculated and used for small-sized data sets. However, their ability to learn cross-modal, contextual and hierarchical representations is limited by the manual features they can build, which is the reason for their movement to deep learning.

5.2. Deep Learning

Deep Learning enabled automatic representation learning, which enabled models to learn meaningful patterns directly from unprocessed or slightly processed data. Recurrent Neural Networks (RNNs), LSTM and Bidirectional LSTM (BiLSTM) models are used to model sequential text, voice and behavioral data, while Convolutional Neural Networks (CNNs) are extensively used for visual, facial, video and physiological data. Hybrid designs can also include multimodal, temporal, and spatial representations. CNN and recurrent networks are still applied to model personality and related human characteristics [55]. Sreevidya et al. [42] introduced an intermediate-layer dual-attention architecture to model the information of all three modalities, whereas Ghassemi et al. [41] focused on the study of multimodal personality detection based on verbal, visual, and audio information that are independent of the subjects. These experiments demonstrate how deep networks can learn complex representations in a multitude of ways in terms of behaviors. Deeper models, however, tend to require more computing resources and can over-fit on small personality datasets, particularly when there are changes to the demographics or domains. Based on these constraints, transformer designs and attention techniques were adopted.

5.3. Transformer-based Models

Models Based on Transformers: A significant step toward attention-based contextual representation learning is represented by transformers. While recurrent structures can capture the interactions between elements that are close in a sequence, they are more appropriate for long textual responses, interviews, and discussions, or multimodal sequences, where such interactions are possible between distant elements in a sequence. Transferring linguistic knowledge from big corpora to comparatively smaller personality datasets is made possible by pretrained models like BERT. Sarikaya and Küçükmanısa [3] demonstrated the applicability of BERT-based models to Big Five prediction on interview transcripts and how the manually-made linguistic characteristics were replaced with pretrained contextual representations. More broadly, Sagenela et al. [100] and Kaur et al. [89] represent new trends towards Transformer and self-supervised designs, with Transformer-based architectures also expanding to visual modelling via Vision Transformers and others. Transformers provide stronger context modelling and transfer learning capabilities, but still have several considerations to consider, such as processing power usage, the need for pretraining data, and limitations in interpretability. Moreover, the predictive accuracy of learned representations does not constitute evidence for their similarity to theoretically important personality traits.

5.4. Large Language Models

Large Language Models. The emergence of LLM has advanced personality detection beyond the simple classification/regression to context-aware and conversational personality inference. Pretrained models can be prompted, few-shot, in-context and fine-tuned to understand interviews, conversations, social media posts and descriptions of behaviors without having to learn only from small, task-specific datasets. Sohail et al. [37], Abd-alrazaq et al. [36] and Voultsiou and Moussiades [101] illustrate how LLMs can be applied in human-centered and psychological contexts. By including heterogeneous information, multimodal LLMs further expand this capability. Chandra et al. [81] investigated multimodal LLM structures, and Sun et al. [64] examined multimodal agent systems with language, vision, audio and other information sources. Given the broad context in which LLM can be applied, LLM might provide valid results even if not psychometrically valid. Thus, issues such as hallucinations, sensitivity to prompts, demographic bias, privacy, and inconsistent thinking should be addressed systematically in the context of independent data, human judgment, validated psychological tools, and longitudinal studies.

5.5. Generative and Agentic AI

Agentic and Generative AI Generative AI takes personality modelling to new heights, from prediction to simulation and even to personality-conditioned adaptation. GANs, VAEs, diffusion models, generative language models, and multimodal generative architectures are potential tools for supporting synthetic data generation, personalized content, and simulated interactions. While synthetic data can also lead to distributional bias and loss of psychological realism, Yazdani et al. [102] pointed out the potential of GANs, VAEs, and diffusion models in addressing the problem of data scarcity. Additionally, personality data can serve as a conditioning variable for customized digital representations thanks to multimodal generative systems. While Albrecht et al. [4] studied customized digital twins using text, voice, and photorealistic avatars, Zhao et al. [103] studied multimodal generative AI and cross-modal alignment. This is exemplified by the creation of multimodal and autonomous entities that are personality-aware with respect to planning, interaction and adaptability by Sun et al. [64] and Ramos et al. [65] respectively.

5.6. Methodological Evolution and Synthesis

The reviewed literature indicates that newer approaches have extended rather than completely replaced earlier methods. Traditional ML remains valuable for small datasets and interpretable baselines; deep learning provides

automated representations for complex visual, acoustic, physiological, and temporal information; Transformers improve contextual and long-range modelling; LLMs extend personality inference to conversational and multimodal settings; and generative and agentic AI enable personality-conditioned generation and adaptive interaction.

The central research question has consequently shifted from whether personality can be predicted from selected features toward whether AI can reliably infer theoretically grounded personality traits from heterogeneous behavioral evidence and use them appropriately for personalization. Addressing this question requires not only increasingly capable algorithms but also high-quality datasets, reliable psychometric annotations, longitudinal evidence, cross-cultural validation, explainability, privacy preservation, and robust generalization. The principal methodological approaches are summarized in Table 3.

All things considered, the methodological trajectory is distinguished not only by growing model complexity but also by increasingly richer representation, contextual modelling, multimodal reasoning, and adaptive personalization. However, psychological validity cannot be established by algorithmic sophistication alone. Therefore, the multimodal fusion techniques that represent and integrate diverse behavioral information for personality trait recognition are examined in the following section.

6. Multimodal Personality Trait Detection

Identification of Multimodality of Personality Traits. Personality is transmitted through many behaviors, including language, speech, gestures, facial expression, physiological responses and the way people interact. In consequence, multimodal learning, the ability to predict personality traits by learning complementary information from two or more modalities, has slowly taken over unimodal prediction. This evolution has progressed from pair-wise combinations of textual, auditory and visual data, to physiological and behavioral integration, attention-based fusion, multimodal Transformers, and finally Large Multimodal Language Models (LMMs). The motivation for multimodal modelling is due to the complementary evidence that can come from several modalities. Text captures semantic and linguistic features, while voice captures acoustic and prosodic qualities, visual data captures facial and behavioral qualities, and physiological signals describe information on interior reactions. Selection of informative modalities is not just as crucial to the recognition of multimodal personality as the representation, alignment, integration and weighting of the modalities.

6.1. Modality Combinations

Combinations of Modality. There are many possible arrangements of modalities that have been studied. Text-audio fusion integrates the voice characteristics, such as pitch, intensity, speaking speed, pauses, and prosody, with the meaning of the text. Yan et al. [31] combined verbal and audio information with visual and physiological information, while Bennett et al. [10] combined voice and text into a Social AI context. Another example of the integration of linguistic, physiological, prosodic and auditory information was given by Remya et al. [97]. In text-visual fusion, particularly relevant in the context of social networks and assessment through interviews, linguistic data is combined with facial, photographic or behavioral information. While Ismail et al. [5] looked at personality inference from Instagram photographs and metadata, Singh and Singh [6] found social media language to be a significant source for personality prediction. Audio-visual fusion involved speech with body language and facial expressions, as shown by Bennett et al. [10], who also incorporated speech and facial expression, and Kathan et al. [11], who introduced weighted late fusion, and Ghassemi et al. [41], who extended the method to incorporate subject-independent personality detection using verbal, visual, and auditory data.

The combination of text, voice and visual information has increased in significance as a multimodal arrangement since then. Sarikaya and Küçükmanısa [3] demonstrated the use of multimodal interview data for predicting the big five, and Sreevidya et al. [42] fused these modalities with an intermediate-layer dual-attention architecture. Other parts of multimodal study are physiological-behavioral fusion, including EEG, ECG, EMG, PPG, HRV, respiration, facial behaviors, voice, posture, gesture and textual data. This more extended integration is exhibited by Zhong et al. [12], Kalateh et al. [45], Muhammad et al. [84], Gasz et al. [85] and Tripathi and Alam [104]. These signals do provide additional information, but should not be interpreted as an indicator of stable personality traits as they are heavily context dependent and psychologically state-dependent.

6.2. Multimodal Fusion Strategies

Strategies for Multimodal Fusion. Multimodal performance is significantly impacted by the selection of the fusion approach. Early fusion is used to fuse the information from the modalities before prediction, which allows for joint learning, but can lead to high-dimensional representations and sensitivity to modality imbalance and missing data [34]. Late fusion models each modality independently and combine the predictions using a variety of methods, such as voting, weighted averaging and stacking. Although this offers modularity, it might not capture interactions across modalities prior to the final decision; Kathan et al. [11] give an example of weighted late fusion. Hybrid or intermediate fusion is used for heterogeneous modalities that involve a combination of modality-specific and common representations in several stages [41-43]. Attention-based fusion, which dynamically weights modalities or temporal segments based on their significance, is becoming more and more popular in more recent methods. Chittepu et al. [83]

demonstrated the broader use of attention and transformer in multimodal human interaction, while Sreevidya et al. [42] applied dual attention for auditory, text and visual information. Attention's weights should not always be taken as psychologically significant explanations, even if they can enhance adaptive integration.

