**Research on the Image Shaping and Communication Effect of Digital Virtual Human under the Perspective of New Media Design**

WangYuXiao 1,a；GaoXiang2\*,b; Yiran Li3,c; Qingcheng Xie4,d

*1. Silla University, Busan,* *46958,* *Korea;*

*2. Huaiyin Institute of Technology，Huai'an,Jiangsu, 223001, China；*

*3. Xi’an Jiaotong-Liverpool University, Suzhou,Jiangsu, 215123, China;*

*4. Xi’an Jiaotong-Liverpool University, Suzhou,Jiangsu, 215123, China;*

*aEmail:* 694743591@qq.com

*bEmail:* [312291316@qq.com](mailto:312291316@qq.com)

cEmail: [gussed@163.com](mailto:gussed@163.com)

dEmail: [3124402244@qq.com](mailto:3124402244@qq.com)

**Abstract:**In order to explore the construction logic and dissemination mechanism of digital avatar image in new media context, we analyze the influence of four design dimensions, namely, visual recognition, interaction abundance, narrative consistency and data operation, on the dissemination effect, and construct multi-factor experimental and standardized evaluation models for empirical research. The results show that visual recognition has the highest explanatory power for the comprehensive effect of communication, followed by interaction depth and narrative consistency, and social presence and proposed social relationship play a significant mediating role. The study concludes that the optimization of the communication effect of digital avatars relies on the synergistic evolution of multidimensional design elements and strategic adjustment mechanism, and can be realized through the feedback of behavioral data to dynamically adjust the power and scene adaptation, and ultimately enhance the user's participation and emotional immersion.

**Keywords:** digital avatar; new media design; communication effect; interaction mechanism

**1 Introduction**

With the continuous evolution of media forms and the continuous expansion of technical boundaries, digital avatars, as a synthesizer integrating graphic generation, intelligent interaction and cultural narrative, are gradually becoming an important communication subject and emotional carrier in the new media environment. Under the mechanism of visual economy domination and algorithm-driven content distribution, the creation of virtual human image not only involves aesthetic expression, but also carries multiple functions such as identity construction, value transmission and user participation. Its communication effectiveness no longer relies on a single dimension of technical presentation, but needs to form a coordinating mechanism between multimodal interaction, narrative logic and data operation. Therefore, there is an urgent need to re-examine the construction logic and dissemination mechanism of digital avatars from the perspective of new media design, in order to respond to the real demand for simultaneous enhancement of communication effect and user experience in the era of media convergence.

**2 Core features of the new media environment**

The new media environment takes digitalization, networking and interactivity as its core features, and its communication form presents the trend of decentralization, diversification and instantaneousness. The production and consumption of information are synchronized in multiple platforms and terminals, and the audience is no longer a passive recipient, but participates in content generation and dissemination through comments, sharing and secondary creation. In this environment, media content highly relies on visual symbols, emotional narratives and immersive experiences to attract attention, while algorithmic recommendation and data analysis play a key role in user reach and accurate push [1]. The combination of virtual people and new media makes its image shaping not only need to be highly refined and stylized in visual presentation, but also to establish a continuous connection with the audience through interactive design, story building and community operation.

**3 Design dimensions of core image shaping of digital avatars**

**3.1 Visual Design Dimension**

In the context of high-frequency interaction and multi-distribution of new media, the visual design of digital avatars should aim at "recognizability × evolvability", and build a closed loop from geometric modeling - material and lighting - style unification - multi-adaptation. Closed-loop: Geometry level adopts bindable topology and LOD (to ensure mobile frame rate and close-up fidelity), and reserves expression/mouth-driven skeleton and hybrid deformation domains; Material level is based on the PBR workflow to unify the metallicity/roughness/normal channel and BRDF parameters, to ensure cross-engine consistency [2]; Lighting level uses environment-probe-based IBL and unified color mapping curves, to reduce the platform color bias; At the style level, the main color, shape vocabulary and iconic features are used to build a consistent "recognition anchor" across scenes, and the amplitude of movements and props interfaces are reserved for interactions and narratives. In order to avoid the imbalance of realization caused by a single aesthetic optimum, optimization can be expressed as multi-objective design:



Lgeom measures topology/binding errors, Ltex constrains normal and roughness consistency, Llight controls color differences under different lighting, and Lstyle maintains cross-end style consistency [3]; α,β,γ,δ are set by the placement scenario and the platform arithmetic. The structured design of this visual layer provides stable image semantics and rendering baseline for interaction mechanism, narrative expression and data-driven operation.

