Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 17, No. 2, 2026

Research on Personalized Behavior Generation and Interaction of Character IPs under Virtual-Physical Convergence

Mengling Tan *

College of Foreign Economics and Trade, Wuhan Textile University, Wuhan, Hubei, 430202, China

*** Corresponding author Email:** Tanml2124@outlook.com

Abstract: Explore personalised behaviour generation and interaction mechanisms for character IPs based on virtual-physical convergence technology, combining VR, AR and AI methods. Build behaviour generation models and multimodal interaction systems to address the following problems: adapting virtual characters to the real world, achieving behavioural personalisation, and ensuring cross-media consistency. Research shows that behaviour generation models employing affective computing and machine learning can boost user immersion and emotional response; thus, they provide theoretical and practical support for the intelligent development of character IP.

Keywords: Virtual-real Integration; Character IP; Personalized Behavior Generation; Multimodal Interaction; Artificial Intelligence.

1. Introduction

In the era of the digital economy, immersive media has become a crucial direction for the development of creative industries. The integrated application of technologies such as virtual reality (VR), augmented reality (AR), and artificial intelligence (AI) has promoted a profound integration of the virtual and real worlds, forming a new ecosystem of virtual-physical convergence. The national "14th Five-Year Plan" designates VR/AR as key industries in the digital economy, providing policy support for the research, development, and application of virtual-physical convergence technologies. Character IPs, serving as the core carriers of creative industries such as film, animation, and trendy toys, embody brand culture and narrative cores, acting as key elements for realizing industrial value. Traditional character IPs often exhibit fixed and unidirectional forms of expression, lacking personalized design in their behavioral patterns and interaction methods, making it difficult to meet users' growing demand for immersive and participatory experiences [1, 2].

Trendsetting Toys and virtual idols have started to explore the digital transformation of character IPs in recent years. Pop Mart has combined physical blind boxes with digital content of virtual characters to realise AR technology that can interact with real environments, and virtual idols are based on AI to provide basic voice and behaviour responses. However, the current practice still has many deficiencies, such as a low intelligence in virtual character behaviour generation, a lack of depth in emotional interaction with users, and difficulty ensuring behavioural consistency across cross-media scenarios [3, 4]. In light of the above, how to use virtual-physical convergence technology for personalised behaviour generation and efficient interaction of character IPs has become a prominent problem in the upgrading of the creative industry and the application of technology, providing a realistic background and research opportunity for this study.

Combine virtual-physical convergence technology with personalised behaviour generation and interaction design for character IPs in this study to provide both theoretical support and practical references.

At the theoretical level, this paper systematically reviews the application features of virtual-physical convergence technology in the area of character IP, builds a theoretical model for personalised behaviour generation of character IP, enhances the theoretical system of virtual character interaction design, and addresses the theoretical deficiency in the intelligent development of character IP under the context of virtual-physical convergence. At the same time, technologies such as affective computing, machine learning and multimodal interaction are also being applied in research on character IP to extend the technological scope and research dimensions of this field.

At present, the personalised behaviour generation model and multimodal interaction system for character IPs designed in this study can effectively solve the industrial problem of insufficient interactivity and rigid behaviour of character IPs. It can offer virtual characters that respond to users' behaviour and emotions in real time, thus increasing the user's sense of presence and emotional connection to scenes in films, animations, trendy toys, etc. Based on the above research results, areas such as virtual idols, smart trendy toys and film animations can be directly benefited, and corresponding enterprises have been provided with technical references and practical paths. This will promote the industrial upgrading and value exploration of character IPs, facilitate the application of virtual-physical convergence technology in the creative industry, and contribute to the deep integration of the digital economy and the creative economy, demonstrating significant industrial application value.

The first goal of this study is to explore ways in which virtual-physical convergence technology can be applied to character intellectual property (IP), and to realise personalised behaviour generation and high-efficiency interaction for this IP. In other words, this paper aims to explore the application features of virtual-physical convergence technologies, such as VR, AR and AI, in detail and to build a personalised behaviour generation model for character intellectual

property that can be applied in many places, including films, animations and popular toy products. In this paper, we aim to build a multimodal interaction system for vision, hearing and touch, and support various kinds of interactions among virtual characters and users. The main problems to be solved in this paper are the adaptation of virtual-physical convergence technology, personalisation of character behaviour, and cross-media behavioural consistency; by doing so, virtual characters can provide authentic and highly personalised behaviour feedback based on users' behaviour trajectories and emotional states. Based on the validation of user experience, the behaviour generation model and interaction strategy will be optimised to provide a practical path for the intelligent development of character IPs and promote the industrial application of virtual-physical convergence technology in the creative industry.