6.3. Multimodal Transformers and Foundation Models

Foundation Models and Multimodal Transformers Research has moved from manually constructed fusion to learned cross-modal representation and alignment with the advent of multimodal Transformers. Transformer architectures rely on self- and cross-attention to connect information between textual, visual, auditory, etc., representations [34, 49]. Unified processing and reasoning over diverse inputs have been made possible by the later development of multimodal LLMs and LMMs. Zhao et al. [103] examined multimodal generative AI and cross-modal alignment, Chandra et al. [81] explored multimodal LLMs, and Sun et al. [64] explored multimodal modelling and agentic systems with language, vision, audio and more. These foundation-model frameworks can enable embedding of personality representations into adaptive communication and decision-making processes and the ability to infer personality from complex interaction histories. Some of the challenges they also pose are cross-modal misalignment, modality dominance, computational needs, hallucinations, bias, privacy, and limited interpretability. More importantly, multimodal foundation models must continue to be based on known trait models and proven assessment techniques since successful correlation of multimodal patterns with personality categories does not establish psychological validity in and of itself. Montessori's overall development and research consequences are covered in a number of ways.

6.4. Overall Evolution and Research Implications

The development of multimodal personality detection can be summed up as follows: Unimodal modelling, pairwise fusion, text–audio–visual integration, Physiological/behavioral fusion, Early/late fusion, Hybrid and attention-based fusion, Multimodal Transformers, Multimodal LLMs and agents. This has shifted the research from determining whether new modalities improve predictive performance to how to align, weight and interpret diverse evidence while maintaining psychological validity. It is particularly important because while many signals may be seen, such as speech, facial expression, physiological responses, and interpersonal behaviors, may be seen as dynamic and context-dependent, personality traits are more stable. Hence, multimodal fusion could be regarded as an attempt to connect across modalities and across time consistent individual features with heterogeneous behavioral data, not only to achieve higher accuracy. This process motivates the following analysis of dataset features, personality annotations, prediction settings, evaluation measures and generalization techniques in Section 7. Table 4 compiles the main fusion techniques and modality combinations that were identified in the literature that was analyzed.

7. Datasets and Evaluation

The reliability of personality trait detection is, besides model architecture, influenced by the quality, variety and psychological validity of the datasets and assessment methods. Operationalization of personality as a latent entity requires the use of self-reports, standardized psychological inventories, expert assessments or related annotation processes. Comparisons between studies are difficult because of differences in the participants, the language/culture of the data, the context in which the data was collected, the method of collection, the way the data was annotated, and the variables of the goal. In this sense, recent surveys and multimodal studies [34, 43, 55, 59, 102] point out the need for developing and evaluating datasets for personality recognition.

7.1. Dataset Characteristics

Features of the Dataset Personality datasets can differ greatly in scale, participant characteristics, language, culture, medium and collection setting. Though more observations are associated with deeper learning, Transformers, and foundation models, it is important to note that this is not necessarily synonymous with the number of participants. Participant-independent evaluation is important to assess the true generalization, as many recordings or interactions from the same participants can include participant-specific information in training and testing. Diversity in datasets is also important as it can influence behavioral and linguistic expression, such as through the influence of age, gender, education, occupation, geographic background and socioeconomic setting [1, 34, 49].

There is also an importance of cross-cultural research [1, 31, 105, 106] and multilingual and underrepresented-language modelling [107] to help distinguish general personality-related patterns from population-specific correlations, as language and culture also affect the expression and understanding of personality. There is also a trade-off between ecological validity and experimental control in data-gathering settings. Data gathered through social media, smartphone devices, wearable devices, education, clinical, and virtual environments are noisy, heterogeneous, and capture more natural behaviors, whereas laboratory and structured-interview datasets provide controlled and synchronized assessments. Some of these include longitudinal smartphone sensing [12], multimodal learning analytics [23] and psychological testing in virtual environments [72]. Therefore, future datasets require a balance of participant variety, multimodal coverage, ecological validity, experimental control and ethical data collection.

7.2. Personality Annotations

Annotations of Personality is the psychological validity of personality prediction, directly impacted by annotation quality. Self-report questionnaires are often used because they are user-friendly and are related to existing psychological classifications, but they could be influenced by social desirability, style of responding and lack of self-awareness [1, 49]. While external assessment of conduct using expert annotation is possible, it is costly and subject to inter-rater variability [31, 49]. Standardized psychological inventory, particularly Big Five measures, provides a more theoretically and psychometrically sound basis for comparison, though there may still be problems of comparability depending on the instrument and scoring method used. [6, 55] Therefore, psychological validity does not imply predictive validity. In studies, the instrument used, the scoring process, the reliability and validity data, as well as the stable or transient nature of the personality labels, should be reported. This separation is quite significant for more complex Transformers and LLM-based systems.

7.3. Prediction Settings and Evaluation Metrics

Usually, personality prediction is made using classification, regression or multi-trait prediction. For classification, which sorts individuals into categories, accuracy, precision, recall, F1-score, and AUC are typical measures of accuracy; however, information in continuous trait scores may be lost if individuals are merely classified into categories [2, 55]. Usually, regression uses MAE, RMSE, and correlation coefficients and preserves the dimensionality of personality. Sarikaya and Küçükmanısa [3] have studied continuous Big Five prediction, while Naz et al. [55] have talked about it in great detail. Multi-trait/multi-output methods can exploit shared representations and simultaneously estimate multiple personality traits in a multi-trait/multimodal learning setup [108]. The choice of metrics should take into account the features of the dataset and the prediction formulation. Macro-F1 is a more balanced categorization metric than accuracy when there is a class imbalance. RMSE gives greater weight to more significant mistakes than MAE, which is directly interpretable. Pearson or Spearman correlation is used to measure the relationship between expected and observed scores, but it cannot stand alone as evidence of minimal prediction error or calibration. Correlation should thus be reported alongside other error measures such as RMSE or MAE [3, 55].

7.4. Comparative and Generalization Evaluation

Evaluation of Comparison and Generalization. Evaluation of Comparison and Generalization, when interpreting reported performance, the following participant factors are considered: demographic information, trait inventories, label distributions, modalities, preprocessing, data partitions, and assessment measures. Cross-subject evaluation is used for novel individuals that have not been previously observed, while within-dataset evaluation would enable controlled comparison of models. Stronger evidence of transferability across populations, languages, and recording settings is provided by cross-dataset, cross-cultural, and multilingual evaluation [1, 31, 105]. In the process of ablation analysis, it is important to determine whether the improvements in multimodal systems come from improving the models or whether the individual modalities provide complementary information [42, 43, 59]. In addition, robustness evaluation should look at domain shifts and missing or noisy modalities, which frequently happen in real-world situations [43, 59, 60]. Reporting confidence intervals, repeated trials, statistical significance, and effect sizes can be used to determine the significance of observed improvements whenever possible [55]. Overall Evaluation. On average, it is hard to conclude personality detection performance without taking into account the dataset and evaluation parameters. While predictive ability has improved as models have become more complex, from typical machine learning to deep learning, Transformers, LLMs, and multimodal foundation models, we have not completely overcome the need for representative data and psychologically appropriate annotations. Three priorities of special importance are the increase in the size and diversity of datasets (in terms of demographics, languages and cultures); using validated psychological instruments with precise trait-state definitions; and moving from single-dataset performance to participant-independent, cross-dataset, cross-cultural, longitudinal, multimodal and robustness-oriented evaluation. Standardized benchmarks, including shared definitions of traits, processes of annotations, independent partitions of the participants, specifications of modalities, baseline models and reporting of results, would improve the reproducibility of results and allow meaningful comparisons of results from different studies.

8. Applications of Personality Trait Detection

Detection of personality traits, which consists of applications that detect personality traits. With the increasing capability of machines and deep learning models to learn personality from textual, visual, acoustic, physiological, behavioral and multimodal information, the task of personality trait detection has been transformed from a prediction task to a process of personalization, adaptation, decision support, and human-AI interaction. Some of the current applications include Human-Computer Interface (HCI), recruitment, career advice, recommender systems, education, healthcare, social media analysis, conversational AI, human-AI cooperation and immersive intelligent settings. In these situations, the value of personality identification is therefore primarily in the exploitation of the inferred individual traits to change the behaviors of the system, suggestions, communication or decision support. Personality assessment and job or career suitability are made by using resumes, psychometric data, interviews, and behavioral data in recruitment and career applications [1, 2].

However, due to the consequences of employment decisions, the evidence of validity, fairness, transparency, informed consent and human oversight must be strong for such systems to be effective [24]. In adaptive systems and HCI, personality information can be utilized to adjust the interface, explanation, communication style, interaction pace, and level of guidance based on individual characteristics [17, 46, 51]. Recent studies have turned towards context-aware and continuous personalization based on behavioral and contextual data [82]. Like in social media analysis, where the difference between users is derived from posts, comments, photographs, metadata and interaction patterns, recommender systems rely on the additional user representation of personality along with traditional behavioral signals [1, 3, 5, 53]. When it comes to an online presence, however, it does not necessarily reveal a consistent personality, but can rather be a form of self-presentation, platform effects, culture, and circumstances, which creates problems regarding user autonomy, privacy, permission, and profiling. Personality detection in education can be used to provide learner profiling, adaptive content, feedback, pacing, personalized learning, in multimodal learning analytics, which include self-reports, behavior logs, eye tracking, facial expressions, physiological signals, and interaction data [38, 52, 71, 88].

