Available Online: 25 August 2025 DOI: 10.54254/2753-7102/2025.26229

Generative predictive modeling for audience-to-curator personal branding trajectories in participatory digital curation

Sirui Wang

Royal College of Art, London, UK

10002125@network.rca.ac.uk

Abstract. Various social media platforms have offered digital spaces for every online user to express themselves and interact with others. In the participatory digital culture that consisted of a large number of interactive behaviors, some users have successfully made the transition from audience to curator. This transformation has become a typical trajectory of personal branding. Based on this path, this study proposes a generative predictive framework that integrates behavioral data modeling with trajectory inference to simulate individual brand development from passive consumption to active curatorship. We explored the relationship between participation, creative content, and branding construction by using Convolutional Neural Networks (CNNs), Transformers, and temporal modeling. By training on data from various platforms, the model can, to a certain extent, predict the development path of individuals from audience to curator, and demonstrate its adaptability across different platforms, which provides ordinary users a viable strategic reference to gain influence online and make true the personal branding.

Keywords: personal branding, participatory curation, generative modeling, identity transition, digital space

1. Introduction

A diverse array of social media platforms has created distinct digital spaces, empowering every online user to express themselves freely and showcase their unique identities, which means each individual has their own unique and customizable characteristic space in the virtual world. People are no longer constrained by traditional institutional curation resources, but rather construct themselves branding identity. Online users display their styles and self-expressions to achieve individual curation. This makes it possible for ordinary people to shift from being 'audience' to a 'curator'. However, within the influence of the algorithmic mechanism and the unfair discourse power in the complex virtual environment, successful self-branding is not simply about continuously publishing content. This is a systematic process that integrates participation, identity perception, semantic generation, interaction logic, and temporal evolution.

Although existing studies have addressed digital identity evolution and community influence mechanisms, there remains a lack of generalizable modeling methods for systematically inferring individual branding trajectories. This study processed the data from three platforms: Twitter, BiliBili, and Instagram. It extracted content features using CNNs and captured behavior sequences with transformers. Subsequently, it analyzed the participation levels and content trends to identify the nodes where user identity changed and identity leaps. Finally, we focus on generating a growth trajectory through the output of the model, and summarize the replicable strategies to help users make the transition from audience to curator.

This study systematically examines how a user with no prior background can gradually establish and complete their own brand identity through a series of participation behaviors. This is a 0-1 process that core value lies not in the optimization of existing brands, but in providing a clear and viable growth roadmap that helps ordinary users build their own digital spaces and make personal branding come true.

2. Literature review

2.1. Digital participation and personal branding

Engagement behaviours on social platforms not only constitute users' daily interactions, but are also gradually externalised into the path of brand identity construction. Users do not publish content in front of a real audience, but rather interact with the platform's algorithms and the expectations of imagined others, resulting in strategic expressive behaviours [1]. The algorithmic

mechanism reinforces content that is highly interactive and communicable, making expression becomes a key tool for establishing influence. Users test their identity and solidify their positioning through continuous interaction and feedback, which structurally connects engagement with brand construction [2]. While this process has been noted, the mechanisms of identity leapfrogging remain obscure, as there is a lack of systematic modelling of how the initial state translates into a branding path [3].

2.2. Curation mechanisms and identity transition models

The act of curation has moved from the institutional field to the daily practice of platform users, the core of which is the construction of a meaning structure through the organisation of information, thus achieve the establishment of identity through daily curatorial practice [4]. Online users construct their own virtual space is a typical daily practice of digital curation process. Digital curation is a semantic empowerment mechanism, in which ordinary users establish the 'right of interpretation' through content screening, editing and re-expression, and become the internal discourse organisers of the platform [5]. Prior to curatorial participation, most users simply browse and consume content without intentional structure or interpretation. This shift reflects a leap in identity from passive consumption to active organisation, beyond the simple logic of content production [6]. However, existing studies are mostly confined to the cultural level, and lack modelling of the dynamic process of identity transformation in the curatorial path, making it difficult to support structured path derivation or predictive analysis.