**3.2 Interaction Design Dimension**

In the interaction design dimension, the interaction system of the digital virtual human should realize efficient information exchange and emotional resonance through the multi-channel input and intelligent response mechanism, the core of which lies in the fusion processing and real-time optimization of the input modalities (voice, text, gesture capture, etc.) [4]. The system can define the interaction efficiency function



where Tlat is the response delay (s), Aacc is the semantic understanding accuracy (0-1), Perspers is the personalized matching degree (0-1), and w1,w2,w3 are the design weights, which are adjusted according to the application scenario. In order to realize immersion and sustainable interaction, a context-based dialogue memory module and an emotion recognition feedback loop should be introduced into the architecture, so that the virtual human can not only maintain semantic consistency in multiple rounds of dialogue, but also dynamically adjust expressions, intonations, and motions, thus shaping the interaction experience with a sense of social presence [5]. In terms of interface and interaction tempo, high-frequency trigger points should be located through interaction traffic heat map analysis to optimize the trigger command set and feedback granularity. For example, in short video scenarios, the interaction triggering delay can be controlled within 200ms to guarantee immediacy, while behavioral preference labels are introduced in personalized responses to achieve continuous experience design across sessions.

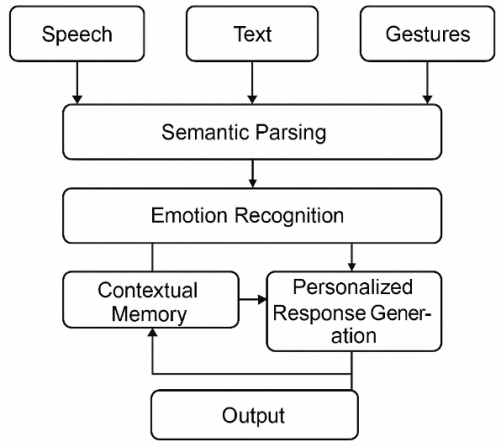


Fig. 1 Schematic diagram of the multi-channel interaction design framework for digital avatars

**3.3 Narrative and Content Design Dimension**

In the narrative and content design dimension, the core image shaping of digital avatars not only relies on the aesthetics and style of a single piece of content, but also depends on the narrative consistency and plot coherence across time and multiple contexts [6]. The design should build a content generation engine based on the three-layer structure of "worldview-role setting-plot driven", and store the character background, values, and behavioral logic in the narrative database in a parameterized way, which is invoked through the event triggering mechanism to ensure that the plots of different communication nodes are consistent with the time line of the digital avatars. This ensures that the plots of different dissemination nodes remain closed in terms of timeline and causal chain. Narrative consistency can be measured by the following formula:



Where Edev denotes the number of events that violate the established settings, Nplot is the total number of episodes, and a higher Cstory∈[0,1] indicates a smaller narrative deviation. In the specific implementation, it can be combined with emotion curve modeling to distribute the plot tension peaks in different interaction nodes to maintain user interest progression. At the same time, the content release rhythm should be matched with the platform algorithm rhythm, which is tuned through Table 1 [7] to realize the two-way supplementation of the main plot and fragmented content, so as to continuously strengthen the image stability and emotional immersion of avatars in the multi-scene environment of new media.