Educational systems need to be aware of the difference between stable personality characteristics and temporary learning states, so as not to misread context-dependent behaviors as a personality trait. Healthcare and mental health applications are increasingly incorporating other multimodal data, such as clinical, behavioral, physiological, neuroimaging, speech, textual, and personality data to facilitate customized assessment and intervention [47, 68, 86, 108]. Because of the potential impact of misinterpretation, personality detection should be utilized along with already existing clinical and psychological assessments in healthcare, not as a substitute for them. It should also adhere to strict standards of privacy, clinical validity, fairness, explainability, uncertainty and safety. Another important use is through dialogue, which provides linguistic, auditory, emotional, and interpersonal information that could aid in personalized communication.

Emotionally intelligent chatbots lay the groundwork for the development of adaptive communication in LLM systems, multimodal LLMs, and agentic systems, which utilize language, vision, audio, and interaction history to empower systems for human-like communication [7, 45, 81, 83]. Conversational systems can thus be customized for personality by modifying their tone, depth of explanation, conversational pacing or social involvement; however, inferred psychological traits should not be seen as definitive or predictable. Further, with human-AI collaboration, it is important to adapt autonomy, communication, explanations, and collaboration to personal characteristics [4, 17, 65, 83]. Likewise, the combination of behavioral, physiological, and contextual information provides opportunities for personalized engagement and evaluation in the virtual and augmented environments [4, 82, 87, 100]. Throughout these application areas, we see a shift from personality assessment and prediction to adaptive human-AI systems and personality-conditioned personalization. Personality evaluation → prediction → suggestion → adaptive interaction → personality-conditioned systems → ongoing human–AI personalization might be used to summarize this trend. It is important to note that the concept of personality has been evolving from a single dimension to a more complex one that is now associated with emotion, cognition, motivation, behaviors, health status, context, and interaction history.

While consumer recommendations can have a relatively low risk of error, hiring or health care predictions can have significant consequences. The loss caused by an incorrect personality inference varies from application to application as well. Consequently, the risk of the application should be considered when deciding the appropriate use of personality-aware AI, and the level of robustness of the use would need to be greater for higher risk applications with regard to validation, transparency, fairness, privacy protection, permission and human oversight. Further, when predictions are incorrect, adaptive systems create a closed-loop personalization process that can lead to self-reinforcing profiles, in which multimodal sensing → personality inference → context interpretation → personalization → user interaction → new behavioral data → model adaptation can create a self-reinforcing loop. In turn, future applications will involve mechanisms to correct inferred traits, user feedback, continuous validation, and uncertainty-aware representations. In general, it becomes clear that the question of prediction is becoming less important than that of applying personality information responsibly and reliably to provide assistance in the form of personalization, which is helpful, transparent and controllable.

9. Challenges and Limitations

Obstacles and Restrictions Despite its advancement in machine learning, deep learning, multimodal fusion, Transformers and large language models, reliable computational detection of personality traits remains an unsolved problem. The most common challenges go beyond the model architecture and involve issues such as dataset quality, validity of annotations, reliability of multimodal data, generalization, interpretability, privacy, fairness, longitudinal stability, and psychological validity, as well as demographic and cultural bias. First, personality is a psychological trait that is generally stable and latent, which is a fundamental challenge for computational models, given that most of them are concerned with the way the concept is manifested in behaviors (in language, in speech, in facial behaviors, in physiological responses, and in interaction

patterns). Emotion, context, culture, environment, and transient psychological states all have an impact on these observable signs. Consequently, models may be able to learn behaviors effectively without reflecting the construct of personality. One of the main issues in the field is this trait-state-context conundrum.

9.1. Dataset and Annotation Limitations

Limitations of Datasets and Annotations: There remains a lack of large, diverse, and psychometrically sound data sets. Although the social media platforms and digital systems generate huge amounts of behavioral data, there are relatively limited data sets that integrate reliable personality evaluations with multiple multimodal observations from the same subjects [43, 77, 80]. Deep-learning, Transformer, and multimodal foundation models, which need a lot of data to develop generalizable representations, are especially affected by this restriction. Also, the number of distinct people may not always be equal to the number of observations. This may be due to participant-specific information being leaked from the training set to the test set because multiple recordings were made from the same participant. Therefore, it is important to have participant-independent evaluation.

A similar problem is found in the form of dataset bias. Populations that do not meet the definition of the intended deployment site may be due to sampling method, platform selection, population characteristics, and conditions of collection. In addition, labels for personality may vary according to psychological inventory or expert rating, or may be self-rating. While observer evaluations may add inter-rater and cultural heterogeneity, self-reports might be influenced by social desirability and self-perception [1, 78]. As a result, models may inherit their ground-truth annotations' flaws. Consequently, the inclusion of participant-level diversity information, repeated observations, standardized psychological tests, clear annotation procedures and comprehensive demographic and cultural information in future datasets is encouraged.

9.2. Cultural, Linguistic, and Cross-Domain Generalization

The languages and cultures influence people's expression of personality. There may be significant differences between groups in terms of vocabulary, discourse style, emotional expression, facial behaviors, gestures, speech prosody, and questionnaire responses. Thus, models that have been trained primarily on a single language or geographic group might learn population-specific correlations [1, 41]. Expanding personality detection beyond English-language dominated datasets is emphasized by multilingual and underrepresented-language research [101, 107].

Likewise, the more modalities, the less it eliminates cultural prejudice; cultural diversity should be incorporated into training and assessment. A related one is cross-domain generalization. A system trained from lab recordings would possibly not be as effective as a system trained from social media text, and a system trained from social media text could not be as effective as a system trained from lab recordings. Changes in the devices, platforms, environment, language, and audience affect significant distribution changes [41, 43, 57, 67]. Thus, strength-based personality recognition requires cross-subject, cross-dataset, cross-platform, multilingual and cross-cultural evaluation, and not solely random train-test splits.

9.3. Multimodal Alignment, Missing Data, and Noise

Noise, Missing Data, and Multimodal Alignment Although multimodal systems offer complementary information, they also present new technical difficulties. There is difficulty in cross-modal alignment and cross-modal synchronization due to the fact that both text, audio and video signals and physiological signals operate at different temporal resolutions. This could result in incorrect links in the modalities and in the content of personality-related information being spread across different periods. [67, 80] Additionally, missing or corrupted modalities are encountered by real-world systems. Physiological sensors might malfunction, cameras might not be available, microphones might be noisy, and users might only give specific types of information.

Chang and Chen [60] and Zhu et al. [43] both point out that the following are still important issues that have not yet been resolved: missing modalities, noisy signals, alignment, and multimodal data quality. Both Chang and Chen [60] and Zhu et al. [43] list the following unresolved problems as important issues: missing modalities, noisy signals, alignment, and multimodal data quality. Thus, future models should consider adaptive modality weighting, modality dropout, uncertainty-aware fusion, cross-modal reconstruction, and robust temporal alignment to enable robust inferences across modalities. Crucially, rather than being viewed solely as a technical issue, missingness should also be looked into as a possible source of socioeconomic and demographic bias.

9.4. Interpretability and Psychological Validity

Psychological Validity and Interpretability. It is challenging to ascertain why a personality prediction has been produced due to the growing complexity of deep neural networks, Transformers, multimodal architectures, and LLMs. In psychological inference, explainability is especially important since it is not enough to have the correct answer, and the model has not learned a psychologically relevant link. Rong et al. [52] emphasized the human-centered evaluation of XAI, including trustworthiness, understandability, usability, and cooperation of the AI. There has also been multimodal clinical modelling research on SHAP-based explanations [108], and more general research on advanced AI has focused on interpretability as an ongoing

challenge [78]. The attention weights or feature importance should not be interpreted as psychological explanations, however. Future systems require stable and psychologically secure explanations that can distinguish between important behavioral cues and explanations of dataset-specific correlation, as well as connect to validated personality traits and computational evidence.

9.5. Privacy, Security, and Ethical Risks

Risk to Security, Privacy and Ethics. Personality detection is the ability to identify PTFs on the basis of behavioral data. The content of text, speech, facial photographs, physiological data, social media usage and medical records can all yield information that was never meant to be collected. It is not morally acceptable to deduce someone's personality technically. Therefore, the issues of permission and privacy are important [1]. These risks are particularly important in the recruitment and healthcare sectors, as personality predictions can have an effect on vital choices [24]. Furthermore, security issues, such as adversarial manipulation, unauthorized inference, data leakage, and others, may impact personality-aware systems [90]. Promising paths are offered by privacy-preserving strategies like federated learning, differential privacy, safe data management, access control, and privacy-by-design architectures. However, informed permission, purpose limitation, openness, and appropriate governance need to be matched by technical privacy protection.

9.6. Fairness and Responsible Use

The role of equity and conscientious use of any prediction is used in recommendations, healthcare, education, employment, or other types of decisions. Personality inference can result in discriminatory outcomes. The error rates may differ across demographic and cultural groups due to representativeness or culturally specific cues to behaviors. The more general issue of whether the inference should be drawn at all is likewise related to fairness. Strict adherence to a technically competent model may not necessarily produce an ethical assignment if people are profiled without their knowledge or consent. The analysis of sub-groups, transparency, oversight of people, and contention/rejection of automated conclusions should therefore be seen as high-risk applications [24, 90].