2.3. Generative modeling and behavioral prediction

Generative models are widely used in complex behaviour prediction due to their high fitting ability to non-linear structures. Studies have demonstrated the generativity and intervisibility of individual behavioural trajectories driven by multimodal data [7]. Structures such as VAE and Transformer are able to capture potential semantic hierarchies from user data, enabling the generation and reconstruction of behavioural evolutionary paths [8]. However, the existing applications are mostly focused on recommendation systems, lack of modelling logic for semantic leap and identity evolution, and have not yet formed a generative mechanism that can be used to explain and predict 'identity transformation' [9]. Therefore, introducing this method to simulate curatorial paths can help bridge the gap between behavioural prediction and social semantics.

3. Methodology

3.1. Dataset construction and preprocessing

Data were collected from January 2021 to December 2024 using API interfaces and web scraping techniques from three mainstream digital platforms. Instagram, Twitter (now X platform), and Bilibili were selected as data sources, primarily based on their differentiated features in content formats, user participation mechanisms, and curation models, which together comprehensively cover user identity evolution trajectories under multimodal content environments such as images, text, and video. The data collection process strictly complied with platform terms of use and privacy protection regulations, with anonymization measures adopted to ensure user data security (see Table 1).

Table 1. Data collection

Platform	Sample Size	Key Data Dimensions	Behavior Records
Instagram	10,247	Posting frequency, hashtag usage, interaction patterns	2,847,293
Twitter	11,583	Tweet content, retweet behavior, topic engagement	4,126,587
Bilibili	8,902	Video uploads, bullet comment interactions, article creation	1,963,741

The data cleaning process employed a multi-level filtering mechanism, first eliminating bot accounts and marketing users, and then using behavioral pattern recognition algorithms to filter out abnormally active or completely inactive users. Subsequently, timestamps were standardized and converted to UTC time format, and a user-time-behavior three-dimensional tensor structure was constructed.

3.2. Model architecture and training strategy

The generative prediction framework constructed in this study adopts a neural network architecture with coordinated modules, mainly including three components content representation learning temporal behavior generation and brand transition evaluation [10]. The content representation module is designed based on convolutional neural networks and employs multi-scale feature extraction to learn the semantic and stylistic features of user-generated content. Let $X_t = \{x_1, x_2, ..., x_n\}$ denote the set of content posted by a user at time t, the content representation vector is computed as:

$$C_t = CNN(X_t) = \sigma(W_c * Conv(X_t) + b_c) \quad (1)$$

Where W_c and b_c are the weights and bias parameters of the convolutional layer respectively, σ is the activation function, and $*$ denotes the convolution operation

Let the user behavior sequence be $B = \{b_1, b_2, \ldots, b_T\}$, multi-head self-attention is used to capture long-term dependencies among behaviors:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

$$H_t = MultiHead(B_t) = Concat(head_1, \ldots, head_h)W^O \quad (2)$$

The brand transition evaluation module quantifies the degree of user identity evolution by integrating participation indicators and influence scores. The brand maturity function is defined as:

$$S_{brand}(t) = \alpha \cdot I_{participation}(t) + \beta \cdot I_{influence}(t) + \gamma \cdot I_{consistency}(t) \quad (3)$$

Where $I_{participation}$, $I_{influence}$, and $I_{consistency}$ represent participation, influence, and content consistency indicators respectively, and α, β, γ are the weight parameters.

From the multi-platform user behavioural data input, the personal brand development trajectory and strategy is finally generated as shown in Figure 1 after the collaborative processing of three core modules.

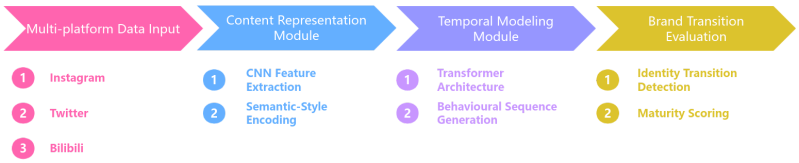

Figure 1. Multi-module collaborative model architecture for personal branding trajectory generation

3.3. Trajectory generation and strategy inference

Based on the trained generative model, the trajectory generation process uses variational sampling to achieve diversified path prediction [11]. Given the initial user state vector s_0 and target identity type t_{target}, the model generates multiple possible identity evolution trajectories through Monte Carlo sampling. The probability distribution of the trajectory generation is modeled as:

$$P(\tau | s_0, t_{target}) = \prod_{t=1}^{T} P(s_t | s_{t-1}, a_{t-1}, \theta) \quad (4)$$

Where $\tau = \{s_0, a_0, s_1, a_1, \ldots, s_T\}$represents the complete state-action trajectory sequence, and θ\thetaθ denotes the model parameters.