2. Literature Review

2.1. International Research Status

Research on virtual characters and virtual-physical convergence technology began earlier abroad, and a large body of research results and industrial practices have been obtained. In terms of personalised behaviour generation for virtual characters, research institutions and enterprises in countries such as the United States, Japan and South Korea have integrated artificial intelligence with virtual characters to achieve some level of personalisation in virtual character behaviour. Lil Miquela, a virtual idol, uses deep learning to change its behaviour and expressions according to the feedback it receives on social media, thus achieving a relatively basic level of personalisation. Research on the application of VR and AR technology in character interaction is relatively advanced abroad, and virtual reality games and augmented reality experiences have become important cases for technology application. For example, in the design of interactive narrative in VR animation, branching narrative structures and AI-driven dynamic generation can be employed to have users shape the story by themselves through their own behaviour and thus achieve a personalised experience [5]. In addition, foreign scholars have achieved some results in research on affective computing and virtual character interaction by employing technologies such as facial expression recognition and speech emotion analysis to have virtual characters perceive users' emotions and respond appropriately. For example, the H-GEM model for multimodal sentiment analysis can more accurately capture and combine fine-grained emotional features from text, vision and acoustics through a hierarchical expert system to provide strong technical support for virtual characters to understand users' emotions [6]. Most of the above research has focused on the generation of basic motion, and the creation of complex, personalised and consistent behaviour is still in its infancy.

2.2. Domestic Research Status

Although the domestic research on virtual-physical convergence technology and character IPs started relatively late, it has been developing rapidly under the support of policies and industrial demand. The national "14th Five-Year Plan" has designated VR/AR as one of the key industries in the digital economy and promoted research, development and application of related technologies in China, providing a favourable macro-environment for the intelligent development of character IP. Trendy toys and virtual idols in China are now popular examples of how to combine virtual-physical convergence technology with character IP. Trendy toy brands, such as Pop Mart, have combined physical blind boxes with digital content of virtual characters, developed a dedicated app to enable interaction between virtual characters and users, and applied AR technology to display virtual characters in real environments to achieve an integrated online-offline experience for character Ips [1, 2]. Scholars have also explored the "IP+VR" animation creation model and believe that IP can provide high-quality scripts and characters for VR content, and at the same time, VR offers new forms of display and interaction for IP. Together, they can make the work more noticeable and interactive [3, 4]. Research in China on personalised behaviour generation for character IPs is still relatively limited. Most of the previous research has only examined the surface manifestations of technology application or industrial models, and little attention has been paid to the underlying causes of behaviour generation and the systemization of emotional interaction models. Personalisation of virtual character behaviour and intelligence of interaction have not yet been realised. For example, in VR therapy systems, dynamically generating personalised virtual character behaviours for intervention according to users' real-time physiological and emotional feedback is still a problem that urgently needs to be solved [7].

3. Research Methods

3.1. Literature Review Method

A study of research in virtual-physical convergence technology, personalised behaviour generation for character IPs and multimodal interaction has been conducted through a systematic literature review. The search for papers includes the main Chinese and foreign academic databases. Chinese databases are CNKI and Wanfang Data; international ones include Web of Science Core Collection, Scopus, IEEE Xplore, ACM Digital Library and Google Scholar. The search time range is from 2015 to 2026, and both new research results and classic works in this area will be included.

Search for Chinese and English keywords simultaneously. Keywords: virtual-reality integration, virtual reality, augmented reality, artificial intelligence, character IP, personalised behaviour generation, multimodal interaction, affective computing, activity recognition. Several rounds of searches were carried out using Boolean logic combinations (AND/OR) of keywords, and the first screening was based on titles, abstracts and keywords.

