Review

Not peer-reviewed version

Large Language Models in the Assessment and Care of Internet Gaming Disorder

Athanasios Kranas * and Vassilios Verykios

Posted Date: 17 December 2025

doi: 10.20944/preprints202512.1552.v1

Keywords: Internet Gaming Disorder (IGD); Large Language Models (LLMs); digital mental health; assessment and treatment; ethics and data privacy

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2025 doi:10.20944/preprints202512.1552.v1

Review

Large Language Models in the Assessment and Care of Internet Gaming Disorder

Athanasios Kranas * and Vassilios S. Verykios

School of Science and Technology, Hellenic Open University, Patras 26335, Greece

* Correspondence: athanasios.kranas@ac.eap.gr; Tel.: +302610367698

Abstract

Internet Gaming Disorder (IGD), recognized in International Classification of Diseases (ICD-11), affects millions—especially adolescents and young adults—and poses challenges that invite scalable innovations in care. This narrative review examined how Large Language Models (LLMs) could support IGD prevention, assessment, treatment, and research. We conducted targeted searches of PubMed, Scopus, Google Scholar, and IEEE Xplore for 2010–October 2025, supplemented by backward/forward citation chasing; English, peer-reviewed clinical, methodological, and review work was prioritized; as a narrative review, we did not apply PRISMA or perform quantitative synthesis; in total, we synthesized over 50 sources. We synthesize peer-reviewed, IGD-specific AI/ML studies with explicit reporting of training approach, validation/performance, dataset size, and model openness. Preliminary improvements observed in adjacent digital-health trials highlight promise yet underscore the need for rigorous, IGD-specific validation; to date, IGD-specific randomized trials remain scarce. Evidence spans transformer-embedding text screening with supervised regression ($r \approx 0.48$), multimodal EEG + neuropsychology classification (≈71% vs. comparison groups), fNIRS deep learning (≈88% vs. healthy), and fMRI-based connectomics/MVPA, with sample sizes $n = 40$–417 and most implementations being research-only (no public code/data). Principal concerns include privacy and data governance, algorithmic bias, inconsistent crisis-escalation performance, and a nascent clinical evidence base. We conclude that LLMs may augment—but should not replace—human clinicians; near-term promise lies in hybrid human-AI pathways, multimodal integrations with wearables and gaming APIs, and rigorous prospective trials to establish safety, effectiveness, and equity in IGD care.

Keywords: Internet Gaming Disorder (IGD); Large Language Models (LLMs); digital mental health; assessment and treatment; ethics and data privacy

1. Introduction

The digital age has ushered in unprecedented connectivity and entertainment, but it has also led to new forms of psychological distress. Among these, Internet Gaming Disorder (IGD) stands out as a behavioral addiction that captivates individuals in virtual worlds, often at the expense of real-life responsibilities and well-being. Defined in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) as a condition warranting further study, and formally included in the ICD-11, IGD involves a pattern of persistent gaming behavior that leads to significant impairment in personal, social, or occupational functioning [1]. Prevalence estimates vary widely by measure and population; meta-analytic studies converge around approximately 3% global prevalence [2].

For example, a young adult might begin gaming as a stress reliever after a long day, only to find hours slipping away into nights, relationships fraying, and academic performance plummeting. This scenario is unfortunately common, reflecting the insidious pull of games designed with psychological hooks such as variable reward schedules—akin to slot machines. Yet, traditional interventions, such as cognitive behavioral therapy (CBT) or pharmacological aids, face barriers: stigma, limited access to specialists, and high dropout rates [3]. This is where artificial intelligence (AI)—specifically Large

Language Models (LLMs)—offers new possibilities, which promise to bridge these gaps through innovative, scalable solutions.

LLMs, exemplified by models like OpenAI's GPT series or Google's PaLM, are neural networks trained on vast datasets to understand and generate human language with fluency that often rivals experts [4]. In psychiatry and psychology, they have shown prowess in tasks ranging from sentiment analysis of social media posts to simulating empathetic conversations [5]. The integration of LLMs with IGD represents a nascent but fertile ground: imagine an AI chatbot that detects early signs of addiction through casual dialogue, or a virtual therapist tailoring CBT exercises to a gamer's specific triggers. This review aims to map this intersection, evaluating how LLMs can enhance our understanding and management of IGD while navigating the ethical minefield.

Why now? The convergence of rising IGD cases—exacerbated by the COVID-19 pandemic's shift to online activities—and rapid AI advancements demands timely exploration [6]. Recent studies indicate that AI tools, including machine learning algorithms, have already been piloted for addiction monitoring, such as tracking spending patterns in gambling disorders [7]. Extending this to LLMs for IGD could democratize mental health support, especially in underserved areas. However, enthusiasm must be tempered with caution; unchecked AI deployment risks amplifying biases or providing misguided advice [8].

This article is a narrative review, not a systematic review. We conducted a targeted literature search across PubMed, Scopus, Google Scholar, and IEEE Xplore for publications from 2010 through October 2025, supplemented by backward/forward citation chasing (snowballing) of key papers. Search terms included: “Internet Gaming Disorder,” “gaming disorder,” “Large Language Models,” “LLM,” “AI in mental health,” “digital addiction,” “chatbot,” “cognitive behavioral therapy/CBT,” “motivational interviewing/MI,” “screening,” “monitoring,” and “relapse prevention.” We prioritized English-language, peer-reviewed clinical, methodological, and review articles and interdisciplinary work bridging mental health and LLM technologies; a limited number of non–peer-reviewed sources (e.g., regulatory or ethical frameworks) were included only when essential for policy/ethics context. In line with the narrative format, we did not apply PRISMA procedures or perform quantitative synthesis or formal risk-of-bias assessment. In total, we synthesized over 50 relevant sources to support the themes of this review. To maximize practical utility, we further compiled IGD-specific AI/ML use cases into a structured evidence summary that reports, for each study, the method of training/fine-tuning, validation and performance, dataset size/characteristics, and model openness (see Table 1 and Appendix A).

Table 1. Representative AI/LLM-based approaches for IGD detection. LLM = large language model; IGD = Internet Gaming Disorder; SVM = support vector machine; CNN = convolutional neural network; fNIRS = functional near-infrared spectroscopy; MVPA = multi-voxel pattern analysis.