Table 1 Plot-pacing matching table

|  |  |  |  |
| --- | --- | --- | --- |
| Plot stage | Mainline node duration (sec) | Frequency of supplementary content (times/day) | Peak emotional weight (%) |
| Beginning buildup | 30-45 | 1 | 20 |
| Conflict development | 60-90 | 2 | 35 |
| Climactic turn | 90-120 | 3 | 40 |
| Finale | 30-60 | 1 | 25 |

**3.4 Data-driven and community-operated dimensions**

In the data-driven and community operation dimension, the key to the design is to incorporate visual parameters and interaction and narrative variables into the same growth-governance closure [8]: event collection → feature engineering → experimental scheduling → strategy dissemination → feedback correction as a chain, and strategically controlling the release frequency, versioning, community tasks, and cross-end distribution. To this end, an objective function can be constructed to constrain the operation intensity and compliance risk:



Where πθ is the scheduling strategy of content/time period/crowd, w is the incentive and community weight; A acquisition, Ret retention, K diffusion coefficient (social invitation × acceptance rate), Cost production and transportation cost, and Risk is measured by the indicators of violation/underage protection, etc.; and the frequency control and compliance model is set as a constraint (e.g., user-level frequency ≤ a set threshold) [9]. Figure 2 shows the path of buried points - feature library - multi-armed gambling/A/B engine - community orchestration (tasks, tiered benefits, co-creation mechanism) - real-time reflow to ensure the consistency and evolvability of avatar images in multiple scenarios. In order to achieve reusability and auditability, Table 2 gives the core indicators and data collection mapping to ensure caliber uniformity and privacy compliance.

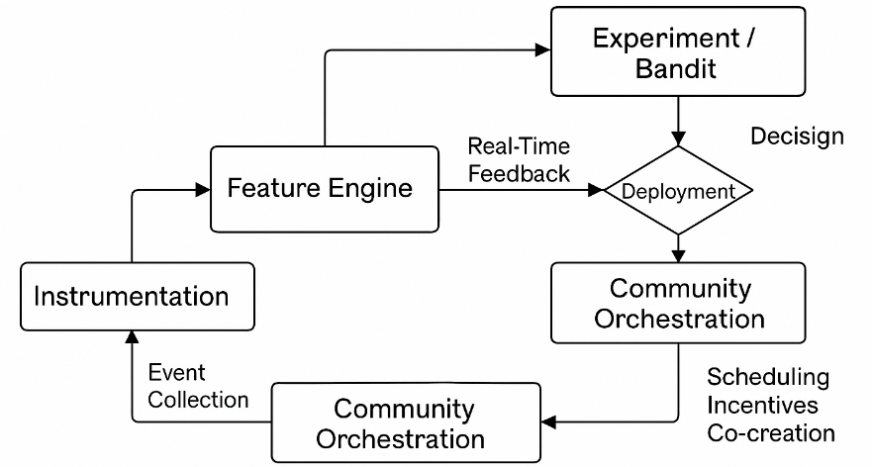


Figure 2 Data-driven community operation closed-loop framework

Table 2 Indicator-data collection mapping (design caliber)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Metric (Symbol) | Operationalization Definition | Data Source | Window | Remarks |
| Acquisition AA | Effective Reach/Target Audience | Platform Exposure and De-emphasis Log | 24h/7d | Filter abnormal traffic |
| Retention Ret(d)Ret(d) | d-day return rate | User session log | 1/7/28d | Layering to crowd profiling |
| K-factor KK | Average number of invites x acceptance rate | Invitation and Conversion Logs | 7d scrolling | Channel attribution required |
| UGC Density GG | UGC Volume/Active Users | Comments / Second Creation Log | 7d | Reflects community activity |
| Risk | Violation/reported rate | Review and Governance Logs | Real-time/Daily | Linkage with Compliance Module |