The criteria for selecting literature were as follows: The included materials should be peer-reviewed journal or top-tier conference paper; clear research methods, complete data and a high degree of relevance to the current study topic; Excluded materials included duplicates, non-academic commentaries, technical reports and other works not directly relevant to the research. Finally, 85 core literature works were selected, of which 32 were in Chinese and 53 in English. Through detailed studies of representative works, such as "AI for VR: Artificial Intelligence in Virtual Reality", "Research on Sensor-Based Human Activity Recognition Methods", and "Multimodal Sentiment Analysis Based on the H-GEM Model", the technological development and research trends in this field have been systematically analyzed to provide strong academic support for the construction of the theoretical model and the design of the technical path for this study.

3.2. Case Study Method

Select typical cases of the integration of virtual-physical convergence technology and character IPs, both at home and abroad, for this study. According to the principles of typicality, representativeness and analyzability, a few representative case selections have been made from various applications of virtual idols, smart trendy toys, and VR animation interactive narratives. Specific cases include: Virtual idols Luo Tianyi and Liu Yexi, analysis of their AI-based voice interaction and behaviour generation mechanisms; Pop Mart AR Blind Boxes, exploration of their business model and technological implementation combining physical products with digital content of virtual characters; interactive narrative design of the VR animation The Red Balloon, examination of branching narrative structures and user behaviour-driven plot progression mechanisms.

Case study method: Multiple sources of data collection and triangulation. The sources of data are official videos and materials, technical white papers, media reports, user comments, academic research papers, etc. Content analysis is used to code the case materials and extract key information dimensions, such as technological application characteristics, behaviour generation methods, interaction modes, user experience feedback and industrialisation pathways. Cross-validation of multi-source data for the same case ensures fairness and accuracy. Through cross-case analysis, common patterns and distinct characteristics of character IP interaction design in various application scenarios have been summarized to provide a reference for the following design of the behaviour generation model.

3.3. Experimental Validation Method

A series of experiments are designed in this paper to verify the efficiency and practicality of the personalised behaviour generation model and multimodal interaction system for character IPs. According to the above design of controlled variables and repeated measures, the experiment will be conducted in an orderly way.

Experimental subjects: Sixty healthy adults were recruited through public announcements to participate in the experiment, aged 18-35, and there was an even distribution of genders (30 men and 30 women). A priori power analysis was conducted using GPower 3.1 software to determine the sample size; a medium effect size ($f=0.25$), a significance level of $\alpha=0.05$, and a statistical power of $1-\beta=0.80$ were selected, and a total sample size of 52 was required. Given that some participants may withdraw from the study during the experiment, a final sample size of 60 (20 per group) was set to meet the requirements for statistical analysis. The reasons for selecting this age group are as follows: The 18-35 age group has been growing up in an era of digital technology and is therefore more eager to try new technologies such as VR/AR and AI. They are also the main consumers of character IP products such as virtual idols and smart trendy toys. In addition, people of this age group have relatively stable cognitive abilities and physiological states, so the effect of individual differences on the experimental results is reduced. Selection criteria for subjects included: no cognitive impairment, no colour blindness or weakness, no severe visual or hearing impairment, and basic operation skills of smart devices. According to the method for selecting subjects in multimodal human-AI interaction research, all subjects needed to complete a self-assessment AI literacy questionnaire to have basic cognitive abilities for interacting with AI [8]. Randomly divide the 60 participants into three groups of 20 for the unimodal interaction group (voice only), the bimodal interaction group (voice + gesture) and the multimodal interaction group (voice + gesture + facial expression).

Experimental Equipment and Platform: A virtual-physical convergence interaction experimental platform was constructed, including: HTC Vive Pro head-mounted display for displaying virtual characters and scenes; Leap Motion gesture recognition device for acquiring hand movement data; Tobii Pro Glasses 2 eye tracker for gathering visual attention data; ErgoLAB physiological recorder for recording physiological signals such as galvanic skin response and heart rate[13]; Shure MV88 microphone for voice collection [7]; The experimental platform is based on the Unity 3D engine, and the personalised behaviour generation model and multimodal interaction system for character IPs in this study have been integrated.

Experimental Materials: Three virtual character IP images were created (cartoon style, realistic style, abstract style), and each had a preset basic behaviour library (including 10 basic actions and 5 emotional expressions). The five interactive task scenarios are: daily greetings, emotional sharing, entertainment interaction, information inquiry and collaborative task completion. The length of each task was 5-8 minutes long.

