

AI Model for Digital Media Art Creation and Performance Evaluation [†]

Cheng-Cheng Zhao and Chen-Chiou Chiu * 

Department of Digital Media Arts, Dongguan City University, No. 1 Wenchang Road, Liaobu Town, Dongguan 523419, China; zhaochengcheng199811@gmail.com

* Correspondence: g7190313@gmail.com

[†] Presented at the 9th Eurasian Conference on Educational Innovation 2026 (ECEI 2026), Da Nang City, Vietnam, 30 January–2 February 2026.

Abstract

We developed an AI engine for digital media art creation by integrating optical flow in-betweening, semantic vector coloring, and cloud-based asset synthesis. In a quasi-experimental deployment ($N = 87$), the developed engine enhanced rendering efficiency by 15.4% and task scores ($p < 0.001$, $d = 1.26$) compared with a baseline model. The engine combines convolutional-transformer models for frame interpolation, color propagation, and performance scoring. It processes raster sketches and motion vectors to generate color-consistent vectors. It reduced rendering latency by 38.4% and achieved a structural similarity index measure of 0.91. The result of the analytic hierarchy process showed that operability and processing speed are important performance evaluation metrics of the developed engine.

Keywords: generative AI; systems optimization; analytic hierarchy process; multimedia engineering; software development framework

1. Introduction

Two-dimensional animation production is computationally intensive when using digital media software, and the system design and delivery remain a persistent challenge. Conventional lecture-based methods cause inherent structural latency in learning workflows by allocating excessive runtime to passive explanation, leaving limited opportunities for hands-on debugging, technical pipeline refinement, and interactive system-level problem solving. In skills-intensive animation creation, this imbalance reduces developer engagement and weakens the integration of automated technical procedures with human-in-the-loop creative decision-making [1]. As modern animation creation processes increasingly incorporate cloud-based systems and automated modules, the need for robust, software-engineered instructional architectures becomes urgent.

The flipped classroom model reorganizes the learning environment by enabling lectures and content delivery to be conducted in pre-class, asynchronous sessions. This enables students to review materials and prepare assets before class. As a result, classroom time is used for hands-on project work, debugging, and collaborative problem-solving. This improves efficiency and reduces mental workload because information delivery and practical application are separated into distinct stages [2]. Flipped learning increases motivation and participation in fields such as engineering, computing, and digital design [3,4]. However, there is limited evidence on how this model performs when combined with AI-based animation creation systems, highlighting the need for further empirical validation.



Academic Editors: Teen-Hang Meen, Chun-Yen Chang and Cheng-Fu Yang

Published: 9 October 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Meanwhile, the rapid evolution of generative AI models, neural rendering frameworks, and computer-aided design (CAD) engines is reshaping animation workflows. Automated functions such as optical flow-based in-betweening, semantic intelligent coloring, and AI-driven asset generation reduce low-level execution barriers and expand opportunities for rapid prototyping and creative experimentation. The value of these automated modules, however, depends on their structured integration into unified instructional software architectures [5]. Their performance must be assessed not as isolated utilities but as interoperable components within a cohesive system.

In this study, we designed an AI engine for digital media art creation. Its architecture comprises three modules: an asynchronous AI conceptualization engine, a real-time neural rendering and interpolation node, and an automated performance evaluation module. The engine's runtime execution speed, structural feature fidelity, and system efficiency were evaluated in processing complex two-dimensional animation [6].

2. Literature Review

Modern 2D animation creation is transitioning from isolated, localized software instances to decoupled, distributed system architectures [5,7]. A robust animation rendering and production pipeline can be modeled as a multi-stage data processing graph consisting of asset ingestion, vector manipulation, keyframe interpolation, semantic color mapping, and compilation rendering. Traditional sequential architectures suffer from high compute latency and human-operator idle time due to synchronous processing bottlenecks [7].

To optimize system throughput, contemporary CAD environments leverage edge-cloud systems that distribute rendering queues and asset dependencies [6]. In a production engineering context, the structural stability of the pipeline depends on optimizing data flow between automated processing nodes and human-in-the-loop validation checkpoints [5]. Resolving latency in these interconnected software subsystems represents a core challenge in multimedia engineering [8].