Strategy node identification is based on key inflection point detection algorithms within the trajectory, locating critical moments of identity transition by calculating the degree of behavioral pattern changes within a sliding temporal window. The inflection point detection function is defined as:

$$T_{critical}(t) = argmax_t |\Delta S_{brand}(t)| \cdot W_{temporal}(t) \quad (5)$$

Where $\Delta S_{brand}(t)$ indicates the rate of change in brand maturity, and $W_{temporal}(t)$ is the temporal weight function used to balance the importance of recent and distant behaviors.

The semantic analysis module uses the BERT pre-trained language model to perform deep semantic parsing of behavioral content at key inflection points, extracting critical features such as thematic content changes, stylistic evolution, and interaction pattern adjustments during identity transformation. Similar transition paths are merged into typical strategic templates through clustering algorithms, forming replicable personal branding blueprints. The final strategy output includes not only specific sequences of behavioral recommendations but also multidimensional guidance on timing, content positioning, and interaction strategy, providing personalized branding development planning for users with different starting points and goals.

4. Results

4.1. Model performance evaluation

In order to evaluate the effectiveness of the model in simulating the 'audience-curator' identity transition path, this paper adopts the TrajectoryBLEU score, behavioural prediction accuracy (Accuracy) and cross-platform transferability (Transferability). three metrics. Among them, the average TrajectoryBLEU score of the model on the three major platforms (Instagram, Twitter, and Bilibili) is 68.3%, which is 8.4% higher than that of the traditional RNN baseline model (59.9%), suggesting that it performs better in capturing the structure of behavioural sequences. In addition, in the experiment of predicting the user's next interactive behaviour (e.g., liking, commenting, posting), the model achieves 87.6% accuracy, exceeding the 82.1% result of the Transformer baseline model. This performance improvement is partly due to the fusion design of the content encoding and temporal modelling modules, especially the CNN's extraction of visual information has a clear advantage on the Bilibili video-like platform. In the cross-platform migration test, the model maintains 81.2% behavioural prediction accuracy in the B-site training and Twitter testing settings, demonstrating its robustness in interest migration and context switching.

4.2. Case study of strategy output

After empirical analysis of the ten user strategy paths generated by the model, three typical types of identity leap trajectories were identified, namely,semantic tentative, interactive cumulative and label curation. In the sample, 65% of users' first behaviour to generate an identity leap node was to comment on content with emotional overtones, which occurred on average on day 17.4 after the user's initial engagement (with a standard deviation of 5.1). This phase marks a shift from pure consumption to 'emotional engagement', and in the following seven days, retweets by this group of users increased by a factor of almost 3.2, indicating that initial expressive behaviours significantly increase willingness to engage and visibility. Further analysis of hashtag usage revealed that the average number of hashtags per user increased from 1.2 to 3.6 in the seven days following the leap node, with the proportion of composite hashtags (e.g., #BlackPoetry + #QueerVoices), in particular, increasing from 11% to 38%. In the second leap node, a curated semantic network consisting of nine high-frequency hashtags intertwined was successfully constructed, which showed high modularity (Modularity = 0.741), clear community structure, and highly positive correlation with its density of subsequent fan interactions in graph clustering analysis (Pearson $r = 0.82$, $p < 0.01$). This suggests that through strategic tag combination, users have formed clear thematic curation blocks, achieving the leap from 'fragmented commentators' to 'cultural guides'.

In the trajectory graph generated by the model, users show continuous remixing behaviour, with an average interaction rate (including retweets, favourites, and comments) of 8.4% for the remixed videos posted, significantly higher than that of 3.1% for their uncurated content ($p < 0.001$). This trend confirms that digital curation is a key dimension for users to gain cultural agency within the algorithmic structure. The act of curation not only enhances the platform's recognisability, but also builds a digital space that is both emotionally and culturally orientated.