Study (Year)	AI Method & Training	Sample & Data	Performance Metrics	Model Open-Source?
Strojny et al., 2024	NLP (Transformer + Ridge Regression) – Supervised learning on open-ended text. Pretrained HerBERT transformer encoder for embeddings (no fine-tuning).	417 gamers (online survey) – 4 free-text responses + IGD scale.	Correlation: Predicted vs. actual IGD scores r = 0.476 (multi-response model).	Yes (Partially) – Used open pretrained transformer (HerBERT); model code for research only (not a deployed tool).
Cho et al., 2024	Digital phenotyping (Regression) – Multivariate linear models with cross-validation. No ML fine-tuning (focused on interpretability).	168 students (53% female, age 13–14) – Tablet sensor metrics during classes.	Variance explained: 23% of IGD score by 5 digital markers; significant group differences (avg. Cohen's d ≈ 0.4).	No – Developed as part of proprietary school system (Dr. Simon).
Lee et al., 2024	Multimodal ML (EEG + Surveys) – Supervised classifiers (Logistic Reg., SVM, Random Forest) on neural and clinical features. Models trained with L1 regularization (feature selection).	191 adults (67 IGD, 58 AUD, 66 healthy) – Resting EEG signals + psychometric data.	Accuracy: 71.2% for IGD vs. AUD classification (best model); salient features: abnormal beta/delta EEG connectivity in IGD.	No – Research-only model (open-access publication, but no released software).
Song et al., 2021	Connectome SVM (fMRI) – Modified CPM algorithm using Support Vector Machine. Supervised learning on brain network matrices.	113 young adults (72 IGD, 41 healthy) – Resting-state fMRI connectivity data.	Accuracy: 78.76% (balanced accuracy = 75.46%; sensitivity = 63.41%; specificity = 87.5%); DMN features also predicted CIAS severity (r = 0.44).	No – Not open-source (analysis code not public; no deployment).
Wang et al., 2023	Deep Neural Network (2D-CNN) – Supervised convolutional network trained on fNIRS time-series images. Compared against other ML/DL algorithms.	40 subjects (24 IGD, 16 healthy) – Prefrontal fNIRS signals during stop-signal inhibition task.	Accuracy: 87.5% (IGD vs. healthy, with 2D-CNN – highest among models); outperformed traditional ML classifiers.	No – Custom experimental model (not released; closed-source prototype).
Ye et al., 2022	MVPA (Multi-voxel Pattern Analysis) – Supervised pattern recognition on fMRI features. Trained to predict continuous IGD severity (regression).	402 individuals (mixed gaming habits; spectrum of IGD severity) – Resting fMRI metrics (ReHo, ALFF).	Outcome: Brain patterns significantly predicted IGD severity (p < 0.001); identified key regions (prefrontal cortex) correlated with symptom level.	No – Research analysis only (no deployed model; code not publicly provided).

This paper is structured as follows: Section 2 provides a comprehensive background on IGD, including its diagnostic criteria, risk factors, and current treatment landscape. Section 3 introduces LLMs, their mechanisms, and established applications in mental health. Section 4 delves into specific integrations for IGD, from diagnostic tools to therapeutic interventions, supported by emerging evidence. Section 5 addresses challenges, including ethical, technical, and clinical hurdles. Section 6 outlines future research directions, and Section 7 concludes with key takeaways.

Through this lens, we not only review existing literature but also propose a framework for responsible integration, ensuring that technology serves humanity's mental health needs. While LLMs are not a standalone cure, their thoughtful application could represent a significant advance in combating IGD.

2. Background on Internet Gaming Disorder

2.1. Definition and Diagnostic Criteria

Internet Gaming Disorder emerges from the broader category of behavioral addictions, where non-substance-related activities hijack the brain's reward system. The American Psychiatric Association (APA) in DSM-5 describes IGD as involving at least five of nine criteria over a 12-month period: preoccupation with gaming, withdrawal symptoms when gaming is restricted, tolerance (needing more time to achieve satisfaction), unsuccessful attempts to control gaming, loss of interest in other activities, continued gaming despite psychosocial problems, deception about gaming extent, using gaming to escape negative moods, and jeopardized relationships or opportunities due to gaming [1].

The World Health Organization's ICD-11 refines this by emphasizing impaired control, increasing priority given to gaming, and continuation despite negative consequences [9]. These criteria distinguish IGD from recreational gaming, which can be beneficial for cognitive skills and social bonding. Notably, IGD often co-occurs with other disorders like depression (up to 75% comorbidity) and anxiety, complicating diagnosis [10].

2.2. Epidemiology and Prevalence

Prevalence varies by method and population; meta-analytic estimates place prevalence around ~3% globally; national data from Korea report 0.8% IGD (12-month) and 8.4% problematic use (1-month) [11]. Meta-analytic work underscores that not all heavy gamers meet IGD criteria [2]. Professional guidance during COVID-19 highlighted increased risks and reported increases in gaming, though magnitudes vary by age and context [6]. Earlier studies have highlighted wide variability in reported prevalence owing to differences in instruments, cut-offs, and populations sampled [13]; more recent meta-analytic work converges near ~3% globally [2].

Risk factors include genetic predispositions (e.g., dopamine receptor variations), environmental influences like peer rejection, and psychological traits such as low grit or high impulsivity [14]. Cross-cultural validation of the IGDT-10 (Ten-Item Internet Gaming Disorder Test) indicates a robust symptom structure across countries [15].

2.3. Etiological Models

IGD's development is multifaceted. The cognitive-behavioral model posits that maladaptive beliefs—e.g., "Gaming is my only escape"—perpetuate addiction [16]. Neurobiologically, fMRI studies show altered reward processing in the ventral striatum, akin to substance use disorders [17]. Social factors, like online communities reinforcing excessive play, add layers [18]. A recent neuroimaging meta-analysis integrating VBM and fMRI studies reports convergent alterations across reward and cognitive-control circuits (ventral striatum, anterior cingulate) in IGD, reinforcing these mechanisms [17].

From an environmental perspective, game design elements—microtransactions, loot boxes—exploit psychological vulnerabilities, drawing parallels to gambling [19].

2.4. Current Treatments and Limitations

Interventions span pharmacological, psychological, and technological approaches. CBT remains the first-line treatment; recent systematic reviews report small-to-moderate benefits [3]. Mindfulness-based therapies improve anxiety and depressive symptoms and support self-regulation [20]. Pharmacologically, bupropion has shown promise in a randomized trial [21].

However, access is limited and help-seeking remains low, with stigma cited as a barrier [22]. Dropout is substantial and relapse remains a concern [23]. Emerging tech interventions, like virtual reality exposure, offer novelty but lack scalability [24].

These challenges highlight a role for LLMs in providing always-available, personalized support to help fill these gaps.

3. Background on Large Language Models

3.1. What Are Large Language Models?

Large Language Models (LLMs) are advanced artificial intelligence systems designed to process, understand, and generate human-like text. Built on transformer architectures, LLMs like OpenAI's GPT-4, Google's PaLM, and Meta's LLaMA leverage vast datasets—often billions of words drawn from books, websites, and social media—to train deep neural networks [4]. These models excel in natural language tasks, including text generation, translation, sentiment analysis, and conversational dialogue. Their strength lies in contextual understanding, enabled by mechanisms like attention, which weigh the importance of words in a sequence to produce coherent responses [25].