9.7. Trait–State Ambiguity and Temporal Variability

The analysis of temporal variability and trait-state ambiguity. Analysis of temporal variability and trait-state ambiguity. Perhaps the most basic conceptual limitation is the hard-to-differentiate distinction between transient states of behavior and enduring personality traits. Stress, exhaustion, emotion, social environment and task demand impact on speech, facial expression, physiological activity and interpersonal behaviors. Consequently, you might get a clear picture of what someone is like now, but not much solid evidence of how they will be in the future. Similarly, a higher physiological arousal does not necessarily mean that someone has a particular personality trait, and a lower level of speaking in one conversation does not necessarily mean that this person is lower on the trait of extraversion. Mehta et al. [1], Zhong et al. [12], and Wiedbusch et al. [33] demonstrate the importance of contextual and temporal information to interpret the behaviors signals. Future models ought to clearly depict: Individual characteristic that tends to be relatively permanent = Trait. No matter which of these the child's problem is, the focus should be on the problem itself. Whatever the problem is, the emphasis should be on the problem. The impact of an event, situation, or environment. This is important in order to avoid personality detection becoming a major way to categorize emotions or behavioral states.

9.8. Lack of Longitudinal Validation

Insufficient Longitudinal Verification. Most contemporary studies use cross-sectional data and provide little information about the consistency of personality predictions across time. Longitudinal data can reveal differences in expression of personality across environments, life stages or situations and whether variations observed are a result of temporary shifts in behaviors or actual differences in personality. In addition to the importance of temporal modelling and model drift as emphasized in adaptive and personalized AI studies [82, 88], Zhong et al. [12] demonstrated the advantage of longitudinal smartphone sensing. Longitudinal evaluation is particularly important for adaptive systems. This constant updating may result in low stability of user profiles or in self-reinforcing user profiles, but may also improve personalization. Thus, the continued personality modelling, the temporal stability of personality models, model drift, frequency of updates, and ways of correcting false personality representations should all be explored in further research.

9.9. Reproducibility and Benchmarking

Benchmarking and reproducibility. Comparative evaluation is still difficult due to the different datasets and personality inventories, preprocessing pipelines, train/test partitions, modalities and evaluation metrics used in personality studies. Proprietary datasets place further limitations on independent lines of replication. Standardized datasets and evaluation methods are critical, as seen in Naz et al. [55] and in multimodal studies in general [43].

Future benchmarks should consist of:

1. Participant-Independent Train/Test Partitions;
2. Standardized, well-documented personality inventories: A Series of Criteria
3. Demographic, Linguistic and Cultural Information;
4. Consistent Classification and Regression Measures;
5. Unimodal and Multimodal Baselines;
6. Cross-Subject and Cross-Dataset Evaluation;
7. Robustness Testing and Missing-Modality Testing;

8. Subgroup Fairness Analysis
9. Confidence Intervals and Statistical Uncertainty; And
10. Reproducible Implementations Under Privacy and Ethical Constraints.

With these benchmarks, one could be more certain that any improvement was due to modelling advances, as opposed to experimental setup changes.

9.10. Integrated Assessment

Combined Evaluation: The problems above are closely related. The smaller the dataset, the higher the risk of annotation bias and demographic bias; the more complex the model, the more difficult it is to be interpretable and reproduce; multimodal data adds modality alignment, missingness and noise; and the lower the cross-domain and cross-cultural generalization. The delicate nature of personality inference raises privacy, security, fairness and ethical issues. At the core, personality is a psychological phenomenon, and computational systems watch behaviors. Factors influencing behaviors at the same time include personality, emotion, context, culture, health and environmental factors. Thus, psychological validity does not necessarily mean that a good prediction is achieved. Consequently, there are three levels of success to evaluate the condition of the field: Accurately anticipating the dataset label is known as predictive success.

Generalization success is the ability to maintain performance across unseen individuals, datasets, cultures, languages and contexts. Predictions that are congruent with the personality construct are indicative of psychological success. The progress of predictive modelling in the present research is very impressive; however, the psychological interpretation and generalization are still relatively underdeveloped. Hence, predictive performance and psychometric validity, robustness, fairness, privacy, interpretability, uncertainty, longitudinal stability, and reproducibility should be taken into account in future personality identification, not targeting a single accuracy or F1-score. Only when paired with more robust psychological theory and thorough assessment will technological innovations like multimodal foundation models, self-supervised learning, federated learning, uncertainty-aware fusion, and continuous learning be beneficial.

10. Open Research Challenges and Future Directions

The course covers the following topics: The challenges described in the previous section indicate that improvements in classification or regression are not enough for the next step, namely, personality trait detection. Future research should focus on psychologically informed, multimodal, longitudinal, generalizable, explainable, privacy-preserving and adaptable representations of individual differences instead of predicting individual personalities in isolation. To improve generalization, it is important to increase the number of behavioral samples and increase the diversity of participants, not necessarily by expanding one to the other; the fundamental goal is to create large-scale, diverse, participant-level datasets with repeated multimodal observations and standardized personality assessments [43, 55, 80]. In future benchmarks, validated personality tests, transparent annotation processes, participant-independent data partitions, demographic and cultural diversity, and, when morally acceptable, complementary modalities such as text, body and facial behaviors, physiological signals and interaction data should be applied. While the generation of synthetic data can help overcome privacy and annotation limitations [102], it should be used in addition to empirical data, and not in place of it, unless the validity of the synthetic data can be demonstrated psychologically and demographically. Furthermore, it is important for evaluation to include cross-cultural, multilingual, cross-subject and cross-domain generalization since the models may learn to make predictions based on population-specific features instead of general personality traits [109].

Studies into multilingual and comprehensive language modelling [101, 107] also reinforce the importance of moving beyond the current scope of personality-aware systems, which are confined to English and to English speakers. Furthermore, the evaluation is particularly important for the subjects that are independent of their identities, as they could benefit from other identity-related attributes such as voice, facial characteristics, recording conditions, etc., or from recurring behavioral patterns [57]. Longitudinal modelling should be considered more, as personality inference based on short observation periods can be misleading due to the mistaken assumption of personality stability of stable traits as short-term psychological states or situational behaviors. These dimensions are separable by multiple behavioral observations [12, 82]. The temporal representation, continuous learning, monitoring for model drift, change detection, and uncertainty-aware updating should become part of future systems. While Large language models, foundation models, Generative AI, multimodal Transformers and agentic systems [65, 81, 83, 93] provide novel approaches to integrating multiple kinds of behavioral data, larger models do not necessarily mean psychological validity. Future foundation models of personality, then, require assessment not only from a computational perspective, but also from a psychometric perspective, including reliability, construct validity, predictive validity, calibration, uncertainty, interpretability, and cross-context validity. At the application level, however, personality inference is transitioning to personality-aware adaptive AI, as validated individual representations can influence a range of applications such as recommendations, education, personalized dialogue, virtual agents, and human-AI interaction [4, 81, 83, 93]. The systems should not only

describe personality as a determinant of behaviors, but also add current state, context, goals and interaction history. To prevent false or self-reinforcing profiles, uncertainty-aware adaptation, user feedback, reversible personalization, profile correction and explicit user control will be key in the implementation. The privacy, fairness, explainability, consent, security and potential misuse concerns with personality inference are serious, particularly in important applications such as hiring, education, healthcare, etc. [24, 90, 108]. Explanation quality should be evaluated based on the following three criteria: fidelity, consistency, and human comprehensibility, but should also clearly address the role that, on the one hand, linguistic, acoustic, visual, physiological, and interaction evidence plays in the prediction of personality traits, and on the other, psychological plausibility [52, 78, 108]. In the case of high-impact applications, informed consent, transparency, human oversight, user control and procedures for rectifying inaccurate personality profiles should apply. Fairness should be measured at the language, cultural and demographic level [24]. In addition to their impact on predictive utility, there are also methodologically defined ways to minimize the need to centralize sensitive behavioral data with privacy-preserving approaches such as federated learning, differential privacy, secure aggregation, decentralized learning and privacy-preserving representation learning [108]. Similarly, standardized assessment tools are needed for comparative studies of different studies.

Future benchmarks should include cross-subject, cross-domain, cross-cultural, longitudinal, external and cross-subgroup, cross-modality, uncertainty, calibration and classification/regression performance reporting, as well as confidence intervals. Moreover, if a processing or an implementation is both morally and practically possible, the promotion of reproducible preprocessing and implementation should be encouraged from small and task-specific datasets to multimodal, longitudinal resources; from monolingual and subject-dependent modelling to multimodal, cross-cultural and participant-independent systems; from static prediction to trait-state-context modelling; from feature-driven approaches to psychologically grounded representations; from task-specific architectures to multimodal foundation models; from personality inference to explainable, privacy-preserving, user-controlled, personality-aware generative AI. Hence, besides accuracy, F1-score, or correlation, the other performances, such as the prediction's psychological validity, generalization ability, robustness, fairness, privacy, interpretability, uncertainty awareness and longitudinal stability, should be considered to assess future progress. The ultimate vision is to unify breakthroughs in machine learning, multimodal AI, foundation models, psychology, behavioral science, human–AI interaction, and responsible AI to design personality-aware intelligent systems that capture individual differences scientifically and leverage them for safe, adaptive, and morally sound interactions.