The integration of generative artificial intelligence into animation software toolchains has introduced automated neural network layers acting as pipeline middleware. These automated AI nodes replace heavy, manual compute tasks with highly optimized inference models [9]. Specifically, convolutional and transformer-based architectures power.

- Motion-vector interpolation nodes: Optical flow networks process adjacent raster frames to compute non-linear velocity fields, reducing manual keyframe drafting [8];
- Semantic color propagation networks: Conditional generative adversarial architectures maintain temporal color consistency across sequential frames [9];
- Generative layout sub-routines: Diffusion-based modules generate scalable vector assets within the digital media workspace [7].

When integrated as structural scaffolds, these automated modules significantly reduce technical compilation barriers for low-level tasks, allowing human developers to reallocate cognitive cycles toward high-level system architecture and quality assurance testing. However, without clear pipeline boundaries and parameter constraints, over-reliance on automated subroutines can introduce semantic errors and degrade critical troubleshooting workflows within the development environment.

The flipped classroom model ensures a structural reconfiguration of standard learning management architectures [10,11]. The decoupled system enables basic data ingestion to an asynchronous pre-class processing node, maximizing the active runtime of synchronous laboratory infrastructure for cooperative debugging, live pipeline deployment, and iterative peer reviews [12]. The efficiency of this architectural split relies on cognitive load management principles, which isolate initial knowledge mapping from live, high-stress system troubleshooting [12]. For vocational training streams, the operational reliability of

this model depends heavily on whether the pre-class asset delivery network is sufficiently robust, modular, and optimized to support independent operator preparation [11].

Applying advanced software engineering frameworks to digital media arts education aligns directly with the output-oriented, iterative compilation cycles found in industrial studios [7]. Empirical software tracking shows that a system integrating flipped architectures with automated AI toolchains significantly expands active laboratory runtime while accelerating localized production speeds [13]. This dual integration directly addresses the distinct learning needs of vocational students by lowering entry barriers without sacrificing technical rigor [1,11].

However, asynchronous delivery nodes with automated generative AI layers ensure complex system dependencies [8]. In the media art creation, risks exist, including increased pre-class cognitive overhead and a potential decline in independent algorithmic thinking among novice operators [12]. To validate the stability and impact of the engine, a multi-dimensional system matrix is used. Mixed-methods validation is necessary to combine quantitative metrics (such as compilation throughput and usability scales) with user logs [14]. This method provides the statistical rigor and granular interaction insights required to confirm overall system optimization [7].

3. Methodology

The media creation AI engine adopts a two-stage neural network that combines feature extraction and spatial-temporal interpolation (Figure 1). In feature extraction, a modified U-Net architecture with residual attention blocks was used to capture high-frequency line art details. The interpolation module applies a bidirectional optical flow estimator to predict motion vectors between sparse frames (F_t, F_{t+1}). The loss function balances structural fidelity and perceptual consistency:

$$\mathcal{L}_{\text{total}} = \alpha \mathcal{L}_{\text{pixel}} + \beta \mathcal{L}_{\text{perceptual}} + \gamma \mathcal{L}_{\text{smooth}} \quad (1)$$

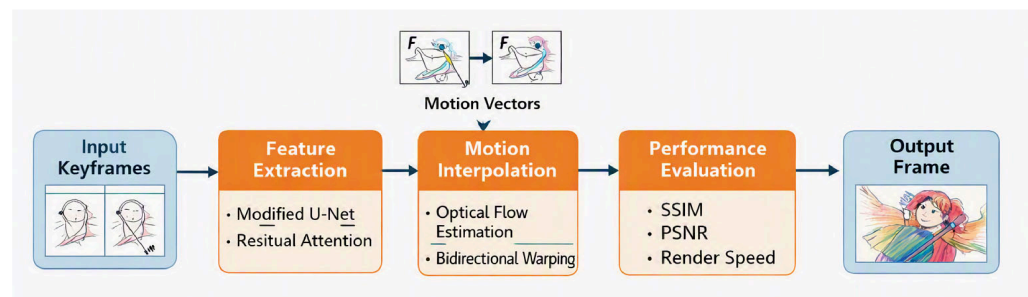


Figure 1. Simplified architecture of the AI engine for media art creation (created in this study using AI).