For instance, when prompted with "How can I manage stress from gaming?", an LLM might draw on patterns in mental health literature to suggest coping strategies, tailoring its tone to the user's emotional state. This adaptability stems from training on diverse corpora, allowing LLMs to mimic human-like reasoning, though they lack true consciousness or emotional awareness [26].

3.2. Mechanisms and Capabilities

The transformer architecture, introduced by Vaswani et al. [25], underpins modern LLMs. It consists of interconnected layers that process input sequentially, using self-attention to prioritize relevant information. For example, in the sentence "Gaming helps me relax, but I can't stop," the model identifies "can't stop" as a key indicator of distress. Training involves supervised fine-tuning (SFT) and reinforcement learning from human feedback (RLHF), which refine the model's alignment with user expectations [27].

LLMs support text generation, enabling coherent narratives and tailored responses—for example, drafting personalized therapy scripts aligned with a patient's goals and language preferences. They also perform sentiment analysis to infer the emotional tone of user inputs, which can help flag IGD-related distress [28]. In dialogue settings, modern models exhibit context retention, carrying information across turns so multi-turn exchanges feel natural and build on prior content. Finally, beyond text, multimodal integration is increasingly available: models such as GPT-4 can accept images in addition to text, and large multimodal models (LMMs) described in WHO's 2024 guidance extend these capabilities to other data types (e.g., audio), though specific functions and safety profiles vary across systems [26,55].

Performance metrics (e.g., perplexity, BLEU) indicate proficiency; GPT-4 achieves near-human performance on several benchmarks [26]. Domain-adapted medical LLMs achieve state-of-the-art accuracy on MedQA (e.g., Med-PaLM 2) and, in human evaluations, physicians often prefer its answers on multiple clinical axes, while specialists still outperform the model—highlighting promise and limits for clinical use [59]. However, their "black box" nature—where decision-making processes are opaque—poses challenges for clinical trust [8].

3.3. LLMs in Mental Health Applications

LLMs have gained traction in mental health, addressing gaps in accessibility and scalability. Outside formal trials, an observational comparison found chatbot responses to public patient questions were preferred to physicians' 79% of the time and rated higher in quality and empathy, suggesting value for drafting supportive patient communications with clinician oversight [58].

Recent reviews document a rapidly growing literature applying LLMs to psychiatry and digital mental health, spanning four main domains: screening and assessment, therapeutic dialogue, sentiment/risk monitoring, and education/support. In screening and assessment, chatbots such as Woebot have shown symptom reductions in randomized trials with college students and can leverage NLP to sustain user engagement [29]. For therapeutic dialogue, AI-driven conversational agents can deliver CBT-consistent interactions that fit within clinician workflows [30], and across small trials recent reviews describe modest symptom improvements [32,33]. In sentiment and risk monitoring, machine-learning approaches to suicidality detection on social media display variable performance and notable methodological heterogeneity [31]. Finally, LLMs are increasingly used for education and support, providing real-time psychoeducation—such as explaining addiction triggers—to help users understand and manage their behavior [32].

Digital therapeutics show promise in substance-use care; however, LLM-specific evidence in gambling disorder remains sparse and requires rigorous trials [7]. These successes suggest potential for IGD, though domain-specific adaptations are needed.

3.4. Advantages and Limitations in Mental Health

LLMs offer several practical advantages for mental-health applications. They provide continuous, around-the-clock access that can reach users in remote or underserved settings [33]. They can also be personalized—via fine-tuning or conditioning on a user's history and cultural context—so that responses better align with individual needs [34]. Finally, unlike individual clinicians, LLM-based systems can scale to support very large cohorts simultaneously, extending reach without linear increases in staffing.

Despite their promise, LLMs carry important limitations. They simulate empathy rather than experience it, which can yield interactions that feel polished yet emotionally superficial or mismatched to a user's state [35]. Their outputs also inherit biases and errors from training data, at times surfacing as inappropriate or stereotyped mental-health advice (for example, gendered assumptions) and becoming more pronounced under distribution shift or ambiguous prompts [8]. Finally, the evidence base remains nascent: most mental-health applications lack large-scale, prospective, longitudinal trials demonstrating clinical outcomes, safety, and cost-effectiveness, with current support leaning on benchmarks, small studies, or short follow-ups [36].

These characteristics make LLMs powerful yet potentially risky tools with transformative potential that require careful oversight to ensure safety and efficacy in IGD contexts.

4. Integrating LLMs with Internet Gaming Disorder

4.1. Diagnostic and Screening Applications

Early detection of IGD is critical, as delayed intervention exacerbates outcomes [3]. LLMs can enhance screening through conversational interfaces integrated into gaming platforms or mobile apps. For example, a chatbot embedded in a gaming client could engage users in casual dialogue, analyzing responses for IGD symptoms like preoccupation or withdrawal.

Notably, Strojny et al. (2024) showed that a transformer-based analysis of gamers' open-ended responses could moderately predict their standardized IGD scores ($r \approx 0.48$) [37]. This suggests that

NLP models can capture meaningful signals of gaming pathology from text. Building on this, other AI-driven screening approaches have leveraged objective data: for instance, machine-learning classifiers using neurophysiological inputs (EEG/fNIRS) differentiate IGD from comparison groups (alcohol use disorder or healthy controls) with around 70–88% accuracy [61,62]. Objective multimodal and neurophysiological classifiers have also shown promise in IGD, including tablet-based digital phenotyping in schools [63], EEG- and neuropsychology-based classification [61], fNIRS deep learning [62], and fMRI connectomics and MVPA severity prediction [57,64] (see Appendix A). Table 1 summarizes these studies, including their training methods, sample sizes, and validation metrics.

Such systems could leverage validated scales, like the Internet Gaming Disorder Scale-9-item Short Form (IGDS9-SF), by converting its 9-item questionnaire into natural dialogue [12,38]. For instance, instead of asking, “Do you feel preoccupied with gaming?”, the LLM might say, “Hey, do you find yourself thinking about gaming a lot, even when you’re offline?” This approach reduces stigma and increases engagement, especially among adolescents wary of formal assessments.

Moreover, LLMs can integrate multimodal data—text, gaming hours, or even biometric inputs from wearables—to improve diagnostic precision. Mobile phone sensors and machine learning can detect substance-use events in the wild [39]. For IGD, this could mean monitoring in-game behavior (e.g., session length) alongside chat interactions to flag at-risk users.