11. Discussion

There is a profound change in the way personality is being detected and the role it plays in AI, as seen in a comprehensive examination of personality trait detection research from 2022 to 2026. In the field, relatively task-specific, single-modality behavioral prediction has now been replaced by multimodal, context-aware, longitudinal, adaptive and foundation-model-based approaches. This change is part of a larger trend in personality prediction, from assigning labels to personality to developing computational models of individual differences. Potential of deep neural networks, Transformers, large language models, multimodal foundation models and generative AI to model behavior data that is heterogeneous across modalities such as text, speech, facial and body behaviors, physiological signals, social media activities, and interactions is higher, but traditional machine-learning techniques remain useful for smaller datasets and interpretable baselines.

Due to the different time resolutions, reliability, noise and psychological significance of various modalities, the integration of more than one modality is not just about adding more data but rather involves issues of effective alignment, fusion, missing modality, and interpretation. The study lays a foundation for a trait-state-context approach to future personality modelling and highlights the importance of distinguishing stable personality states from transient emotional and contextual experiences. Improved prediction accuracy can be achieved with increasingly sophisticated models, although this is not necessarily an indicator of psychological validity or cross-domain generalization. Three complementary characteristics should be used to gauge progress: predictive success, generalization across people, across cultures, across domains, and across situations; and psychological success, through representations that are theoretically significant and empirically validated.

Some of the main reasons why data quality is still a big problem are the lack of participant diversity, inconsistent personality annotations, small populations, not enough longitudinal observations, multimodal synchronization, and missing data. Most importantly, generalization in a small number of behavioral observations is not always found when observed by a large number of participants. This has expanded the application space to include things like recruitment, education, recommender systems, healthcare, conversational AI, human–computer interaction, immersive environments and adaptive human–AI collaboration. However, the threats of personality inference are quite different across these contexts. Hence, it is required to have a stronger validation, transparency, fairness, privacy protection, permission, uncertainty estimation and human oversight in high-impact applications, as compared to low-risk personalization systems. The growing body of research indicates that personality prediction is being replaced by personality-aware intelligence, that is, incorporating the

state, preferences, behaviors and interaction history of the environment to assist with adaptive systems.

It is important to note that the principles of trustworthy AI must be followed throughout the development life cycle, as there is also the potential for false personality assumptions to be propagated by subsequent decisions and interactions. Several aspects should be considered in the evaluation of future personality-aware systems, such as predictive performance, psychological validity, generalization, robustness, fairness, privacy, interpretability, calibration, uncertainty, and longitudinal stability, to mention a few. To complement the future advances in multimodal AI, Transformers, LLMs, generative models, and intelligent systems, all the evidence from 2022 to 2026 points to the need for rigorous psychological theory, representative and longitudinal data, cross-cultural validation, responsible data practices, and efficient human oversight. In future, instead of looking for the optimal algorithm, the detection of personality traits will be based on combining advanced computational modelling with reliable psychological testing to develop accurate, meaningful, generalizable, privacy-protecting and ethically responsible representations of individual differences.

12. Conclusion

This survey uncovered a significant transition from single-modality prediction models and systems to multimodal, context-aware, longitudinal, foundation-model-based, and adaptive personality-aware prediction models and systems in the development of personality trait identification systems in the range of 2022–2026. Technological progress from machine learning to deep neural networks, Transformers, large language models, multimodal foundation models, generative AI, and agentic systems has enabled the analysis of many different types of behavioral data, such as textual, speech, body and facial, physiological, social media, and interaction data. While all of these changes have occurred, model complexity and psychological validity do not necessarily go hand in hand, particularly when considering the different levels of lifelong personality and momentary emotional or environmental conditions. Personality detection is increasingly being applied in various fields, such as recruitment, education, recommender systems, healthcare, conversational AI, human–computer interaction, and immersive environments, as well as in personalized human–AI interaction. However, there are serious concerns about fairness, privacy, transparency, consent and human oversight in these applications. Some of the primary challenges are limited and biased datasets, inconsistencies in annotations, cultural and linguistic bias, weak cross-domain generalization, insufficient longitudinal evidence, multimodal synchronization and missing data, limited interpretability, privacy and security concerns, and inability to replicate. Therefore, multimodal, longitudinal, explainable, privacy-preserving, equitable, and reliable personality-aware AI backed by robust psychometric validation, uncertainty estimation, a variety of datasets, external evaluation, and responsible data practices should be the main emphasis of future research. Overall, the field is shifting from unimodal prediction towards psychologically motivated and reliable human-AI systems. The ultimate vision is to create reliable, transferable, understandable, and morally sound computational models of individual variation, not simply in terms of predictive accuracy. Therefore, the detection of personality traits is emerging as a multi-disciplinary research area that has a close relationship with artificial intelligence, machine learning, psychology, multi-modal computing, human–computer interaction and responsible AI. Future development of this field will rely as much as anything on advances in computational modelling, psychological measurement, data diversity, rigorous validation, privacy protection, and responsible real-world deployment.