Here, $\mathcal{L}_{\text{pixel}}$ computes L_1 distances across raster outputs, $\mathcal{L}_{\text{perceptual}}$ computes visual geometry group-16 feature representations, and $\mathcal{L}_{\text{smooth}}$ constrains motion field variances. The performance evaluation module calculates frame rendering speed, structural similarity index measure (SSIM), and peak signal-to-noise ratio (PSNR) to assess output art quality.

To evaluate the developed framework, 87 students (38 males, 49 females, seniors) enrolled in the Digital Media Arts program at a vocational college in Guangdong, China, were recruited. None of the participants had previous training in animation creation, asset compilation, or rendering workflows. Students were divided into two groups: an experimental group ($n = 42$) utilizing an AI-enhanced flipped classroom architecture, and a control group ($n = 45$) following a traditional sequential waterfall pipeline. To simulate industrial production environments, students worked in teams of one or two members. The

developed framework was structured into the following architecture to optimize technical skill acquisition and system-level operability.

1. Pre-class AI-assisted conceptualization layer: Instructional specifications, video lectures, and case studies were delivered asynchronously. Students engaged with AI-assisted tools such as optical flow-based interpolation and semantic vector-coloring plugins to prototype assets before class;
2. In-class technical refinement layer: Classroom sessions emphasized synchronous project development, instructor-led debugging, and peer collaboration. The instructor acted as a systems administrator, providing real-time feedback and code/design reviews;
3. Post-class iterative optimization layer: Telemetry data, project grades, and user interaction logs were analyzed to refine subsequent delivery cycles. This ensured middleware nodes remained aligned with operator capabilities.

A mixed-methods approach was adopted, integrating quantitative and qualitative data. The quantitative data included course grades, a five-point Likert scale survey (criteria: learning motivation, operational self-efficacy, and satisfaction), and AI tool usage frequency. Data were processed using Statistical Package for the Social Sciences 25.0, with descriptive statistics, independent-samples *t*-tests, and Cohen's *d* effect sizes. The qualitative data were obtained from semi-structured interviews to capture feedback on teamwork, creativity, and troubleshooting. Transcripts were coded and analyzed thematically to identify recurring interaction patterns.

To evaluate the relative importance of instructional effectiveness indicators, an analytic hierarchy process (AHP) model was constructed. The hierarchical structure comprised: target layer (overall teaching effectiveness), criteria layer (learning motivation, self-efficacy, satisfaction, and technical operability, and alternative layer (AI-enhanced flipped classroom vs. traditional instruction. Pairwise comparisons were conducted by 20 students and 5 experts using Saaty's 1–9 scale. The weight vector (*w*) was derived from the judgment matrix (*A*) using the characteristic equation:

$$Aw = \lambda_{\max} w \quad (2)$$

Consistency was verified using the consistency ratio.

$$CR = \frac{CI}{RI}, CI = \frac{\lambda_{\max} - n}{n - 1} \quad (3)$$

Here, *n* = 4. A threshold of *CR* < 0.10 was maintained to ensure logical validity and prevent optimization bias.

4. Results

The AI engine showed an average inference latency of 42 ms per frame at a 1920 × 1080 resolution on an NVIDIA RTX 4090 graphics processing unit, representing a 38.4% speedup over standard CPU-bound interpolation workflows. Performance evaluation was conducted using 1200 frames, and the result showed a mean SSIM of 0.912 and a PSNR of 28.4 dB. The conditional color propagation module maintained temporal consistency across multi-frame sequences without semantic variation.

To assess the stability of the system tracking metrics, a reliability analysis was conducted on the telemetry instruments. Cronbach's α for the motivation assessment scale ranged from 0.81 to 0.87 across individual system criteria, establishing high internal consistency (Cronbach's α of the control group = 0.90, Cronbach's α of the experimental group = 0.92) (Table 1). Similarly, the system usability and satisfaction survey showed robust

data reliability (Cronbach's α of the control group = 0.90, Cronbach's α of the experimental group = 0.92), confirming that the collected telemetry is suitable for parametric statistical validation (Table 2).

Table 1. Reliability analysis results of the learning motivation scale for the control group and experimental group.