4.2. Therapeutic Interventions

LLMs hold promise for delivering scalable, evidence-informed interventions for IGD, with Cognitive Behavioral Therapy (CBT)—the most established treatment—readily adapted into modular, conversational formats. However, it is important to note that no peer-reviewed clinical trial of an LLM-driven therapy for IGD has been published to date, reflecting the generally nascent evidence base for AI in addiction treatment. Preliminary successes have been observed in adjacent fields – for example, fully automated chatbots delivering cognitive-behavioral therapy for depression and anxiety have shown efficacy in randomized controlled trials (RCTs) – but in IGD, similar implementations remain untested. Future IGD trials should pre-register primary outcomes (e.g., gaming hours, functional impairment, relapse) and report calibration and external validation where prediction components are used. This underscores the need for rigorous trials to validate AI-guided interventions (e.g. motivational interviewing chatbots or personalized game-reduction programs) on user outcomes such as gaming time, psychosocial functioning, and relapse rates. In practice, a system can personalize CBT exercises to a user’s triggers—such as escapism or social pressure—by eliciting brief histories and then generating targeted practice, for example stress-management tasks and reframing prompts for a student who reports “I game to avoid school stress” [40,44]. To sustain engagement, CBT elements can be embedded in interactive, game-like narratives that award progress for completing mindfulness or behavioral-activation “quests”; early digital CBT applications demonstrate feasibility, though robust IGD-specific trials remain limited [32,33]. Complementing skills work, LLMs can be guided by motivational interviewing (MI) principles to help users surface and resolve ambivalence—for instance, posing reflective questions (“What do you enjoy about gaming, and what might you gain by cutting back?”) and summarizing change talk to support goal setting [42].

Real-time intervention is another frontier. LLMs integrated into gaming platforms could detect excessive sessions (e.g., 6+ hours) and prompt breaks or coping strategies, akin to Google’s Digital Wellbeing features [43].

4.3. Monitoring and Relapse Prevention

Continuous monitoring is vital given measurable recurrence in longitudinal cohorts. Prospective cohort evidence in adolescents shows IGD recurs in 16.1% over 24 months, underscoring the need for relapse-sensitive monitoring strategies [48]. LLMs can track user progress via regular check-ins, analyzing text for signs of distress or relapse (e.g., increased gaming references). Emerging work in

doi:10.20944/preprints202512.1552.v1

adjacent domains is exploring AI-driven relapse prevention, but rigorous evidence remains limited [36].

For IGD, LLMs could integrate with wearables or gaming APIs to monitor playtime, sleep patterns, or even social media activity, creating a holistic profile. If a user's gaming spikes alongside phrases like "I'm so stressed," the LLM could initiate a supportive dialogue or escalate to a human clinician. This hybrid model ensures AI handles routine tasks while clinicians address complex cases.

To orient the reader to practical use, Table 2 summarizes where LLMs can add value across the IGD care pathway—screening/assessment, therapeutic support, monitoring/relapse prevention, research/analytics, and governance/safety—and links each task to an example implementation and representative sources. The table is intended as a concise map from clinical questions to technical capabilities, highlighting both opportunities and the references that substantiate them.

doi:10.20944/preprints202512.1552.v1

Table 2. LLM-augmented IGD care pathway: tasks, example implementations, and representative evidence.

Care stage	LLM-enabled task	Primary risks	Safeguards / implementation notes	Implementation readiness
Screening / Early detection	Triage chat + short screener routing (e.g., IGDS9-SF prompts, risk-language detection)	False positives/negatives; misclassification of subthreshold cases	Validate against gold standards; calibrate thresholds; human review for positives	Medium
Diagnostic assessment & comorbidity review	Structured history; comorbidity prompts (ADHD/AUD/anxiety/depression)	Overconfidence; hallucinations; minority bias	Structured prompts; require citations; clinician-in-the-loop; bias testing	Low–Medium
Psychoeducation & motivational enhancement	Explain IGD mechanisms; goals; MI-style reflections	Inaccurate advice; uncalibrated empathy; “therapy” drift	Restrict scope; scripted safety language; clear AI disclosures	Medium
CBT skills coaching & relapse prevention	Homework adherence; cognitive restructuring; craving plans; play-time contracts	Over-reliance; poor escalation	Escalation trees; relapse signals reviewed by clinician; EHR integration	Low–Medium
Continuous monitoring	Summarize diaries/telemetry; flag risk phrases; weekly reports	Privacy leakage; false alarms; clinician burden	Minimize data; privacy-by-design; precision tuning; opt-in	Low
Crisis & safety (adjunct only)	Detect suicidality cues; present resources; encourage contacting humans	Unsafe responses; wrong numbers; delayed escalation	Location-aware hotlines; immediate human handoff; prohibit “how-to”; red-team tests	Low
Governance & compliance	Consent logging; model versioning; bias/robustness reports; DPIAs	Non-compliance; poor traceability	Follow WHO LMM guidance; EU AI Act staged timelines; audit trails; human accountability	Medium

4.4. Research and Data Analysis

LLMs can accelerate IGD research by analyzing large-scale datasets, such as forum posts or in-game chat logs, to identify behavioral patterns. For instance, a 2024 study used open-ended text responses and NLP to screen for gaming disorder [37]. LLMs and generative methods can be used to create synthetic patient cohorts to support hypothesis generation and methods development, potentially accelerating early-stage research, but they do not replace the need for clinical trials [46]. In qualitative research, LLMs are being explored to assist qualitative coding of interview transcripts [47].

4.5. Case Studies and Preliminary Evidence

Preliminary evidence from digital mental-health chatbots suggests symptom reductions in small RCTs (e.g., Woebot for depression/anxiety among students) [29]; recent evaluations of AI agents' handling of suicidality highlight safety gaps [56]; early NLP work shows feasibility for IGD risk screening [37]; and WHO guidance outlines guardrails for LMM safety [55].

5. Challenges and Ethical Considerations

5.1. Ethical Concerns

Integrating LLMs into IGD care surfaces a set of ethical challenges that must be addressed upfront. Data privacy is paramount: gaming platforms and companion apps routinely collect sensitive information—such as playtime patterns and in-game chat logs—and, without robust encryption, clear consent protocols, and strict access controls, incorporating these streams into conversational systems heightens breach risk under frameworks like the GDPR [49]. Recent disclosures by a digital-mental-health provider illustrate the stakes, with a large-scale privacy incident underscoring how quickly trust can be lost when safeguards fail [50]. Algorithmic bias is a second concern. Models trained on skewed or incomplete data may misclassify users or offer inappropriate guidance; moreover, cultural differences in gaming norms and known cross-national measurement disparities can tilt both classification and prevalence estimates, making cross-cultural validation and the use of global meta-analytic evidence essential before deployment across regions [2,15]. A third risk is over-reliance: some users may prefer an always-available AI to human clinicians, which can subtly dehumanize care or delay escalation. Survey data reflect this tension—while a 2024 nationally representative YouGov poll found that 55% of 18–29-year-old Americans would feel comfortable discussing mental-health concerns with a confidential AI chatbot, the figure drops to about one-third (34%) among adults overall, with misdiagnosis and data-safety worries commonly cited [51].