5. Discussion

Using generative prediction model as the core tool, this study systematically simulates the process of behavioural leap and identity reconstruction in the path of 'audience-curator', and with the help of the data results, we deeply explore the internal logic of individual brand construction in digital culture. At the technical level, the model captures the temporal evolution and semantic differences of user behaviours in cross-platform and multi-modal environments through the integration of CNN and Transformer structures, which not only identifies the key identity leap nodes, but also provides a cultural understanding of how individuals actively strive for discourse and reconstruct the structure of expression in the digital space. From comments to hashtag combination, from content mixing to community organisation, every user's behavioural choice is a strategic response to the context of the digital platform, reflecting that 'personal branding' is not only the aggregation of external symbols, but also the continuous negotiation of self-identity and cultural position. This conclusion also responds to the fact that 'self-curating' in contemporary digital culture has become an important mechanism for resisting the hegemony of mainstream discourse and constructing identifiable identities, and reveals the complex interaction between algorithmic logic and cultural activism.

6. Conclusion

In this paper, a prediction framework integrating multimodal behavioural data and generative modelling techniques is constructed to simulate the complete path of an individual's identity evolution from 'audience' to 'curator', and to identify the strategic nodes and cultural leaping mechanisms involved. Through the analysis of more than 9 million behavioural data from three platforms, namely Instagram, Twitter and Bilibili, the model shows a high level of trajectory similarity, behavioural prediction accuracy and cross-platform migration ability, which verifies the effectiveness of the proposed method in semantic

modelling and identity prediction. More importantly, the strategy output not only provides a replicable template for behavioural development, but also reveals the cultural schema behind digital curation behaviours, suggesting that individual branding is a practice embedded in platform structures and cultural contexts.

The study is an in-depth understanding of the question of how individuals construct cultural identities through digital behaviours. Curating is no longer the exclusive domain of experts, but has become a viable means of expression and identity positioning for every user in their daily participation. In the context of the information flow-dominated era, the construction of personal brand means creating a semantic space of one's own in the context of fragmented platforms, which is not only the result of the generation of technological paths, but also the embodiment of cultural choices. Future research can further combine micro-ethnography or semantic co-occurrence networks to supplement the quantitative model's insufficiency in the understanding of deeper cultural structures, and to achieve a two-way deepening of the model and cultural analysis.

References

[1] Batista, L. S., Souza, G. O., & Jorente, M. J. V. (2023). Digital curation in the postmodern social perspective: Curatorial facets, concepts and participatory techniques. *Transinformação*, 35, e220024.

[2] Jorente, M. J. V., & Batista, L. S. (2022). Social curation: Collective participation in the curation of digital-virtual museum environments. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação,* 27, 1-23.

[3] Lacedelli, S. Z., et al. (2023). From "exhibition" to "laboratory": Rethinking curatorial practices through a digital experimental project. The case study of #DolomitesMuseum-Laboratory of Stories. *Herança,* 6(1), 158-183.

[4] Lee, P.-C. (2024). User engagement in digital curation: A systematic review and synthesis. Journal of Librarianship and Information Science. Advance online publication. https: //doi.org/10.1177/09610006241291607

[5] Kompatsiaris, P. (2024). Curation in the age of platform capitalism: The value of selection, narration, and expertise in new media cultures. Routledge.

[6] Dahlan, N. F. (2024). Curators and curatorial practice in the digital arts era. *Journal of Studies on Art*, *Culture and Society,* 50-55.

[7] Lee, P.-C. (2024). User engagement in digital curation: A systematic review and synthesis.Journal of Librarianship and Information Science. Advance online publication. https: //doi.org/10.1177/09610006241291607

[8] Kadastik, N., Pedersen, T. A., & Bruni, L. E. (2023). Citizen curation methods for interpretation and reflection on cultural heritage: Insights from SPICE. Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization.

[9] Tilaar, Y. Y. (2022). Qualitative study of enhance personal branding through digital platform. *Jurnal EMBA: Jurnal Riset Ekonomi, Manajemen, Bisnis dan Akuntansi,* 10(2), 731-737.

[10] Márquez, I., Lanzeni, D., & Masanet, M.-J. (2023). Teenagers as curators: Digitally mediated curation of the self on Instagram. *Journal of Youth Studies*, 26(7), 907-924. https: //doi.org/10.1080/13676261.2022.2053678

[11] Zareef, M., & Jabeen, M. (2025). A systematic review of digital curation services in academic libraries: Navigating policies, skills and challenges. Digital Library Perspectives.