References

[1] Yash Mehta et al., "Future-Generation Personality Prediction from Digital Footprints," *Future Generation Computer Systems*, vol. 136, pp. 322-325, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Priyanka Kamble, and Umesh Kulkarni, "An Innovative Approach of Personality Recognition for E-Recruitment," *2022 5th International Conference on Advances in Science and Technology (ICAST)*, Mumbai, India, pp. 324-328, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Büşra Sarikaya, and Ayhan Küçükmanısa, "TB5I: A Multimodal Turkish Interview Dataset and AI-Based Big Five Personality Trait Prediction," *IEEE Access*, vol. 14, pp. 38038-38056, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[4] Constanze Albrecht et al., "Future You: Designing and Evaluating Multimodal AI-generated Digital Twins for Strengthening Future Self-Continuity," *Proceedings of the 31st International Conference on Intelligent User Interfaces*, Paphos Cyprus, pp. 1901-1916, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[5] Heba Ismail, Maryam Alhefeiti, and Ashraf Khalil, "Image and Metadata-Driven Personality Inference for Career Recommendation: A Social Media-based AI Framework for Adolescents," *Discover Artificial Intelligence*, vol. 6, no. 1, pp. 1-25, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[6] Simarpreet Singh, and Williamjeet Singh, "AI-based Personality Prediction for Human Well-being from Text Data: A Systematic Review," *Multimedia Tools and Applications*, vol. 83, pp. 46325-46368, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Asifa Nazir et al., "Alzheimer's Disease Diagnosis using Deep Learning Techniques: Datasets, Challenges, Research Gaps and Future Directions," *International Journal of System Assurance Engineering and Management*, pp. 1-35, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Amani M. Aldosari, Sumayyah E. Sharaf, and Fahd M. Aldosari, "Role of Artificial Intelligence in Analyzing Human behavior and Predicting Personality Traits and Personality Disorders," *Scientific Reports*, vol. 16, pp. 1-16, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[9] Ashish Bagwari et al., "CBIR-DSS: Business Decision Oriented Content-Based Recommendation Model for E-Commerce," *Information*, vol. 13, no. 10, pp. 1-31, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Casey Bennett et al., "Exploring Data-Driven Components of Socially Intelligent AI through Cooperative Game Paradigms," *Multimodal Technologies and Interaction*, vol. 6, no. 2, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Alexander Kathan et al., "A Personalised Approach to Audiovisual Humour Recognition and its Individual-level Fairness," *Proceedings of the 3rd International on Multimodal Sentiment Analysis Workshop and Challenge*, Lisboa Portugal, pp. 29-36, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Mengyu Zhong et al., "Unimodal vs. Multimodal Prediction of Antenatal Depression from Smartphone-based Survey Data in a Longitudinal Study," *Proceedings of the 2022 International Conference on Multimodal Interaction*, Bengaluru India, pp. 455-467, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Patrick Terhürne et al., "Validation and Application of the Non-Verbal Behavior Analyzer: An Automated Tool to Assess Non-verbal Emotional Expressions in Psychotherapy," *Frontiers in Psychiatry*, vol. 13, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Dongdong Liu et al., "Measuring Depression Severity based on Facial Expression and Body Movement using Deep Convolutional Neural Network," *Frontiers in Psychiatry*, vol. 13, pp. 1-13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Andreas Holzinger et al., "Personas for Artificial Intelligence (AI) an Open Source Toolbox," *IEEE Access*, vol. 10, pp. 23732-23747, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Luisa Brinkschulte et al., "Perspectives on Socially Intelligent Conversational Agents," *Multimodal Technologies and Interaction*, vol. 6, no. 8, pp. 1-23, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Carolin Wienrich et al., "Inhibitors and Enablers to Explainable AI Success: A Systematic Examination of Explanation Complexity and Individual Characteristics," *Multimodal Technologies and Interaction*, vol. 6, no. 12, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Dnyaneshwar Madhukar Bavkar, Ramgopal Kashyap, and Vaishali Khairnar, "Multimodal Sarcasm Detection via Hybrid Classifier with Optimistic Logic," *Journal of Telecommunications and Information Technology*, vol. 89, no. 3, pp. 97-114, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Martha T. Teye et al., "Evaluation of Conversational Agents: Understanding Culture, Context and Environment in Emotion Detection," *IEEE Access*, vol. 10, pp. 24976-24984, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Ghazala Bilquise, Samar Ibrahim, and Khaled Shaalan, "Emotionally Intelligent Chatbots: A Systematic Literature Review," *Human Behavior and Emerging Technologies*, vol. 2022, pp. 1-23, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] Roger Azevedo et al., "Lessons Learned and Future Directions of MetaTutor: Leveraging Multichannel Data to Scaffold Self-Regulated Learning with an Intelligent Tutoring System," *Frontiers in Psychology*, vol. 13, pp. 1-23, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] Anjeela Jokhan et al., "Increased Digital Resource Consumption in Higher Educational Institutions and the Artificial Intelligence Role in Informing Decisions Related to Student Performance," *Sustainability*, vol. 14, no. 4, pp. 1-17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Vaishali U. Gongane, Mousami V. Munot, and Alwin D. Anuse, "Detection and Moderation of Detrimental Content on Social Media Platforms: Current Status and Future Directions," *Social Network Analysis and Mining*, vol. 12, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Anna Lena Hunkenschroer, and Christoph Luetge, "Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda," *Journal of Business Ethics*, vol. 178, pp. 977-1007, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Chiara Ullstein et al., "AI-Competent Individuals and Laypeople Tend to Oppose Facial Analysis AI," *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, Arlington VA USA, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] Umut Arioz et al., "Scoping Review on the Multimodal Classification of Depression and Experimental Study on Existing Multimodal Models," *Diagnostics*, vol. 12, no. 11, pp. 1-24, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Paweł Jemioło et al., "Datasets for Automated Affect and Emotion Recognition from Cardiovascular Signals Using Artificial Intelligence— A Systematic Review," *Sensors*, vol. 22, no. 7, pp. 1-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Jing Ling Tay, Ziqiang Li, and Kang Sim, "Effectiveness of Artificial Intelligence Methods in Personalized Aggression Risk Prediction within Inpatient Psychiatric Treatment Settings-A Systematic Review," *Journal of Personalized Medicine*, vol. 12, no. 9, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[29] Lorenzo Picinali et al., "The SONICOM Project: Artificial Intelligence-Driven Immersive Audio, From Personalization to Modeling [Applications Corner]," *IEEE Signal Processing Magazine*, vol. 39, no. 6, pp. 85-88, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[30] Sue E. Kase et al., "The Future of Collaborative Human-Artificial Intelligence Decision-Making for Mission Planning," *Frontiers in Psychology*, vol. 13, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] Wen-Jing Yan, Qian-Nan Ruan, and Ke Jiang, "Challenges for Artificial Intelligence in Recognizing Mental Disorders," *Diagnostics*, vol. 13, no. 1, pp. 1-10, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[32] Abdul Rehman Javed et al., "Artificial Intelligence for Cognitive Health Assessment: State-of-the-Art, Open Challenges and Future Directions," *Cognitive Computation*, vol. 15, pp. 1767-1812, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[33] Megan Wiedbusch et al., "Measuring Multidimensional Facets of SRL Engagement with Multimodal Data," *Unobtrusive Observations of Learning in Digital Environments: Examining Behavior, Cognition, Emotion, Metacognition and Social Processes Using Learning Analytics*, pp. 141-173, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[34] Summaira Jabeen et al., "A Review on Methods and Applications in Multimodal Deep Learning," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 19, no. 2s, pp. 1-41, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[35] Priya Bhatt et al., "Machine Learning for Cognitive Behavioral Analysis: Datasets, Methods, Paradigms, and Research Directions," *Brain Informatics*, vol. 10, pp. 1-37, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[36] Alaa Abd-alrazaq et al., "Large Language Models in Medical Education: Opportunities, Challenges, and Future Directions," *JMIR Medical Education*, vol. 9, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[37] Shahab Saquib Sohail et al., "The Future of GPT: A Taxonomy of Existing ChatGPT Research, Current Challenges, and Possible Future Directions," *SSRN*, pp. 1-22, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[38] V. Kowsalya, and C. Divya, "An Analogous Review of the Challenges and Endeavor in Suspense Story Generation Technique," *Advances in Data Science and Artificial Intelligence*, IIT Patna, India, pp. 99-111, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[39] Smita Khairnar et al., "Face Liveness Detection Using Artificial Intelligence Techniques: A Systematic Literature Review and Future Directions," *Big Data and Cognitive Computing*, vol. 7, no. 1, pp. 1-35, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[40] Yonghao Xu et al., "AI Security for Geoscience and Remote Sensing: Challenges and Future Trends," *IEEE Geoscience and Remote Sensing Magazine*, vol. 11, no. 2, pp. 60-85, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[41] Sina Ghassemi et al., "Unsupervised Multimodal Learning for Dependency-Free Personality Recognition," *IEEE Transactions on Affective Computing*, vol. 15, no. 3, pp. 1053-1066, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[42] P. Sreevidya, J. Aravinth, and Sathishkumar Samiappan, "Intermediate Layer Attention Mechanism for Multimodal Fusion in Personality and Affect Computing," *IEEE Access*, vol. 12, pp. 112776-112793, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[43] Xianxun Zhu et al., "A Review of Key Technologies for Emotion Analysis Using Multimodal Information," *Cognitive Computation*, vol. 16, no. 4, pp. 1504-1530, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[44] Na Chen et al., "Multimodal Heterogeneous Data Sensing and Communication Integration for CIoT," *IEEE Transactions on Consumer Electronics*, vol. 71, no. 3, pp. 7454-7472, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[45] Sepideh Kalateh et al., "A Systematic Review on Multimodal Emotion Recognition: Building Blocks, Current State, Applications, and Challenges," *IEEE Access*, vol. 12, pp. 103976-104019, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[46] Fei Gao et al., "Intelligent Cockpits for Connected Vehicles: Taxonomy, Architecture, Interaction Technologies, and Future Directions," *Sensors*, vol. 24, no. 16, pp. 1-35, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[47] Anish Bhandarkar et al., "Deep Learning based Computer Aided Diagnosis of Alzheimer's Disease: A Snapshot of Last 5 Years, Gaps, and Future Directions," *Artificial Intelligence Review*, vol. 57, no. 2, pp. 1-62, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[48] Vincenzo Dentamaro et al., "Enhancing Early Parkinson's Disease Detection through Multimodal Deep Learning and Explainable AI: Insights from the PPMI Database," *Scientific Reports*, vol. 14, pp. 1-18, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[49] Alexander Henlein et al., "An Outlook for AI Innovation in Multimodal Communication Research," *15th International Conference Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management*, Washington, DC, USA, pp. 182-234, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[50] Gokul Yenduri et al., "GPT (Generative Pre-Trained Transformer)— A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions," *IEEE Access*, vol. 12, pp. 54608-54649, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[51] Zijie Ding et al., "Current Status and Trends of Technology, Methods, and Applications of Human–Computer Intelligent Interaction (HCII): A Bibliometric Research," *Multimedia Tools and Applications*, vol. 83, no. 27, pp. 69111-69144, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[52] Yao Rong et al., "Towards Human-Centered Explainable AI: A Survey of User Studies for Model Explanations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 4, pp. 2104-2122, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[53] Sachin Kumar et al., "Multimodality Fusion Aspects of Medical Diagnosis: A Comprehensive Review," *Bioengineering*, vol. 11, no. 12, pp. 1-32, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[54] Phan The Duy et al., "A Study on Adversarial Sample Resistance and Defense Mechanism for Multimodal Learning-Based Phishing Website Detection," *IEEE Access*, vol. 12, pp. 137805-137824, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[55] Anam Naz et al., "Machine and Deep Learning for Personality Traits Detection: A Comprehensive Survey and Open Research Challenges," *Artificial Intelligence Review*, vol. 58, no. 8, pp. 1-57, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[56] Yuntao Hong, and Zongzheng Xia, "AI-Driven Innovations in Psychological Assessment: Multimodal Data, Intelligent Analytics, and Ethical Challenges," *Proceedings