To mitigate these risks, implementation should align with the FUTURE-AI international consensus guideline—six principles that emphasize fairness, universality, traceability, usability, robustness, and explainability—and draw on its thirty best practices spanning development, validation, deployment, and post-market monitoring in healthcare contexts [60].

5.2. Technical Challenges

From a technical standpoint, several constraints shape what is feasible today. Generalization remains a challenge: out-of-the-box models can miss IGD-specific nuances and terminology, often necessitating fine-tuning or targeted adaptation on gaming-related corpora to perform reliably in this domain [36]. Interpretability is a second hurdle. The opaque nature of contemporary LLMs complicates clinical validation and oversight, making explainable AI techniques essential to clarify how outputs are produced and to support auditability in practice [52]. Finally, resource intensity—the computational and energy costs of training and deploying these systems—can be substantial, which risks exacerbating access gaps, particularly in low-income regions and resource-constrained clinical settings [53].

5.3. Clinical Limitations

Clinically, three limitations constrain the safe use of LLMs in IGD care. First, while models can simulate empathic language, they cannot reproduce the human connection that is often therapeutic in itself—an especially salient gap for patients with social deficits [30]. Second, the evidence base is thin: most evaluations rely on small samples or self-report outcomes, and rigorous randomized controlled trials remain scarce, leaving questions about durability, real-world effectiveness, and safety unanswered [36]. Third, crisis escalation remains inconsistent; systems should reliably detect and prioritize risk language (e.g., suicidal ideation) and route users to human support and emergency resources, yet recent evaluations show variable responses and a lack of standardized, audited escalation procedures [56]. Accordingly, IGD-specific, preregistered trials with external validation, model cards, and data/code availability (where feasible) are priorities to move beyond research-only prototypes.

6. Future Directions

As the fields of artificial intelligence and mental health continue to evolve, the integration of Large Language Models (LLMs) with Internet Gaming Disorder (IGD) offers numerous innovative possibilities. While current applications demonstrate promising preliminary results, the path forward requires a concerted effort to address existing gaps through advanced research, technological enhancements, and interdisciplinary collaboration. This section outlines key future directions, emphasizing scalable, ethical, and evidence-based approaches to harness LLMs for IGD management.

6.1. Hybrid Human-AI Therapeutic Models

One of the most compelling avenues is the development of hybrid models that combine LLM capabilities with human oversight. In these systems, LLMs could handle initial screenings, routine check-ins, and basic interventions, while escalating complex cases—such as those involving severe comorbidity with depression or suicidal ideation—to licensed clinicians. This tiered approach not only optimizes resource allocation but also mitigates risks associated with AI's limitations in empathy and nuanced judgment. For instance, a future platform might use LLMs to generate personalized CBT scripts based on user inputs, which a therapist then reviews and refines during sessions. Early explorations in mental health suggest that such hybrid human-AI models may improve adherence, but robust trials are still needed [36]. Extending this to IGD, pilot programs could integrate LLMs with telehealth services, allowing real-time collaboration where the AI analyzes gaming data and suggests interventions, while the human provides emotional support.

Moreover, reinforcement learning techniques could enable LLMs to learn from clinician feedback, iteratively improving their responses. Imagine an LLM trained on anonymized IGD case logs, adapting its motivational interviewing style to better resonate with gamers' unique vernacular and motivations. Research on LLMs in psychiatry highlights the potential for such adaptive systems to enhance clinical reasoning, but calls for rigorous testing in behavioral addictions [36]. Future studies should prioritize RCTs comparing hybrid models to traditional CBT, measuring outcomes like reduced gaming hours, improved social functioning, and long-term relapse rates.

6.2. Multimodal and Immersive Integrations

Advancing beyond text-based interactions, multimodal LLMs—incorporating audio, video, and sensor data—could revolutionize IGD interventions. For example, integrating LLMs with wearables or gaming hardware to monitor physiological markers like heart rate variability during play sessions could provide real-time insights into stress triggers. As multimodal LLMs (e.g., GPT-4) can integrate images, audio and sensor data [26], they could be used to synthesize multiple streams and predict addiction trajectories. In IGD, this might manifest as an AI companion that analyzes voice tone in-game chats for signs of frustration or withdrawal, then intervenes with calming prompts or suggests alternative activities.

Virtual and augmented reality (VR/AR) integrations represent another frontier. LLMs could power immersive therapeutic environments where users confront gaming triggers in a controlled setting, similar to exposure therapy for phobias. A gamified VR module, guided by an LLM-narrated storyline, could teach coping skills through interactive quests, blending entertainment with treatment to boost engagement among reluctant adolescents [41]. Preliminary work on AI in addiction suggests that such immersive tools could reduce cravings by simulating real-world consequences of excessive gaming [24]. However, developing these requires collaboration between game developers, psychologists, and AI experts to ensure cultural sensitivity and accessibility.

6.3. Longitudinal Research and Validation

To establish LLMs' efficacy in IGD, future research must shift from small-scale pilots to large, longitudinal studies. Current evidence is hampered by short follow-up periods and self-reported measures; multi-year RCTs tracking biomarkers, neuroimaging, and real-world behaviors are essential. For instance, a global cohort study could deploy LLM-based apps to thousands of gamers, using machine learning to correlate intervention usage with outcomes like academic performance or mental health scores. Insights from LLMs in mental health education and assessment underscore the need for diverse datasets to minimize biases, particularly in underrepresented populations [33]. Future methodological reviews of LLM-based IGD prediction tools should explicitly follow standardized reporting checklists for prediction-model reviews [54].

Additionally, exploring LLMs' role in etiological research could uncover novel insights. By analyzing vast datasets from gaming forums and social media, LLMs might identify emerging patterns, such as how new game mechanics (e.g., AI-driven opponents) exacerbate addiction risks. A study on evolutionary perspectives in behavioral addictions hints at using AI to model how digital environments exploit human instincts, informing preventive designs [16]. Future directions should include explainable AI (XAI) frameworks to demystify LLM decisions, fostering trust among clinicians and regulators.