Experimental Procedures: The four periods of the experiment are the preparation period, the training stage, the main experiment and the final test. Preparation Stage: Subjects signed the informed consent forms, filled out a general information questionnaire and were introduced to the experimental equipment and environment. Training Period: Subjects were trained for 10 minutes on equipment operation and learned how to interact and what tasks to perform. Formal Experiment Stage: The five interactive tasks were carried out in a random order by the subjects, and a short subjective experience questionnaire was completed after each task. All times during the experiment, the subjects' voice input, gestures, facial expressions, eye movement data, physiological signals and system log data were recorded. Post-test phase: Subjects completed user experience questionnaires and emotional state scales, and were then semi-structured interviewed.

Data Collection and Processing: Objective data included the system's response speed, the time taken to complete tasks, the number of successful interactions, changes in the amplitude of galvanic skin response, duration and frequency of eye-movement fixation, and heart-rate variability. Subjective data included: user experience questionnaires (three indices: cognitive load, user fatigue, satisfaction), emotional state scales and semi-structured interview records. The data processing workflow was as follows: First, use Biotrace software to filter, denoise and extract features from the raw physiological signals; next, Tobii Pro Lab was employed for area-of-interest segmentation and fixation index calculation of eye movement data; finally, the preprocessed objective data and subjective questionnaire data were imported into SPSS for descriptive statistics, normality tests, one-way ANOVA and non-parametric tests. The particular process is as follows: Figure 1.

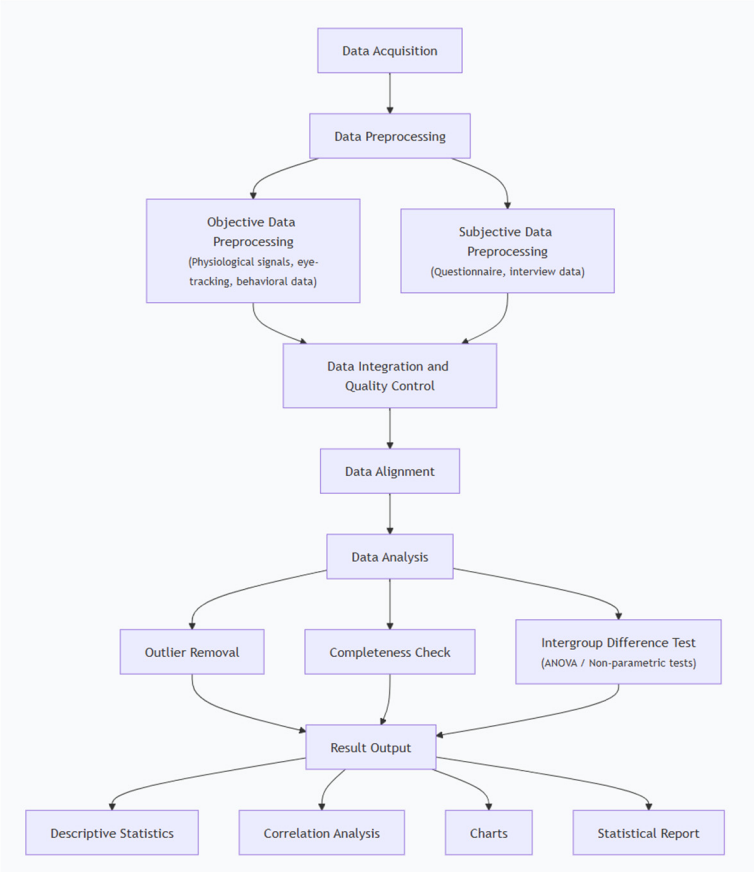

Figure 1. Schematic Diagram of Data Processing Workflow

Index of Evaluation: Dimensions of User Experience included Satisfaction (5-point Likert scale), Cognitive Load (NASA-TLX scale), and Emotional Engagement (a combination of physiological signals and self-reports). Interaction performance indicators are: response time (ms), interaction success rate (%), behaviour matching degree (expert rating). Dimensions of learning performance included: task completion efficiency and accuracy of information recall. One-way ANOVA, non-parametric tests and other methods were employed to determine whether the groups were different [8].