of the 2025 International Conference on Artificial Intelligence and Smart Manufacturing*, Chongqing China, pp. 854-859, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[57] Siham Essahraui et al., "Human Behavior Analysis: A Comprehensive Survey on Techniques, Applications, Challenges, and Future Directions," *IEEE Access*, vol. 13, pp. 128379-128419, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[58] Yash Raj Shrestha, and Vivianna Fang He, "Integrating Multimodal Data and Machine Learning for Entrepreneurship Research," *Strategic Entrepreneurship Journal*, pp. 1-38, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[59] Sadam Al-Azani, and El-Sayed M. El-Alfy, "A Review and Critical Analysis of Multimodal Datasets for Emotional AI," *Artificial Intelligence Review*, vol. 58, no. 10, pp. 1-65, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[60] Kuei-Chung Chang, and Sheng-Quan Chen, "A Survey of Multimodal Emotion Recognition: Fusion Techniques, Datasets, Challenges and Future Directions," *International Journal of Biometrics*, vol. 17, no. 5, pp. 485-510, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[61] Fan Li et al., "A New Era in Human Factors Engineering: A Survey of the Applications and Prospects of Large Multimodal Models," *International Journal of Human–Computer Interaction*, vol. 41, no. 18, pp. 11867-11880, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[62] Yujia Zhao et al., "AI-Enabled Text-to-Music Generation: A Comprehensive Review of Methods, Frameworks, and Future Directions," *Electronics*, vol. 14, no. 6, pp. 1-53, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[63] Abdulmotaleb El Saddik et al., "Unleashing Creativity in the Metaverse: Generative AI and Multimodal Content," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 21, no. 7, pp. 1-43, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[64] Yu-Zhu Sun et al., "Multimodal Agent AI: A Survey of Recent Advances and Future Directions," *Journal of Computer Science and Technology*, vol. 40, no. 4, pp. 1046-1063, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[65] Mayk Caldas Ramos, Christopher J. Collison, and Andrew D. White, "A Review of Large Language Models and Autonomous Agents in Chemistry," *Chemical Science*, vol. 16, no. 6, pp. 2514-2572, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[66] Sobhana Jahan et al., "Federated Explainable AI-Based Alzheimer's Disease Prediction with Multimodal Data," *IEEE Access*, vol. 13, pp. 43435-43454, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[67] Ruoyu Wan et al., "Current Status and Future Directions of Artificial Intelligence in Post-Traumatic Stress Disorder: A Literature Measurement Analysis," *Behavioral Sciences*, vol. 15, no. 1, pp. 1-32, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[68] Yutaro Fuse et al., "Artificial Intelligence in Clinical Data Analysis: A Review of Large Language Models, Foundation Models, Digital Twins, and Allergy Applications," *Allergology International*, vol. 74, no. 4, pp. 499-513, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[69] Rabiah Aslam Ansari et al., "Artificial Intelligence-Guided Neuromodulation in Heart Failure with Preserved and Reduced Ejection Fraction: Mechanisms, Evidence, and Future Directions," *Journal of Cardiovascular Development and Disease*, vol. 12, no. 8, pp. 1-28, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[70] Susanne de Mooij et al., "A Systematic Review of Self-Regulated Learning through Integration of Multimodal Data and Artificial Intelligence," *Educational Psychology Review*, vol. 37, no. 2, pp. 1-27, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[71] Emi Kalita et al., "Educational Data Mining: A 10-year Review," *Discover Computing*, vol. 28, no. 1, pp. 1-25, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[72] Yijie Zheng et al., "Research and Implementation of a Multi-Modal AI-Driven Psychological Assessment and Analysis System Based on Unity," *Proceedings of the 2025 2nd International Conference on Image Processing, Intelligent Control and Computer Engineering*, Zhengzhou China, pp. 580-587, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[73] Süeda Özkaya, Santiago Berrezueta-Guzman, and Stefan Wagner, "How LLMs Are Shaping the Future of Virtual Reality," *IEEE Access*, vol. 13, pp. 193335-193355, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[74] Fatema Rahimi, Abolghasem Sadeghi-Niaraki, and Soo-Mi Choi, "Generative AI Meets Virtual Reality: A Comprehensive Survey on Applications, Challenges, and Future Direction," *IEEE Access*, vol. 13, pp. 94893-94909, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[75] Elias Dritsas, and Maria Trigka, "Machine Learning and Data Science in Social Sciences: Methods, Applications, and Future Directions," *IEEE Access*, vol. 13, pp. 105334-105352, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[76] Michaela Benk et al., "Twenty-four Years of Empirical Research on Trust in AI: A Bibliometric Review of Trends, Overlooked Issues, and Future Directions," *AI & Society*, vol. 40, no. 4, pp. 2083-2106, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[77] Algirdas Laukaitis et al., "FAIR-VID: A Multimodal Pre-Processing Pipeline for Student Application Analysis," *Applied Sciences*, vol. 15, no. 24, pp. 1-21, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[78] Pejman Peykani et al., "Large Language Models: A Structured Taxonomy and Review of Challenges, Limitations, Solutions, and Future Directions," *Applied Sciences*, vol. 15, no. 14, pp. 1-52, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[79] Guodong Chen et al., "The Study of Human-AI Co-creation Design under Generative Artificial Intelligence: Cognition, Process, Method, and Outcome," *Journal of Engineering Design*, vol. 37, no. 4, pp. 1710-1751, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[80] Abdulaziz Alharbi et al., "Artificial Intelligence-Enabled Conversational Agents in Tourism & Hospitality: A Systematic Literature Review & Future Research Directions," *Asia Pacific Journal of Tourism Research*, vol. 31, no. 1, pp. 21-52, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[81] J. Chandra et al., *11 - Applications and Future Directions in Multimodal Large Language Model: Opportunities and Challenges*, Challenges and Applications of Generative Large Language Models, pp. 219-241, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[82] Eunseon Jo, and Han-jin Lee, "Intelligent Personality-Aware AR Gait Training Using Smart Glasses: Personalized Multimodal Feedback for Next-Generation Digital Rehabilitation," *21st International Conference on Multimedia Information Technology and Applications*, Jeju, South Korea, pp. 153-164, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[83] Sireesha Chittepu, Sheshikala Martha, and Debajyoty Banik, "Comprehensive Survey on Multimodal Emotion Detection with Advances, Challenges, and Future Directions in Pairwise Conversations," *Cognitive Processing*, vol. 27, pp. 275-290, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[84] Ghulam Muhammad et al., "EEG-based Multimodal Emotion Recognition: Recent Progress, Challenges, and Future Directions," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 22, no. 2, pp. 1-28, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[85] Rafał Gasz et al., "A Review of Current Capabilities and Future Directions in Machine-based Emotion Recognition," *Cognitive, Affective, & Behavioral Neuroscience*, vol. 26, pp. 1393-1419, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[86] Xu Tian et al., "Interaction of Artificial Intelligence, Mental Disorders, and Diverse Data Modalities: Potential Treatment Management based on the "method–disease–data" Axis" *Neural Regeneration Research*, vol. 21, no. 9, pp. 3919-3932, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[87] Heysem Kaya, and Gizem Sogancioglu, "Towards Responsible Multimodal Modeling for Mental Healthcare," *27th International Conference Speech and Computer*, Szeged, Hungary, pp. 3-22, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[88] Dhanashree Vipul Yevle, Palvinder Singh Mann, and Dinesh Kumar, "AI Based Diagnosis of Chronic Obstructive Pulmonary Disease: Acomparative Review," *Archives of Computational Methods in Engineering*, vol. 33, pp. 2413-2450, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[89] Sandeep Kaur, Usha Mittal, and Ankita Wadhawan, "A Systematic Review of Deep Learning Approaches for Brain Tumor Segmentation in MRI: Trends, Challenges and Future Directions," *Archives of Computational Methods in Engineering*, vol. 33, no. 2, pp. 3035-3058, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[90] Partha Pratim Ray et al., "A Review of TRiSM Frameworks in Artificial Intelligence Systems: Fundamentals, Taxonomy, Use Cases, Key Challenges and Future Directions," *Expert Systems*, vol. 43, no. 3, pp. 1-53, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[91] Nathan Foster, "A Comprehensive Survey on Artificial Intelligence Applications Across Critical Domains: Healthcare, Finance, Cybersecurity, and Beyond," *Journal of Artificial Intelligence and Machine Learning Review*, vol. 7, no. 1, pp. 120-135, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[92] Lubnaa Abdur Rahman et al., "Generative AI in Precision Nutrition: A Review of Current Developments and Future Directions," *Nutrients*, vol. 18, no. 6, pp. 1-31, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[93] Avijit Paul et al., "Empowering Photodynamic Therapy with Artificial Intelligence: Current Trends and Future Directions," *Frontiers in Oncology*, vol. 16, pp. 1-33, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[94] Athanasios Zikopoulos et al., "Artificial Intelligence in Recurrent Pregnancy Loss: Current Evidence, Limitations, and Future Directions," *Journal of Clinical Medicine*, vol. 15, no. 2, pp. 1-20, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[95] Arfan Ahmed et al., "Machine Learning Models to Detect Anxiety and Depression through Social Media: A Scoping Review," *Computer Methods and Programs in Biomedicine Update*, vol. 2, pp. 1-9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[96] Mohsen Ghorbian, Saeid Ghorbian, and Mostafa Ghobaei-Arani, "Advancements in Machine Learning for Brain Tumor Classification and Diagnosis: A Comprehensive Review of Challenges and Future Directions," *Archives of Computational Methods in Engineering*, vol. 33, no. 1, pp. 1373-1408, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[97] M. S. Remya et al., "Artificial Intelligence for Speech Classification and Enhancement of Speech and Language Disorders: Techniques, Applications, and Future Directions," *IEEE Access*, vol. 13, pp. 177136-177159, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[98] Ziqing Zhang, "AI-Powered Intelligent Speech Processing: Evolution, Applications and Future Directions," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 2, pp. 918-928, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[99] Weiqi Xu, and Fan Ouyang, "A Systematic Review of AI Role in the Educational System based on a Proposed Conceptual Framework," *Education and Information Technologies*, vol. 27, pp. 4195-4223, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[100] Vijaya Kumar Sagenela et al., "Artificial Intelligence in Diabetic Retinopathy Detection: A Decade of Progress from Machine Learning to Transformer-Based Frameworks," *Archives of Computational Methods in Engineering*, vol. 33, pp. 7693-7718, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[101] Evdokia Voultsiou, and Lefteris Moussiades, "A Systematic Review of AI, VR, and LLM Applications in Special Education: Opportunities, Challenges, and Future Directions," *Education and Information Technologies*, vol. 30, pp. 19141-19181, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[102] Shamim Yazdani et al., "Generative AI in Depth: A Survey of Recent Advances, Model Variants, and Real-World Applications," *Journal of Big Data*, vol. 12, no. 1, pp. 1-43, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[103] Qin Zhao, and Shah Nazir, "English Multimode Production and Usage by Artificial Intelligence and Online Reading for Sustaining Effectiveness," *Mobile Information Systems*, vol. 2022, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[104] Richa Tripathi, and Mohd Rashid Alam, *Multimodal AI in Suicide Risk Prediction: Behavioral, Physiological, and Digital Biomarkers*, Multimodal AI in Healthcare, Springer, pp. 415-446, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[105] Afnan A. Shah, "Enhancing Hajj and Umrah Rituals and Crowd Management Through AI Technologies: A Comprehensive Survey of Applications and Future Directions," *IEEE Access*, vol. 12, pp. 161820-161841, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[106] Dessy Harisanty et al., "Cultural Heritage Preservation in the Digital Age, Harnessing Artificial Intelligence for the Future: A Bibliometric Analysis," *Digital Library Perspectives*, vol. 40, no. 4, pp. 609-630, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[107] Nuzhat Khan et al., "Systematic Literature Review of Machine Learning Models and Applications for Text Recognition," *IEEE Access*, vol. 13, pp. 177647-177670, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[108] Liang Gao et al., "AI-Assisted Design of Advanced Polymeric Materials: Challenges and Solutions," *Advanced Materials*, vol. 38, no. 7, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[109] Manuel Paniagua-Gómez, and Manuel Fernandez-Carmona, "Trends and Challenges in Real-Time Stress Detection and Modulation: The Role of the IoT and Artificial Intelligence," *Electronics*, vol. 14, no. 13, pp. 1-31, 2025. [CrossRef] [Google Scholar] [Publisher Link]