6.4. Ethical Frameworks and Policy Development

Ethical innovation must parallel technological advances. Developing standardized guidelines for LLM deployment in IGD, aligned with the EU Artificial Intelligence Act (Regulation (EU) 2024/1689) [45], could address privacy, bias, and accountability. For example, requiring "human-in-the-loop" for high-stakes decisions and mandating transparency in training data. Ethical reviews of AI in mental health emphasize the importance of informed consent and data sovereignty, especially for vulnerable groups like minors [49]. Future policies might incentivize open-source LLM models tailored for mental health, reducing costs and promoting equity.

Global collaborations are crucial, given IGD's varying prevalence across cultures. Initiatives like WHO-led consortia could standardize LLM adaptations, ensuring they account for linguistic and societal differences. Moreover, addressing emerging risks, such as AI-generated addictive content (e.g., deepfake games), will be vital as LLMs evolve.

6.5. Preventive and Public Health Applications

Shifting focus upstream, LLMs could power preventive tools embedded in educational curricula or gaming platforms. Chatbots delivering psychoeducation on healthy gaming habits could reach millions, potentially curbing IGD onset. A scoping review on treatments for IGD stresses the need for early interventions, where LLMs might excel in scalable outreach [3]. Public health campaigns using

LLM-generated content, like personalized infographics or videos, could raise awareness, drawing from successful AI applications in addiction prevention [7].

In summary, these directions promise to transform IGD from a reactive treatment model to a proactive, AI-augmented ecosystem. Realizing this vision demands investment in research infrastructure, ethical safeguards, and cross-sector partnerships.

6.6. Limitations of This Review

As a narrative (non-PRISMA) review, our synthesis may be subject to selection and reporting biases; effect sizes and metrics were taken as reported by each study, and most models lacked external validation or publicly released code/data. Although we updated our targeted search through October 2025, the evidence base remains small and heterogeneous, especially for therapeutic interventions; meta-analytic conclusions were therefore not attempted.

7. Conclusion

In summary, across the IGD-specific AI/ML studies reviewed, we found training approaches including transformer-based text regression, multimodal EEG models, CNNs on fNIRS data, and fMRI MVPA methods. Validation/performance ranges from $r \approx 0.48$ for open-ended text screening to ~71–88% accuracy for physiology-based classification against comparison groups; dataset sizes span $n = 40$–417; model openness is generally limited (research-only, no public code/data). Therapeutic RCTs targeting IGD with LLMs remain absent, highlighting a near-term need for prospective trials and external validation.

The intersection of Large Language Models and Internet Gaming Disorder represents a pivotal moment in mental health innovation, offering tools that could make interventions more accessible, personalized, and effective. As we have explored, IGD's rising prevalence—driven by immersive digital landscapes—calls for novel approaches beyond traditional therapies. LLMs, with their prowess in natural language understanding and generation, are well-positioned to help fill this gap through diagnostic chatbots, therapeutic simulations, and ongoing monitoring. Preliminary evidence from pilots in addictions and mental health suggests reductions in symptoms and improved engagement, yet these gains must be balanced against ethical pitfalls like privacy breaches and algorithmic biases [8,50].

Ultimately, while LLMs are not a standalone cure, their integration could democratize care, particularly for underserved populations. By fostering hybrid models and rigorous validation, we can ensure that AI enhances—rather than supplants—human compassion. As technology advances, so too must our commitment to ethical stewardship, paving the way for a future where gaming can enrich lives without trapping individuals in harmful patterns. This review synthesizes the current landscape and outlines a path forward, urging stakeholders to collaborate in realizing LLMs' full potential for combating IGD.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADHD	Attention-Deficit/Hyperactivity Disorder
AI	Artificial Intelligence
ALFF	Amplitude of low-frequency fluctuation
APA	American Psychiatric Association
AUD	Alcohol Use Disorder
BLEU	Bilingual Evaluation Understudy (MT metric)
CBT	Cognitive Behavioral Therapy

doi:10.20944/preprints202512.1552.v1

COVID-19	Coronavirus Disease 2019
DPIA(s)	Data Protection Impact Assessment(s)
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
EHR	Electronic Health Record
EU	European Union
FUTURE-AI	International consensus guideline for trustworthy, deployable AI
fNIRS	Functional near-infrared spectroscopy
GDPR	General Data Protection Regulation
GPT-4	Generative Pre-trained Transformer 4
ICD-11	International Classification of Diseases, 11th Revision
IGD	Internet Gaming Disorder
IGDS9-SF	Internet Gaming Disorder Scale–Short Form (9-item)
IGDT-10	Ten-Item Internet Gaming Disorder Test
LLM(s)	Large Language Model(s)
LMM(s)	Large Multimodal Model(s)
NLP	Natural Language Processing
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT(s)	Randomized Controlled Trial(s)
ReHo	Regional Homogeneity
RLHF	Reinforcement Learning from Human Feedback
SFT	Supervised Fine-Tuning
SVM	Support Vector Machine
VBM	Voxel-Based Morphometry
VR/AR	Virtual Reality / Augmented Reality
WHO	World Health Organization

Appendix A. Expanded Details for IGD-Specific AI/ML Studies

A1. Strojny et al., 2024 — Text/NLP Screening (JMIR Serious Games)

- Model/training: Transformer embeddings (HerBERT) + ridge regression; no task-specific fine-tuning.
- Dataset: *n* = 417 gamers after quality control; four open-ended responses + Gaming Disorder Test.
- Validation/performance: Pearson r = 0.476 using all four responses; single-item r = 0.274–0.406.
- Open vs closed: Used an open-source LM (HerBERT); code/data not publicly released.
- Limitations: Correlational (screening signal, not diagnosis); self-report criterion; no external validation.

A2. Cho et al., 2024 — School Digital Phenotyping (JMIR Mental Health)

- Model/training: Multiple regression with bootstrapping + MANOVA; interpretability prioritized over prediction.
- Dataset: *n* = 168 adolescents (85 IGD, 83 non-IGD); tablet sensor metrics during classes.
- Validation/performance: Five digital markers explained 23% of IGDS variance; group differences ~Cohen's d ≈ 0.40 (moderate).
- Open vs closed: Methods published; no public dataset/model.
- Limitations: Screening (not confirmatory diagnosis); single-school context; no ML classifier.

A3. Lee et al., 2024 — EEG + Neuropsychology Classification (Comprehensive Psychiatry)

- Model/training: L1-SVM, Random Forest, and L1-logistic regression; unimodal EEG, unimodal neuropsych, and multimodal feature sets.

- Dataset: IGD *n* = 67, AUD *n* = 58, HC *n* = 66; resting-state EEG + psychometrics.
- Validation/performance: Best model (logistic regression, multimodal) for IGD vs AUD: accuracy = 71.2%; salient delta/beta source connectivity + demographics.
- Open vs closed: Open-access article; code/data not released.
- Limitations: Case–control; no external validation; limited clinical utility metrics (e.g., AUC, calibration).