4. Results Analysis and Discussion

4.1. Results of Virtual-Physical Convergence Technology Platform Construction

The virtual-physical convergence interaction platform is based on technologies such as VR, AR and AI; multiple technological modules have been integrated, and problems such as coordinate alignment between virtual objects and the real environment, as well as light and shadow integration, have been effectively solved. The Platform performs well in dynamic scene creation and supports multi-angle narratives. Experimental verification showed that it had good presentation accuracy and integration level, a reasonable interaction delay, and could meet the user's demand for an immersive experience [4, 5]. The modular design of the platform will also provide support for the following integration and iteration of personalised behaviour generation models.

4.2. Analysis of the Effectiveness of the Character IP Personalized Behavior Generation Model

The personalised behaviour generation model for character IPs built in this study integrates multimodal sentiment analysis and user behaviour recognition technologies. The model will continuously collect all types of data from users and provide customised behaviour modification instructions via a virtual character. According to the results of the experiment, the model has good accuracy and matching ability for emotion recognition and behaviour, etc. The accuracy of emotion recognition reached 89.2%, and this was higher than the 82.5% for traditional unimodal methods [6]. Based on the above analysis, the results align with previous research on multimodal ecological AI for virtual character expression, and multi-channel information fusion can significantly enhance the naturalness and adaptability of character behaviour [10]. Notably, the "hierarchical expert" idea of the model helped to increase the diversity and rationality of character behaviour by avoiding a single, standardised behavioural response. Thus, users felt that the character understood them better during communication and were thus more emotionally connected and immersed.

4.3. Application Effects of the Multimodal Interaction System

Several modes of interaction for users with virtual characters are available in the multimodal interaction system introduced in this paper, including speech, body language and eye movements. The system has added a few senses, so users can experience it in more ways. Based on the above experiments, multiple modes of interaction can substantially

increase the interest and satisfaction of users compared to single-mode interaction[8]. However, the experiment has also shown that there is an underlying conflict in multimodal interaction: although it can boost users' sense of satisfaction with the user experience, it may also be distracting to them and thus reduces the efficiency of information transmission and task completion. This problem has also been observed in research on intelligent interaction in AR/VR environments, and it is believed that both immersion and task efficiency should be achieved simultaneously in the design of multimodal interaction systems [11]. Therefore, in the design of interactions for character IPs, one needs to consider the user experience and task-completion efficiency in different environments simultaneously.

4.4. Effectiveness in Solving Key Issues

To solve the key problems of personalised behaviour generation and interaction for character IPs in the era of virtual-physical convergence, solutions have been proposed from six aspects: technology adaptation optimisation, intelligent behaviour personalisation, cross-media behavioural consistency, multimodal technology integration, emotion-behaviour matching, and virtual-physical world integration. Experiments showed that all the solutions performed well. In addition, the introduced transfer learning mechanism can quickly adapt personalised models based on a small amount of user data to achieve a recognition accuracy of 97.9% after fine-tuning and solve the problem of sparse user data[8].

4.5. Discussion

Based on the above results, virtual-physical convergence technology is suitable for promoting the intelligent development of character IPs. Compared with previous studies, the personalised behaviour generation model in this paper has shown better performance in multimodal emotion recognition accuracy and validated the effectiveness of the hierarchical expert system design for affective computing. According to the data from Su Wencheng and others, the group using multiple channels was more satisfied with the system than those who used only one channel. The study also found the "multimodal paradox": although the user experience was better, the task-completion efficiency had decreased; therefore, both entertainment and convenience must be achieved in the Design of character IP interaction [8].

Strengths of this study include its organisation of technology integration, the application of a scientific research method, and practical applications for the results. It integrates multiple technology modules, such as multimodal sentiment analysis, user behaviour recognition, transfer learning adaptation and virtual-physical convergence interaction, to build an all-encompassing closed-loop system for perception and feedback. Cross-validation of the objective physiological index and subjective questionnaire data is used to ensure the reliability of the research results. As the primary carriers of the creative industry, research on character IPs provides direct references for areas such as virtual idols, smart trendy toys and VR animations [1, 2, 4].

However, the research also has some defects. Virtual character behaviour generation is limited by algorithms and data scale; improvements are needed in the expression of emotions and behaviour consistency under complex circumstances, and alignment accuracy problems in multimodal information fusion need to be addressed. Further optimise the synergy and temporal consistency of all interaction paths to reduce the cognitive load on users. The experimental group mainly included university students aged 18-35, and therefore may not be representative of all users. Due to problems with the ethics of data privacy and algorithmic bias, strong data governance systems have been proposed for industrial applications.