**4 Empirical Study on the Communication Effectiveness of Digital Avatars**

**4.1 Research Design**

In order to test the mechanism of design variables on the dissemination effect, this study adopts a multifactorial experiment among 8 groups of subjects, taking the unified virtual human parent as the baseline, manipulating only three independent variables on variable slots: visual recognition (high/low controlled by anchor points and color mapping constraints), interaction abundance (configured by multimodal and personalized pipeline as high/low), narrative consistency (set by the rule engine as high/low Cstory ), platform-side operational parameters are used as covariates and stratified randomization to control release time and population; user stratification is based on usage habits and device patterns to avoid learning effects. The online allocation strategy adopts a contextual dobby framework, and the scheduling probability of a given user context feature ϕ(x,a) with candidate condition a is:



where η is the learnable parameter and τ is the temperature to balance exploration/exploitation; the offline session ensures attributability with a blocked control. Subjective scales collect social presence, proposed social relationship, trust and attitude (immediate and follow-up), and objective logs record reach, stay, interaction, second creation and diffusion paths, and are aggregated by user level to provide metrics and model estimation [10]; the data flow follows the closed-loop "buried point→feature library→experimentation/distribution→return flow" (Figure 3). The data flow follows the closed loop of "buried point→feature library→experiment/distribution→reflow" (Figure 3). For uniformity, the variable-operationalization- data source mapping is shown in Table 3, which is used to drive stimulus generation and evaluation computation, and provides an auditable update interface for subsequent strategy weights (including α,β,γ,δ and λ vectors).

Table 3 Independent variable and measure mapping (experimental caliber)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Factor | Level Coding | Manipulation (design leverage) | Primary Indices | Log/Source |
| Visual Identity | High / Low | Constant anchors, style slot switching, consistent color mapping across ends | First few seconds retention, standardized viewing duration | Playback and exposure logs |
| Interaction Richness | High / Low | Modal count/response pipeline/personalized strategy configuration | Interaction Rate, Conversation Rounds, Trigger Depth | Interaction and Conversation Logs |
| Narrative Consistency | High / Low | Rule Engine Constraints Cstory and Worldview Knowledge Base | Consistency Perception, Intent to Revisit | Comment semantics + questionnaire |

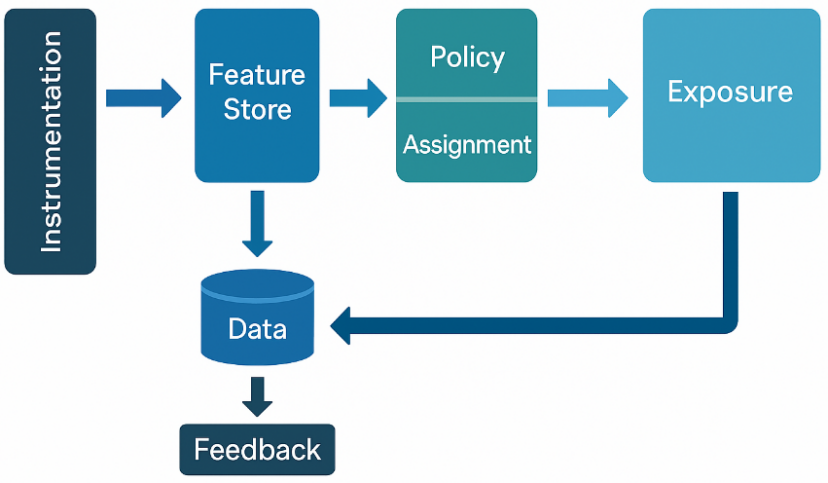


Figure 3 Schematic of research process and data flow

**4.2 Construction of communication effect evaluation index system**

In order to take over the four dimensions of visual, interactive, narrative and data operation, this study designs the communication effect as a measurement model of "four domains and two-level indicators": attention (A), participation (E), psychological mechanism (P), diffusion and transformation (D). The domains are aggregated from dual sources of platform logs and scales, and are calibrated and standardized for comparability across platforms and time periods (z-score). The comprehensive evaluation uses weighted synthetic metrics to drive pathways and mediation tests:



