# AI-Assisted UML Modeling for Serious Mental Illness Crisis Management

Balancing Automation and Human Oversight: A Comparative Analysis of UML Diagramming Methods

Kerry Gilder
Applied Health Informatics Program
Fordham University
New York, USA
e-mail: kgilder@fordham.edu

John Chelsom
Applied Health Informatics Program
Fordham University
New York, USA
e-mail: jchelsom@fordham.edu

*Abstract*— **This study assesses the effectiveness of Artificial Intelligence (AI) generated Unified Modeling Language (UML) diagrams in illustrating the treatment pathways for Serious Mental Illness (SMI) crises. AI-assisted tools, including ChatUML, ChatGPT, Claude, and DeepSeek, were assessed for accuracy, clarity, efficiency, and cost. A sample SMI use case was used to compare six AI-generated UML diagrams against a human-created benchmark. Results show that AI tools can improve diagram creation efficiency, with ChatUML using Claude 3.5 Sonnet and DeepSeek Reasoning with R1 accessed through Perplexity performing best. The limitations in the other AI-generated outputs demonstrate the need for human oversight to ensure precision in healthcare applications. The findings suggest that the use of generative AI in system planning and design would accelerate the development of pathways for healthcare providers managing SMI crises, with the potential to extend to other care settings.**

***Keywords-artificial intelligence; UML diagrams; serious mental illness; PlantUML; AI-assisted modeling*.**

## I. Introduction

The management of SMI crises can be a challenge for healthcare providers and patients alike, requiring detailed organization and procedures to identify accurate care pathways [1]. To make this process clearer, the individual care teams can coordinate services using UML Use Case diagrams that serve as a blueprint for teams to communicate and track relevant data. These diagrams have been found to be supportive in understanding the interactions between healthcare providers and services.

UML has been proven to be an effective tool for visualizing software design. It has been a valuable method for graphical view of system relationships in software engineering. UML is published by the Object Management Group (OMG) currently as UML 2.5.1 as a standardized modeling language for software engineering and system design [2]. A review of 128 papers of UML applications by Koç et al. [3] concluded that UML was a useful aid for design, modeling and class diagrams that support the identification of development requirements and system scope for computer science and industry applications. The visual ability to quickly display a system's requirements has been shown to improve analysis and design by breaking down complex steps into manageable components [3]. A detailed guide to implementing complex medical information systems recommends using UML to simplify the modeling of component relationships and grouping activities within the stages of development lifecycle. UML diagrams show the relationship between users and the system, a key function in care pathway development where user roles can be complex in healthcare applications [4].

As AI tools became more available for developers, Cámara et al. [5] evaluated ChatGPT's ability to perform modeling tasks and act as a modeling assistant. They found that while there were enhancements provided to UML models, there were limitations including lack of consistency and syntactic and semantic issues. However, they reported that the correctness of the models produced when using ChatGPT with PlantUML was much higher than ChatGPT alone, where PlantUML models made fewer syntactic errors. They conclude with the encouragement of incorporating AI models into model-based software engineering for the betterment of the modeling profession. Their findings reinforce the experience in this study, where AI models with PlantUML produced reliable outcomes.

PlantUML is an open-source UML diagramming tool that allows a user to begin with a text-based language that provides simple model representation of complex systems that are often used in the healthcare sector. It is favored for its effectiveness for healthcare modeling tasks and proves to be reliable in being syntactically correct in healthcare applications. PlantUML's popularity is attributed to its support of all UML diagrams, including class, use case, activity, sequence, component, deployment, and object diagrams [5]. The iterative flexibility of PlantUML supports agile development methods that use incremental modeling without locking in costly upfront design decisions. A recent study on Large Language Model (LLM) generation of UML models reported that PlantUML was the preferred tool for AI assisted UML diagram creation due to its widespread use and representation in LLM training data [6].