A4. Wang et al., 2023 — fNIRS Deep Learning (Biomedizinische Technik/Biomedical Engineering)

- Model/training: Prefrontal fNIRS during stop-signal task; seven ML/DL algorithms compared; 2D-CNN performed best.
- Dataset: *n* = 40 (24 IGD, 16 HC).
- Validation/performance: Hold-out accuracy = 87.5% (2D-CNN), outperforming traditional ML.
- Open vs closed: Methods described; code/data not public.
- Limitations: Small sample; single site; hold-out split (no cross-site external validation).

A5. Song et al., 2021 — Resting-State fMRI Connectome SVM (Addiction Biology)

- Model/training: Modified connectome-based predictive modeling (CPM) with SVM; classification + regression.
- Dataset: IGD *n* = 72, HC *n* = 41; resting-state fMRI.
- Validation/performance: DMN-based model accuracy = 78.76% (balanced accuracy = 75.46%; sensitivity = 63.41%; specificity = 87.5%); DMN features also predicted CIAS severity (r = 0.44).
- Open vs closed: Data available on request; MATLAB scripts mentioned; no public repo.
- Limitations: Predominantly young adult males; no external replication; cultural/generalizability considerations.

A6. Ye et al., 2022 — MVPA Predicting IGD Severity (Journal of Affective Disorders)

- Model/training: MVPA on resting-state fMRI features (ReHo, ALFF) to predict continuous IGD severity.
- Dataset: *n* = 402 participants spanning the IGD severity continuum.
- Validation/performance**:** Neural patterns significantly predicted severity; high-weighted regions then probed by graph/causality analyses.
- Open vs closed: No public code/data noted.
- Limitations: Cross-sectional; thresholds for clinical decision-making not established; potential site/scan heterogeneity.

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; American Psychiatric Publishing: Arlington, VA, USA, 2013.
2. King, D.L.; Delfabbro, P.H.; Griffiths, M.D.; et al. Global prevalence of gaming disorder: A systematic review and meta-analysis. Aust. N. Z. J. *Psychiatry* **2021**, *55*, 553–568.
3. Zajac, K.A.; et al. Treatments of internet gaming disorder: A systematic review of the evidence. Expert Rev. *Neurother*. **2020**, *20*, 85–93.
4. Brown, T.; et al. Language models are few-shot learners. *Adv. Neural Inf. Process. Syst*. **2020**, 1877–1901.
5. Wei, J.; et al. Emergent abilities of large language models. *arXiv* **2022**, 2206.07682.
6. Király, O.; et al. Preventing problematic Internet use during the COVID-19 pandemic: Consensus guidance. *Compr. Psychiatry* **2020**, *100*, 152180.
7. Marsch, L.A.; et al. Digital therapeutics for substance use disorders: The state of the science and vision for the future. *Am. J. Psychiatry* **2021**, *178*, 396–399.
8. Bender, E.; et al. On the dangers of stochastic parrots: Can language models be too big? *ACM FAccT* **2021**, 610–623.

9. World Health Organization. International Classification of Diseases (ICD-11). 2019.
10. Paulus, F.C.; et al. Internet gaming disorder in children and adolescents: A systematic review. *Dev. Med. Child Neurol*. **2018**, *60*, 645–659.
11. Ko, Y-M.; Lee, E.S.; Park, S. Prevalence, correlates, and comorbidities of internet gaming disorder and problematic game use: National mental health survey of Korea 2021. *Front. Psychiatry* **2024**, *15*, 1442224.
12. Pontes, H.; Griffiths, M. Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. *Comput. Hum. Behav*. **2015**, *45*, 137–143.
13. Kuss, D.; Griffiths, M. Internet gaming addiction: A systematic review of empirical research. *Int. J. Ment. Health Addict*. **2012**, *10*, 278–296.
14. Potenza, M.N. Neurobiology of substance use in adolescents and potential therapeutic effects of exercise for prevention and treatment of substance use disorders. *Birth Defects Res*. **2017**, *109*, 1711–1729.
15. Király, O.; et al. Ten-Item Internet Gaming Disorder Test (IGDT-10): Measurement invariance and cross-cultural validation across seven countries. *Psychol. Addict. Behav*. **2017**, *31*, 391–403.
16. King, D.L.; Delfabbro, P.H. The cognitive psychology of Internet gaming disorder. *Clin. Psychol. Rev*. **2014**, *34*, 298–308.
17. Niu, X.; Jiang, W.; Liu, X.; et al. Structural and functional brain alterations in Internet Gaming Disorder: A meta-analysis. *Front. Psychiatry* **2022**, *13*, 1029344.
18. Griffiths, M. Internet gambling: Issues, concerns, and recommendations. *CyberPsychol. Behav*. **2003**, *6*, 557–568.
19. Zendle, D.; Cairns, P. Video game loot boxes are linked to problem gambling: Results of a large-scale survey. *PLoS One* **2018**, *13*, e0206767.
20. Hofmann, S.G.; et al. The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *J. Consult. Clin. Psychol*. **2010**, *78*, 169–183.
21. Song, J-H.; et al. Comparative study of the effects of bupropion and escitalopram on Internet gaming disorder. *Psychiatry Clin. Neurosci.* **2016**, *70*, 967-975.
22. Young, K.S. Treatment outcomes using CBT-IA with Internet-addicted patients. *J. Behav. Addict*. **2013**, *2*, 209–215.
23. Chen, Y.; Lu, J.; Wang, L.; Gao, X. Effective interventions for gaming disorder: A systematic review of randomized controlled trials. *Front. Psychiatry* **2023**, *14*, 1098922.
24. Maden, Ç.; Bayramlar, K.; Arıcak, O.T.; Yagli, N.V. Effects of virtual reality–based training and aerobic training on gaming disorder, physical activity, physical fitness, and anxiety: A randomized, controlled trial. *Ment. Health Phys. Act*. **2022**, *23*, 100465.
25. Vaswani, A.; et al. Attention is all you need. *Adv. Neural Inf. Process. Syst*. **2017**, 5998–6008.
26. OpenAI. GPT-4 technical report. *arXiv* **2023**, 2303.08774.
27. Christiano, P.; et al. Deep reinforcement learning from human preferences. *arXiv* **2017**, 1706.03741.
28. Gozalo-Brizuela, R.; Garrido-Merchán, E.C. ChatGPT is not all you need: A state-of-the-art review of large generative AI models. *arXiv* **2023**, 2301.04655.
29. Fitzpatrick, K.; et al. Delivering cognitive behavior therapy using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Ment. Health* **2017**, *4*, e19.
30. Luxton, D.D. Artificial intelligence in psychological practice: Current and future applications and implications. *Prof. Psychol. Res. Pract*. **2014**, *45*, 332–339.
31. Nikmehr, G.; Bilbao-Jayo, A.; Almeida, A. Categorizing and assessing aspects of suicidal ideation detection approaches: A systematic review. *Comput. Hum. Behav. Rep*. **2025**, *19*, 100733.
32. Torous, J.; Bucci, S.; et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psych.* **2021**, *20*, 318–335.
33. Lim, S.M.; Shiau, C.W.C.; Cheng, L.J.; Lau, Y. Chatbot-Delivered Psychotherapy for Adults With Depressive and Anxiety Symptoms: A Systematic Review and Meta-Regression. *Behav. Ther*. **2022**, *53*, 334–347.
34. Abd-Alrazaq, A.; et al. An overview of the features of chatbots in mental health: A scoping review. *Int. J. Med. Inform*. **2019**, *132*, 103978.