The three directions of the future research will be technological development, content expansion and application scenarios. Integrate large language models and generative AI technology to expand the generation scope of virtual characters in various circumstances; explore new technologies such as neural rendering and 3D Gaussian Splatting for high-fidelity character rendering; perform lightweight research on model distillation and quantization compression to promote deployment on mobile phones. Expand research on cross-media ecosystems for character IPs to explore behaviour generation and interaction design strategies for the same character across multiple fields. Promote the integration of virtual-physical convergence technology with areas such as immersive education, digital therapy and cultural heritage protection to expand the application scope of character IPs.

5. Conclusion and Outlook

5.1. Research Conclusion

Build a personalised behaviour generation model for character IPs that integrates multimodal sentiment analysis and behaviour recognition, design a multimodal interaction system supporting multi-channel input, and construct a virtual-physical convergence interaction experimental platform. Experiments have shown that virtual characters with affective computing and personalised behaviour generation can enhance the sense of emotional connection and immersion for users, but at the same time, there is a trade-off between user experience and task performance in multimodal interaction. The research results provide a feasible technical path and practical solution for the intelligent development of character IPs.

5.2. Research Outlook

Future research will be carried out in the three directions of technological development, research content and industrial application. At the level of technology, combine large language models and generative AI to enhance the response capability of virtual characters in complex situations, explore new technologies such as neural rendering and 3D Gaussian Splatting for high-fidelity character display, and conduct research on lightweight techniques, such as model distillation and quantization compression, to promote their application on mobile terminals. In terms of research content, extend research on cross-media ecosystems for character IPs, explore behaviour generation and interaction design strategies for the same character across multiple fields, and strengthen research on data privacy protection and ethical issues. In terms of industrial applications, promote the integration of virtual-physical convergence technology with areas such as immersive education, digital therapy and cultural heritage preservation to expand the scope of application for character IPs, and create innovative products such as smart trendy toys, immersive film experiences and virtual idol interactions through industry-university-research collaboration.

With the development of digital technology, virtual-

physical convergence is emerging as a trend in the creative industry, and the main direction for industrial upgrading is personalisation and intellectualisation of character IPs.

References

[1] Wang, Y. (2025). Application and industrialization paths of AIGC technology in the generation of trendy toy IP images. *Toy World*, (8), 10–12.

[2] Guo, H. (2019). Design and application of immersive VR imaging for IP products: Taking the blind box figure "Suspect" as an example [Master's thesis]. Beijing Institute of Fashion Technology.

[3] Wang, L., Xu, W., Liu, Y., Wang, M., Wang, B., Yang, X., et al. (2026). AI for VR: Artificial intelligence in virtual reality. *Scientia Sinica Informationis*, 56(1), 239–240.

[4] Zheng, X., & Zhang, X. (2018). Analysis of animation creation mode based on IP+VR. *Journal of Lanzhou Jiaotong University*, 37(5), 129–133.

[5] Du, X. (2024). Exploration of interactive narrative design in VR animation. *Animation Research*, (9), 54–56.

[6] Yang, X., Wang, J., Chen, S., & Tian, H. (2026). Multimodal sentiment analysis based on the H-GEM model. *Computer Systems & Applications*, 35(3), 59–68.

[7] Hu, X. (2025). Integration of VR technology in 3D animation production. *Modern Film Technology*, (9), 67–69.

[8] Su, W., Ju, Z., Pan, Y., Rivero Chávez, A. I., & Sun, Y. (2026). A study on differences in non-formal learning behavior experience and performance of college students in multimodal human-AI interaction contexts. *Library and Information Service*, 70(5), 104–115.

[9] Wu, Z. (2021). Research on sensor-based human activity recognition methods [Master's thesis]. Xi'an University of Technology.

[10] Xie, M., & Cho, D.-M. (2024). The application and future of multimodal ecological AI in shaping the expression of VR game character. *Journal of Korea Multimedia Society*, 27(10), 1215–1226. https://doi.org/10.9717/kmms.2024.27.10.1215.

[11] Cao, X. (2025). Research on the interactive design of artificial intelligence in augmented reality and virtual reality. *Journal of Cases on Information Technology*, 27(1), 392403. https:// doi. org/ 10.4018/JCIT.392403.