Where Z. is the standardized score for the corresponding domain and the weights w. are obtained from validated factor analysis loadings or cross-validated predictive validity learning (no preset values). The key calibers within the domains are as follows: the attention domain defines reach quality in terms of standardized viewing time NWT=WatchTime/VideoLength vs. first 3-second retention; the engagement domain measures interaction strength in terms of interaction rate ER=(Like+Comment+Share)/Impressions vs. depth of commenting (average floor/number of sub-strings); and the psychological domain employs the social Presence and Proposed Social Relationship Scale scores and introduces the narrative consistency Cstory already defined in the previous section as a structural constraint; the Diffusion and Conversion domain characterizes the radius of outreach and persistence with the K-factor K=Invites/User×AcceptanceRate and the revisit rate Ret(d). To ensure reuse and auditing, Table 4 gives the indicator-operationalization-data source mapping with fixed time windows (e.g., 24h/7d) and de-emphasis rules at the acquisition level to ensure that the experimental data and model estimation are articulated under the same measurement benchmark.

Table 4 Indicator-Operationalization-Data Source (Assessment System Caliber)

|  |  |  |  |
| --- | --- | --- | --- |
| Domain | Metric (symbol) | Operationalization Definition | Data Source/Window |
| Attention | NWTNWT, 3-sec Retention | Viewing duration/film length; ≥3s/exposure | Playback Log/24h |
| Engagement | ERER, Comment Depth | (Likes + Comments + Retweets)/Exposure; Average Floor | Interaction log/24h,7d |
| Psych | Social Proximity, Mimetic Society, Cstory | Scale score; narrative coherence constraints | Questionnaire + comment semantics/follow-up |
| Diffusion | K, Ret(d) | Invitation x acceptance rate; d-day return | Sharing/account logs/7d,28d |

**4.3 Data analysis and results**

Based on the aforementioned variable measurement and standardized assessment system, this study firstly carried out the reliability test of scale-like indicators, with Cronbach's α above 0.83 and KMO value up to 0.89, and passed the Bartlett's spherical test (p < 0.001), and after confirming that the scale structure was appropriate, two-factor analysis of variance (MANOVA) was used for the After confirming the adaptability of the scale structure, a two-way analysis of variance (MANOVA) was used to validate the main effect and interaction effect of the three factors of visual recognition, interaction richness and narrative consistency, and the results showed that all dimensions showed significant differences in the indicators of the four domains of communication, especially in the interaction between the attention domain and the psychological domain, with an F-value of higher than 4.70, and the level of significance reached 0.01. Further, the "design variables→psychological mechanism→communication indicators" model was constructed with the help of structural equation modeling (SEM). Further, with the help of structural equation modeling (SEM), a mediation path model of "design variables→psychological mechanisms→communication indicators" was constructed, as shown in Figure 4, in which Presence and Para-social Interaction constituted a double-mediation pathway, with standardized path coefficients of 0.25 or above, and the significance of the mediation was verified by the bootstrap method (95% CI without 0). CI without 0). In addition, to verify causal robustness, propensity score matching (PSM) was used to test with alternative indicators, and the results trended in the same direction, enhancing external validity. In order to clarify the contribution weight of each variable to the communication effect, this study combines XGBoost and SHAP value analysis, and the results are shown in Table 5, the explanatory power of recognition, interaction depth, and narrative consistency to the total effect is 34.1%, 27.6%, and 22.3%, respectively, which provides the data support for the adjustment of the weights and the dynamic evolution in the subsequent strategy optimization.