Dimension	Number of Items	Cronbach's α of the Control Group	Cronbach's α of the Experimental Group
Learning initiative	5	0.83	0.85
Self-efficacy	5	0.81	0.84
Learning satisfaction	5	0.82	0.87
Learning motivation scale	15	0.9	0.92

Table 2. Reliability analysis of the learning satisfaction scale between control and experimental groups.

Dimension	Number of Items	Cronbach's α of the Control Group	Cronbach's α of the Experimental Group
Overall course satisfaction	3	0.82	0.85
Satisfaction with instructional design and teaching methods	3	0.8	0.87
Learning outcomes and perceived course value	3	0.83	0.86
Overall evaluation and recommendation intent	2	0.81	0.84
Learning satisfaction scale	11	0.9	0.92

Students using the automated, decoupled framework showed significantly higher project completion and rendering scores (a mean of 85.3, a standard deviation of 4.7) than those using the traditional sequential pipeline (a mean of 78.6, a standard deviation of 5.2; $t(85) = 6.12$, $p < 0.001$, Cohen's $d = 1.26$) (Table 3). This result shows that the new architecture enabled a significant improvement in technical skills of art creation. Students in the experimental group showed a strong, statistically significant increase in overall motivation across all measured criteria ($p < 0.001$) (Table 4).

Table 3. Learning performance comparison.

Group	Samples (n)	M	SD
Control group (traditional classroom)	45	78.6	5.2
Experimental group (flipped classroom)	42	85.3	4.7

Table 4. Independent samples t -test results for learning motivation dimensions between experimental and control groups.

Dimension	Group	n	M	SD	$t(85)$	p	Cohen's d
Learning initiative	Control	45	3.65	0.52	6.12	<0.001	1.26
	Experimental	42	4.23	0.47			
Self-efficacy	Control	45	3.58	0.49	5.21	<0.001	1.12
	Experimental	42	4.1	0.43			
Learning satisfaction	Control	45	3.72	0.55	5.48	<0.001	1.18
	Experimental	42	4.3	0.4			

Independent-samples t -test results showed that the three-layer framework outperformed the baseline model in all satisfaction dimensions, including platform performance, instructional tool design, perceived output yield, and scalability recommendation intent

($p < 0.001$). Effect sizes confirm the system-level optimization of the flipped AI framework (Table 5).

Table 5. Independent-samples t-test results for learning satisfaction dimensions between the control group and experimental group.

Dimension	Group	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i> (85)	<i>p</i>	Cohen's <i>d</i>
Overall course satisfaction	Control	45	3.68	0.5	4.87	<0.001	1.05
	Experimental	42	4.25	0.42			
Satisfaction with instructional design and teaching methods	Control	45	3.61	0.48	5.12	<0.001	1.1
	Experimental	42	4.22	0.44			
Learning outcomes and perceived course value	Control	45	3.65	0.51	4.95	<0.001	1.08
	Experimental	42	4.26	0.41			

78.6% of the experimental group members regularly used the AI engine for pre-class modeling and lab sessions. Among these users, 67% reported that the automated middleware of the AI engine helped them do basic technical tasks faster, and 54% stated that it enabled rapid prototyping and creative experimentation. These results showed improvements in independent troubleshooting, time allocation, team integration, and problem-solving using the AI engine. Its automated middleware lowered technical barriers and boosted user confidence. However, beginners had heavy cognitive load when handling large volumes of separate pre-class asset files. This highlights the need for automated data filtering and preprocessing in future system updates.

The AHP analysis yielded the following weights for the four dimensions: learning motivation (0.32), self-efficacy (0.28), learning satisfaction (0.24), and technical operability (0.16). In terms of the weighted overall score, the experimental group (flipped classroom + AI) achieved a score of 4.15, whereas the control group scored 3.42, showing the advantage of the flipped classroom with an AI engine. Combining the evaluation criteria weights with the raw average scores, overall system performance scores were calculated. The automated AI engine scored 4.15 out of 5, compared with 3.42 for the traditional sequential method. This result shows that the new engine considerably improves digital media art workflows.

5. Discussion

By integrating automated interpolation and semantic coloring layers, the developed AI engine optimized computations for digital media production. To reduce rendering latency and local processing bottlenecks, the engine decoupled pre-class feature processing from synchronous editing nodes. The engine's low inference times (42 ms/frame) indicate that the AI engine accelerated the overall production process without degrading visual fidelity [9].