35. Weizenbaum, J. ELIZA—A computer program for the study of natural language communication between man and machine. *Commun. ACM* **1966**, *9*, 36–45.
36. Torous, J.; Bucci, R. Integrating artificial intelligence into psychiatric practice: Opportunities and challenges. *Psychiatr. Serv*. **2020**, *71*, 1295–1298.
37. Strojny, P.; Kapela, K.; Lipp, N.; Sikström, S. Use of four open-ended text responses to help identify people at risk of gaming disorder: Preregistered development and usability study using NLP. *JMIR Serious Games* **2024**, *12*, e56663.
38. Pontes, H.M.; Griffiths, M.D. Assessment of Internet Gaming Disorder in clinical research: Past and present perspectives. *Clin. Res. Regul. Aff.* **2014**, *31*, 35–48.
39. Bae, S.; et al. Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults. *Addict. Behav*. **2018**, *83*, 42–47.
40. King, D.L.; et al. Cognitive behavioral therapy for problematic video game players: Conceptual considerations and practice issues. *J. CyberTher. Rehabil*. **2010**, *3*, 221–230.
41. Layous, K.; et al. Delivering happiness: Translating positive psychology intervention research for treating depressive disorders. *J. Altern. Complement. Med*. **2011**, *17*, 675–683.
42. Miller, W.R.; Rollnick, S. Motivational Interviewing: Helping People Change, 3rd ed.; Guilford Press: New York, NY, USA, **2012**.
43. Digital Wellbeing. Available online: https://wellbeing.google/ (accessed on 10 September 2025).
44. Young, K.S.; de Abreu, C.N. Internet Addiction: A Handbook and Guide to Evaluation and Treatment; John Wiley & Sons: Hoboken, NJ, USA, **2010**.
45. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Available online: https://data.europa.eu/eli/reg/2024/1689/oj (accessed on 1 September 2025).
46. Smolyak, D.; Bjarnadóttir, M.V.; Crowley, K.; Agarwal, R. Large language models and synthetic health data: Progress and prospects. *JAMIA Open* **2024**, *7*, ooae114.
47. Raza, M.M.; Venkatesh, K.P.; Kvedar, J.C. Generative AI and large language models in health care: Pathways to implementation. *NPJ Digit. Med*. **2024**, *7*, 62.
48. Jo, S.-J.; Lee, H.K.; Rumpf, H.-J.; Jeong, H.; Yim, H.W. Recurrence of internet gaming disorder in Korean adolescents: a 24-month follow-up study. *Eur. Child Adolesc. Psychiatry* **2025**, *34*, 1637–1647.
49. European Union. General Data Protection Regulation (GDPR). 2016.
50. Cerebral, Inc. Notice of HIPAA Privacy Breach. Available online: https://medcitynews.com/wp-content/uploads/sites/7/2023/03/cerebral-data-breach.pdf (accessed on 2 October 2025).
51. YouGov. Can an AI Chatbot be your therapist? A third of Americans are comfortable with the idea. Available online: https://yougov.com/en-us/articles/49480-can-an-ai-chatbot-be-your-therapist (accessed on 3 September 2025).
52. Samek, W.; et al. Explaining deep neural networks and beyond: A review of methods and applications. *Proc. IEEE* **2021**, *109*, 247–278.
53. Strubell, E.; et al. Energy and policy considerations for deep learning in NLP. *arXiv* **2019**, 1906.02243.
54. Pineiro, J.M.; et al. Reporting checklist for methodological reviews of prediction models in health care. *BMJ* **2021**, *372*, n6.
55. World Health Organization. Ethics and governance of artificial intelligence for health: Guidance on large multimodal models. 2024.
56. Pichowicz, W.; Kotas, M.; Piotrowski, P. Performance of mental health chatbot agents in detecting and managing suicidal ideation. *Sci. Rep*. **2025**, *15*, 31652.
57. Ye, S.; Wang, M.; Yang, Q.; Dong, H.; Dong, G.-H. Predicting the severity of Internet Gaming Disorder with resting-state brain features: A multi-voxel pattern analysis. J. Affect. Disord., 2022, 318, 113–122.
58. Ayers, J.W.; Poliak, A.; Dredze, M.; et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern. Med*. **2023**, *183*, 589–596.
59. Singhal, K.; Tu, T.; Gottweis, J.; et al. Toward expert-level medical question answering with large language models. *Nat. Med*. **2025**, *31*, 943–950.

60. Lekadir, K.; Feragen, A.; Fofanah, A.J.; et al. FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ* **2024**, *388*, e081554.
61. Lee, J.-Y.; Song, M. S.; Yoo, S. Y.; et al. Multimodal-based machine learning approach to classify features of internet gaming disorder and alcohol use disorder: A sensor-level and source-level resting-state electroencephalography activity and neuropsychological study. *Comprehensive Psychiatry*, **2024**, *130*, 152460.
62. Wang, Q.; Li, Y.; Su, H.; et al. Deep neural network to differentiate internet gaming disorder from healthy controls during stop-signal task: a multichannel near-infrared spectroscopy study. *Biomed Tech (Berl)*. **2023** Apr 27, *68*, 457-468.
63. Cho, K.; Kim, M.; Cho, Y.; et al. Digital Phenotypes for Early Detection of Internet Gaming Disorder in Adolescent Students: Explorative Data-Driven Study. *JMIR Ment. Health*, **2024**, *11*, e50259.
64. Song, K.-R.; Potenza, M.N.; Fang, X.-Y.; et al. Resting-state connectome-based support-vector-machine predictive modeling of Internet Gaming Disorder. *Addiction Biology*, **2021**, *26*, e12969.