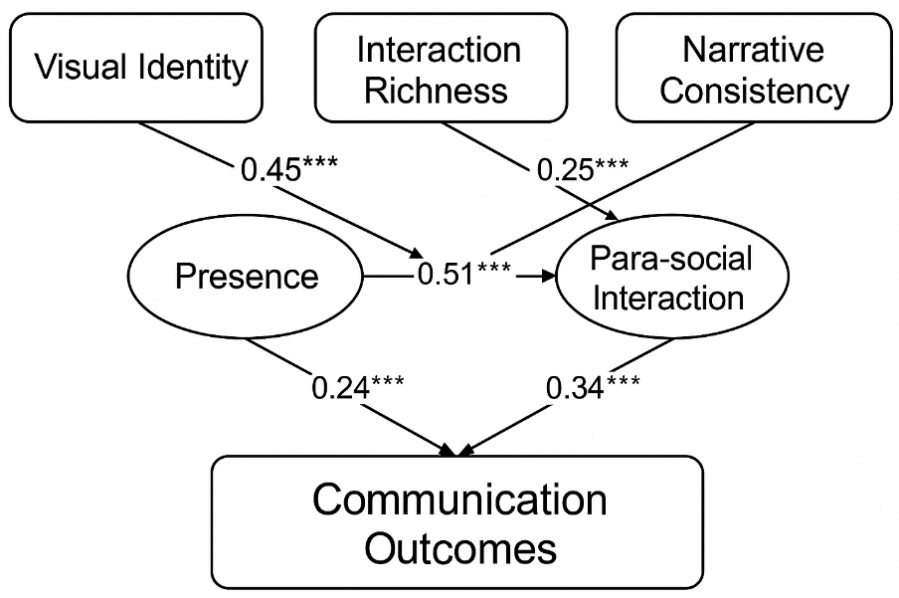


Fig. 4 Path model of design dimension-psychological mechanism-communication effect

Table 5 Distribution of the explanatory power of each design dimension on the total effect of communication (based on SHAP analysis)

|  |  |  |
| --- | --- | --- |
| Feature | Contribution to Ceff(%) | Importance Rank |
| Visual Identity | 34.1 | 1 |
| Interaction Richness | 27.6 | 2 |
| Narrative Consistency | 22.3 | 3 Community Dynamics |
| Community Dynamics | 11.4 | 4 Residual (Unexplained) |
| Residual (Unexplained) | 4.6 | - |

**5 Digital avatar image optimization strategy**

Based on the quantitative analysis of the influence paths of multidimensional design factors, a set of dynamically adjustable digital avatar image optimization strategy framework is proposed, aiming at achieving continuous iteration and scene adaptation of communication effects. Specifically, the proposed four major design dimensions are transformed into adjustable strategy parameters, and the following optimization model is constructed with the integrated communication effect Ceff as the objective function:



Where V is the visual recognition score, I is the interaction performance, N is the narrative consistency index, O is the operation activity index, and the four dimension weights α,β,γ,δ can be dynamically adjusted according to the feedback from different platforms. In terms of strategy implementation, it is suggested to adopt the closed-loop mechanism of "version evolution + behavioral data reflow", and through different design configurations of A/B distribution, extract the indicator differences in user behavior and psychological feedback contacts (Table 6) to drive the weight update and personalized distribution. At the same time, we keep the "identification anchors" stable in the visual design (e.g., main colors, style lines), and set the interaction mechanism, narrative rhythm and operation touch points as "variable slots" to form a scene-based strategy template library (Figure 5), so as to achieve cross-platform consistency and local optimal strategy coexistence. The coexistence of cross-platform consistency and locally optimal strategies ultimately leads to stronger adaptability and growth potential of the digital avatar image in real communication.

Table 6 Strategy Optimization Indicator and Contact Mapping Table (Strategy Execution Interface)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dimension | Adjustable Parameters | User touchpoints | Data Source | Update Period |
| Visual Identity | Style slot/anchor density | First frame dwell, recognition correctness | Visual Log | Weekly |
| Interaction Depth | Response granularity/sentiment modeling | Session rounds, hit rate | Conversation logs | Day |
| Narrative Coherence | Worldview density/plot hooks | Commentary semantics, renewed viewing behavior | Content + commentary corpus | Weeks |
| Operational Mechanisms | Distribution frequency / community incentives | Second creation rate, invitation success rate | Platform return log | Weeks |