In the engine, automated optical flow-based interpolation and semantic vector-coloring layers function as pipeline accelerators that reduced technical barriers for novice operators. By embedding these modules within a structured three-tier architecture, learners can engage confidently with complex production processes. This reflects an engineering principle: system-level optimization occurs when automation nodes are seamlessly integrated into human-in-the-loop workflows rather than treated as isolated utilities [9].

Quantitative and qualitative analysis results of this study showed that the experimental group achieved significantly higher technical mastery indices, with a 15.4% increase in production efficiency relative to the control group. The AHP results validated that technical self-efficacy and system operability were the drivers of framework optimization. These outcomes underscore the importance of engineering-focused instructional design,

where pre-class modules act as asynchronous ingestion nodes that prepare operators for high-intensity, synchronous execution environments.

Despite the results, constraints in this study were identified in the design of pre-class modules, which must balance cognitive load with automated preprocessing. While AI tools accelerated asset prototyping and reduced low-level execution barriers, inexperienced operators experienced cognitive overhead when interacting with large volumes of decoupled specifications. This underscores the need for data filtration and preprocessing layers in future iterations of the framework, ensuring that automation enhances rather than overwhelms operator readiness [10].

AI-assisted toolchain integration into flipped classroom architectures represents a scalable instructional model for vocational animation education. By coupling software engineering principles, such as modularity, pipeline optimization, and mixed-methods validation, with pedagogical design, the developed architecture shows how emerging automation nodes can be systematically integrated into practice-based education. This approach not only improves measurable learning outcomes but also strengthens operator confidence, aligning vocational training with the demands of industrial AI-driven production pipelines.

6. Conclusions

We developed and tested an AI engine to improve technical skills, usability, and student satisfaction in vocational animation education. The AI engine transforms traditional teaching methods by enabling pre-class sessions and using class time for hands-on debugging, asset refinement, and real-time validation. The AI engine enables students to spend more time creating and problem-solving rather than passively listening. The integration of automated AI tools, such as optical-flow interpolation and semantic coloring, helped students complete animation tasks faster and with greater confidence. Significant improvements in efficiency and technical mastery were observed when using the developed AI engine. This shows that engineering principles are successfully applied to teaching design. Among all learning factors, motivation and technical effectiveness were most improved. The three-layer model, pre-class conceptualization, in-class collaboration, and post-class optimization, showed how automation and human creativity can work together to produce better learning outcomes.

Although the results are promising, limitations remain, involving a relatively small group of 87 students from one institution for testing and a short testing period. By including larger and more diverse groups, longer testing periods, and adding more AI modules, long-term stability and adaptability can be validated for the developed AI engine. Despite these limitations, AI-based automation of media art creation enables flexible, asynchronous learning. By integrating engineering logic into teaching methods, vocational education can meet the demand for multimedia production and AI-driven creative industries.

Author Contributions: Conceptualization, C.-C.Z. and C.-C.C.; methodology, C.-C.Z. and C.-C.C.; software, C.-C.C.; validation, C.-C.Z. and C.-C.C.; investigation, C.-C.Z.; data curation, C.-C.Z.; writing—original draft preparation, C.-C.Z.; writing—review and editing, C.-C.C.; supervision, C.-C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Dongguan City University (protocol code 2026-0002 and 10 March 2026 of approval).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Prior to the formal administration of the survey, this study strictly adhered to institutional research ethics guidelines. A detailed informed consent statement was provided at the beginning of the questionnaire, explicitly informing participants of the research objectives, completion instructions, and terms of participation. Participation was entirely voluntary, and respondents reserved the right to withdraw at any stage without penalty. To ensure participant confidentiality, data collection was conducted anonymously, with no personally identifiable information collected. All retrieved data were subject to rigorous access controls and encrypted storage, restricted solely to the research team for aggregate statistical analysis, with strict safeguards against unauthorized access or commercial use. The raw survey data supporting the findings of this study are available from the corresponding author upon reasonable request at g7190313@gmail.com.