Appendix

Table 2. Gives a summary of the primary data sources

Data source	Typical data	Personality information	Main advantage	Main limitation	Representative studies
Text/linguistic	Posts, transcripts, essays, conversations	Linguistic style, semantics	Rich semantic information; scalable	Language/domain bias	[1, 3, 54,78]
Social media	Posts, comments, images, metadata	Self-expression, interests, interaction	Naturalistic behavioral traces	Privacy, platform/context bias	[1, 5, 54]
Speech/audio	Interviews, voice, prosody	Vocal style, communication cues	Complements text	Accent, speaker and state effects	[3, 10, 11, 31, 92]
Facial/visual	Images, expressions, gaze, appearance	Non-verbal behaviors	Rich visual information	Privacy, demographic bias	[3, 5, 21, 41, 42]
Video/behavioral	Gestures, posture, gait, gameplay	Temporal behaviors	Captures dynamics	Computational/temporal complexity	[10, 11, 23, 33, 83]
Physiological	EEG, ECG, EMG, PPG, HRV	Affective/internal responses	Complementary bio signals	Noise, context/state effects	[3, 12, 31, 45, 85, 86]
Psychological/clinical	Questionnaires, self-reports, clinical measures	Trait scores/reference measures	Psychologically grounded	Self-report bias, privacy	[1, 23, 32, 47, 48, 57]
Longitudinal/mobile	Smartphones, wearables, repeated observations	Stability and variation	Temporal/ecological modelling	Missing data, attrition	[1, 12, 43]
XR/immersive	VR/AR, smart glasses, 3D interaction	Contextual behaviors	Interactive behavioral evidence	Limited datasets, transferability	[73, 61, 101]
Multimodal	Text + audio + video + physiology	Integrated representation	Complementary information	Alignment, missing modalities, scale	[3, 10, 23, 34, 41, 43, 60]

Table 3. Machine-Learning and Deep-Learning methodologies for personality detection

Method	Representative techniques	Main role	Strengths	Key limitations	Representative references
Traditional ML	SVM, RF, Logistic Regression, Decision Tree	Classification/regression from engineered features	Efficient; suitable for small datasets; interpretable baselines	Feature engineering; limited representation learning	[2, 3, 56]
Regularized Regression	Ridge, Lasso, ElasticNet	Continuous personality-trait prediction	Suitable for trait scores; relatively interpretable	Limited nonlinear modelling	[3]
Deep Learning	CNN, RNN, LSTM, BiLSTM	Visual, acoustic, textual and temporal representation learning	Automatic feature learning; complex pattern modelling	Data/computation requirements; interpretability	[41, 42, 56, 85]
Hybrid/Attention Models	CNN-LSTM, multimodal and dual-attention networks	Joint spatial, temporal and multimodal modelling	Combines complementary representations	Architectural complexity	[41, 42]
Transformers	BERT and related models	Contextual language and multimodal representation	Transfer learning; long-range contextual modelling	Computation; bias; interpretability	[3, 56]
Vision	ViT and related models	Visual/facial/video	Global visual modelling	Data and	[90, 102]

Transformers		representation		computational demands	
LLMs	GPT-family and related models	Conversational and semantic personality inference	Few-shot learning; broad contextual capability	Hallucination; bias; privacy; psychometric validity	[36, 37, 103]
Multimodal LLMs	Vision-language and multimodal models	Joint reasoning across heterogeneous modalities	Unified multimodal representation	Alignment, modality bias, and high computational cost	[65, 82]
Generative AI	GAN, VAE, diffusion, generative models	Synthetic data and personality-conditioned generation	Addresses data scarcity; enables simulation	Synthetic-data validity; bias; controllability	[104, 105]
Agentic AI	Multimodal/autonomous LLM agents	Personality-aware planning and adaptation	Dynamic personalization and interaction	Safety; reliability; governance	[65, 66]

Table 4. Multimodal fusion strategies for personality trait detection

Fusion strategy	Modalities	Main contribution	Strengths	Key limitations	Representative references
Text–Audio	Text + speech	Combines linguistic and vocal cues	Captures content and delivery	State/context effects	[10, 31, 92]
Text–Visual	Text + images/facial behaviors	Combines linguistic expression and visual self-presentation	Complementary semantic and visual information	Privacy; self-presentation bias	[5, 6]
Audio–Visual	Speech + facial/video	Models' verbal delivery and non-verbal behaviors	Rich interpersonal representation	Synchronization; missing data	[10, 11, 41]
Text–Audio–Visual	Text + speech + video	Joint communication representation	Integrates verbal, vocal, and visual evidence	Dimensionality; alignment	[3, 10, 41, 42]
Physiological–Behavioral	Bio signals + behavioral data	Combines internal responses and observable behaviors	Richer individual-state information	Noise; context; trait–state ambiguity	[12, 45, 85, 86]
Early Fusion	Feature-level multimodal inputs	Joint learning from combined features	Simple, direct interaction modelling	Dimensionality; modality imbalance	[34]
Late Fusion	Modality-specific predictions	Decision-level integration	Modular and flexible	Cross-modal interactions may be lost	[11, 43]
Hybrid/Intermediate	Multiple fusion levels	Combines modality-specific and shared representations	Flexible and potentially robust	Architectural complexity	[41-43]
Attention-Based Fusion	Multiple modalities	Dynamically weights informative modalities/segments	Adaptive cross-modal integration	Complexity; interpretation challenges	[42, 84]
Multimodal Transformers	Text, audio, visual and other signals	Learns cross-modal contextual relationships	Strong contextual and alignment capability	Data/computation; alignment	[43, 50, 82]
Multimodal LLMs/LMMs	Text, image, audio, video	Unified multimodal reasoning	Flexible foundation-model paradigm	Hallucination; bias; privacy; validity	[65, 82, 105]
Multimodal Agents	Language, vision, audio, behavioral data	Personality-aware planning and adaptation	Dynamic personalization	Safety; reliability; governance	[65, 66]

Table 5. Major challenges and limitations in personality trait detection

Challenge	Core problem	Effect on detection	Required research direction	Representative references
Dataset scarcity	Few large participant-level multimodal datasets with validated personality labels	Overfitting and poor generalization	Large, diverse, longitudinal datasets	[43, 56, 81]
Annotation bias	Self-report, expert and inventory differences	Weak or inconsistent ground truth	Standardized, psychometrically validated annotation	[1, 79]
Cultural/language bias	Unequal linguistic and cultural representation	Poor cross-population generalization	Multilingual and cross-cultural datasets	[1, 41, 103, 109]
Cross-domain shift	Changes in platforms, environments and populations	Performance degradation outside training conditions	Cross-dataset and cross-domain benchmarks	[41, 58, 43]
Multimodal alignment	Different temporal resolutions and asynchronous signals	Spurious cross-modal associations	Temporal alignment and cross-modal attention	[43, 81]
Missing/noisy modalities	Incomplete or corrupted observations	Reduced real-world reliability	Robust and uncertainty-aware fusion	[43, 81, 83]
Interpretability	Complex models provide limited insight into predictions	Difficult psychological validation	Psychologically grounded XAI	[53, 79, 110]
Privacy/security	Sensitive traits inferred from behavioral data	Unauthorized profiling and data leakage	Federated/privacy-preserving learning and governance	[1, 91, 110]
Fairness/ethics	Demographic bias and inappropriate profiling	Discrimination and loss of autonomy	Subgroup evaluation, consent and human oversight	[91, 111]
Trait–state ambiguity	Behavior reflects both stable traits and temporary states	Models may detect state rather than personality	Trait–state–context modelling	[1, 12, 33]
Longitudinal limitations	A few repeated observations over time	Limited evidence of stability and drift	Longitudinal and continual evaluation	[12, 83]
Reproducibility	Inconsistent datasets, protocols and metrics	Difficult cross-study comparison	Standardized benchmarks and reporting	[43, 56, 81]