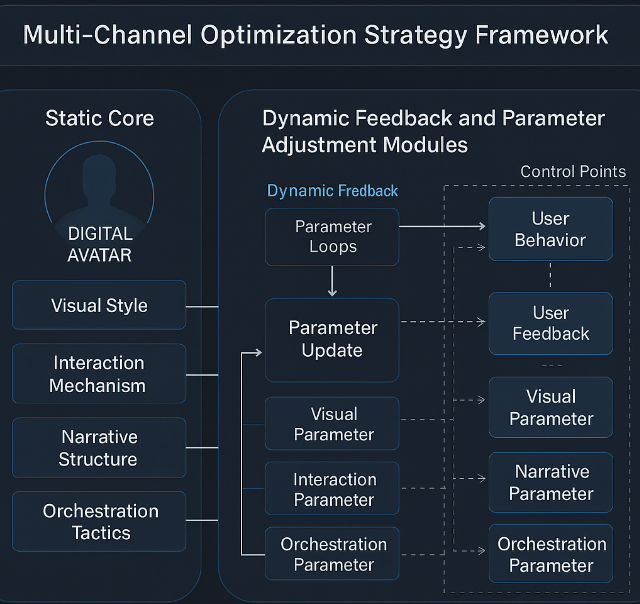


Fig. 5 Structure of multi-dimensional optimization strategy for digital avatars

**6 Conclusion**

The construction of digital avatar image presents a highly integrated design and communication logic in new media multifaceted scenarios, where visual identification, interaction mechanism, narrative structure and data operation form a systematic closed loop, which jointly support the enhancement of its communication effectiveness. Multidimensional experimental design and path modeling reveal the significant influence of design variables on psychological mechanisms and communication indicators, and verify the effectiveness and adaptability of image optimization strategies. Based on the combination of theoretical construction and empirical verification, a set of evolvable and adjustable strategy models is proposed, which enriches the research paradigm in the field of digital communication and virtual image. Due to the limitations of experimental platform and sample distribution, data extrapolation still needs to be handled with caution. In the future, it can be further expanded to multi-context and multi-cultural cross-platform adaptation research, deepen the coupling mechanism between personalized response mechanism and narrative evolution model, and provide a solid foundation for the intelligence and emotional immersion of avatar system.

**References**

1. Jin P, Liu Y. Fluid space: Digitisation of cultural heritage and its media dissemination[J]. Telematics and Informatics Reports, 2022, 8: 100022.
2. Chen J, Pan L, Zhou R, et al. Shaping and optimizing the image of virtual city spokespersons based on factor analysis and entropy weight methodology: A cross-sectional study from China[J]. Systems, 2024, 12(2): 44.
3. Ahmedien D A M. A drop of light: an interactive new media art investigation of human-technology symbiosis[J]. Humanities and Social Sciences Communications, 2024, 11(1): 1-20.
4. Li S, Chen J. Virtual human on social media: Text mining and sentiment analysis[J]. Technology in Society, 2024, 78: 102666.
5. Ciriello R, Gal U, Hannon O, et al. Responsible social media use: how user characteristics shape the actualisation of ambiguous affordances[J]. European Journal of Information Systems, 2024: 1-23.
6. Foyet M, Child B. COVID-19, social media, algorithms and the rise of indigenous movements in Southern Africa: perspectives from activists, audiences and policymakers[J]. Frontiers in Sociology, 2024, 9: 1433998.
7. Merino M, Tornero-Aguilera J F, Rubio-Zarapuz A, et al. Body perceptions and psychological well-being: a review of the impact of social media and physical measurements on self-esteem and mental health with a focus on body image satisfaction and its relationship with cultural and gender factors[C]//Healthcare. MDPI, 2024, 12(14): 1396.
8. Oyighan D, Okwu E. Social media for information dissemination in the digital era[J]. RAY: International Journal of Multidisciplinary Studies, 2024, 10(1): 1-21.
9. Rossetti G, Stella M, Cazabet R, et al. Y social: an llm-powered social media digital twin[J]. arXiv preprint arXiv:2408.00818, 2024.
10. Xiao Y, Chen R, Xiao Q, et al. Enhancing Trust or Fostering Misjudgment? Assessing the Impact of Emerging Geographic Information Displays on Social Media Users’ Information Trust[J]. International Journal of Human–Computer Interaction, 2025: 1-17.