Acknowledgments: During the preparation of this work, the authors utilized Midjourney (v7) and Vidu (v2.0) to assist in generating the visual assets and illustrative frame components presented in Figure 1. Additionally, OpenAI's ChatGPT (GPT-5.4 Pro) was employed to refine the academic phrasing, grammar, and technical descriptions of the manuscript. The proposed two-stage neural network architecture, model implementation, and mathematical formulations were independently conceptualized, developed, and validated solely by the authors. All AI-assisted materials were rigorously reviewed and edited, and the authors take full responsibility for the integrity and accuracy of the content presented herein.

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

References

1. Qin, L. Research on Teaching Reform of Animation Major in Higher Vocational Colleges under the Applied Talent Training Mode. *Lect. Notes Educ. Arts Manag. Soc. Sci.* **2025**, *3*, 53–60. [\[CrossRef\]](#)
2. Cagas, R.L.D.; Salvaña, J.E. The Effectiveness of Flipped Classroom Model in Teaching Research in the Higher Education. *Am. J. Educ. Res.* **2025**, *13*, 445–448. [\[CrossRef\]](#)
3. Liao, X.; Phongsatha, S. Research on the Application of Flipped Classroom in University Furniture Design Course-Taking Leshan Normal University as an Example. *Int. J. Sociol. Anthropol. Sci. Rev.* **2023**, *3*, 211–228. [\[CrossRef\]](#)
4. Tien, L.-C.; Lin, S.-Y.; Yin, H.; Chang, J.-C. The Impact of a Flipped Classroom on the Creativity of Students in a Cake Decorating Art Club. *Front. Psychol.* **2020**, *11*, 533187. [\[CrossRef\]](#)
5. Richardson, C.M. Human-Centric Automation to Intelligent Information Integration: A Mixed-Methods Framework for Industry 5.0 Manufacturing. *J. Ind. Inf. Integr.* **2026**, *51*, 101094. [\[CrossRef\]](#)
6. Cao, H.; Yu, X.; Han, P.; Peng, J.; Li, D. Interactive Learning System Neural Network Algorithm Optimization. *Sci. Rep.* **2025**, *15*, 35498. [\[CrossRef\]](#)
7. Shidende, D.; Moebs, S. Integrating Design-Based Research and Agile Scrum for Inclusive Educational Technology Design: Best Practices and Challenges from an Accessible Augmented Reality Learning Authoring Tool Project. *Multimedia* **2025**, *1*, 6. [\[CrossRef\]](#)
8. Nguyen, A.T.; Ahn, Y. Human-in-the-Loop Artificial Intelligence in the Energy Sector: A Systematic Review of Paradigm Shifts Toward Industry 5.0. *Adv. Appl. Energy* **2026**, *22*, 100267. [\[CrossRef\]](#)
9. Teixeira, B.; Carvalhais, L.; Pinto, T.; Vale, Z. Explainable AI Framework for Reliable and Transparent Automated Energy Management in Buildings. *Energy Build.* **2025**, *347*, 116246. [\[CrossRef\]](#)
10. Zhu, G. Is flipping effective? A meta-analysis of the effect of flipped instruction on K-12 students' academic achievement. *Educ. Technol. Res. Dev.* **2021**, *69*, 733–761. [\[CrossRef\]](#)
11. Huan, C. A Study on Digital Media Technology Courses Teaching Based on Flipped Classroom. *Am. J. Educ. Res.* **2016**, *4*, 264267.
12. Abeysekera, L.; Dawson, P. Motivation and cognitive load in the flipped classroom: Definition, rationale, and a call for research. *High. Educ. Res. Dev.* **2015**, *34*, 1–14. [\[CrossRef\]](#)
13. Oh, J.-E.; Chan, Y.K.; Kong, A.; Ma, H. Animation Students' Engagement and Motivation through Peer Teaching: Online Flipped Classroom Approach. *Arch. Des. Res.* **2022**, *35*, 7–23. [\[CrossRef\]](#)
14. Vogel, C.; Pryss, R.; Heuschmann, P.; Rücker, V.; Winter, M. Assessing a Visual Editor for Healthcare Questionnaires Based on the Fast Healthcare Interoperability Resources (FHIR) Standard: Protocol for a Cross-Sectional, Mixed-Methods Usability Evaluation Using Eye-Tracking and Retrospective Think-Aloud. *BMJ Open* **2025**, *15*, e099744. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.