In researching available tools for UML generation, ChatUML was identified as a publicly available tool that describes itself as an AI-powered diagram generator that allows users to create and edit PlantUML diagrams using natural language conversations. It was launched in 2020 and generates UML diagrams through conversational interfaces



 ISBN: 978-1-68558-270-8 

through the ChatUML website and provides initial credits at no cost, with fees for additional use [7]. Earlier versions provided only ChatGPT for AI assistance but currently it offers a selection of commercially available AI tools to create PlantUML code for diagram generation. In addition to OpenAI's ChatGPT 4o, Anthropic's Claude Sonnet 3.5 and DeepSeek's R3 models were chosen for this study as current available AI tools inside ChatUML. The study will also evaluate each AI tool outside of the ChatUML shell for comparative effectiveness. Performance metrics of UML output for the SMI application will include technical accuracy, diagram clarity, time efficiency, and user costs. A manual diagram will be created as a benchmark to compare with AI tool's added value. The SMI use case will be used to create a consistent prompt across all tools, with comparative metrics used to provide a recommended toolset for developing UML. The SMI use case was developed using the Medicaid Innovation Accelerator Program (IAP), which helps manage beneficiaries with complex care needs and assists state agencies with data and workflow management of beneficiaries with SMI [8].

Complex SMI crisis management requires clear care pathways that UML diagrams can assist in structuring. This study aims to assess AI UML modeling tools for their effectiveness in generating accurate, efficient, and usable diagrams. This research will identify reliable and cost-effective AI-assisted solutions by comparing six AI-generated UML use case diagrams against a manual benchmark. The goal is to recommend an accessible AI tool that supports both system developers and healthcare providers in improving care coordination, decision-making, and workflow efficiency within the SMI treatment framework.

As AI has rapidly evolved, its tendency to hallucinate and generate false output that misrepresents the intended prompt has increased. Specific barriers with ethical, technological, liability and regulatory, workforce, social, and patient safety concerns have been cited, with the conclusion that human intervention is required to address barriers before AI can be safely and successfully applied widely in healthcare settings [9]. A Human-In-The-Loop (HITL) approach to augmenting AI in healthcare applications has been described, with human guided expertise to ensure safe application of AI in healthcare to lessen the fears and concerns that can stifle adoption of enabling technology [10]. Maintaining a human in the loop for AI applications is both a control for AI errors and a method of enhancing AI effectiveness in healthcare applications.

The objective of this research is to measure and validate the speed and efficiency that AI tools can provide in supporting and enhancing the human generation of UML diagramming that balances automation with human oversight to ensure accuracy and compliance with healthcare standards.

This investigation is organized to include a review of related works and an exploration of current approaches in the field of healthcare modeling in Section II. It then outlines the methodology for a comparative analysis in Section III, presents the results in Section IV with observations in Section V, and concludes with an evaluation of the findings and opportunities for future improvement in Section VI.

## II. Related Works

The use of Unified Modeling Language (UML) in healthcare settings has a history of studies that show the value in visually demonstrating system processes and provider interactions [3]. Recently, the integration of AI into modeling tasks has significantly advanced. Cámara et al. [5] explored ChatGPT's performance with UML modeling and noted improved outcomes when combined with PlantUML, while noting minor syntax issues. Conrady and Cabot [6] displayed how Large Language Models (LLMs) can generate UML diagrams from visual prompts in their work. These studies suggest that while AI can enhance modeling efficiency, there are still accuracy and consistency limitations.

The current study leverages these findings and focuses on the application of AI-generated UML diagrams in behavioral health, particularly for Serious Mental Illness (SMI) crisis management. It also includes multiple LLMs and use environments (ChatUML vs. native interfaces), as well as a benchmarking process against a human-generated diagram.

AI-assisted UML generation can utilize several tools, such as ChatGPT, Claude, and DeepSeek, as explored here. These tools, along with others, vary in their accuracy and usability. Prior studies have explored UML generation in generic modeling contexts [5], [6], but their application to healthcare-specific use cases or high-stakes environments such as serious mental illness (SMI) crises is less common. In the behavioral health use case explored here, existing solutions lack precision and struggle to perform consistently across dynamic, multi-role scenarios. A notable limitation of these tools is their current inability to consistently meet regulatory requirements and standards, making them unreliable in real-world application.

This study evaluates how current AI tools perform when applied to a standardized behavioral health use case. The ability to use these tools in real time with human input is also considered. The goal is to determine whether these tools, as they develop, can reliably support SMI response planning.

## III. Methods

To meet the research objective of determining the practicality of using emerging AI tools to generate a UML use case diagram for the care of a client experiencing an SMI crisis, the approach was to employ data simulation, manipulation, and processing. IAP's list of select behavioral health procedures for adult Medicaid beneficiaries with SMI was extracted for the simulated data input. For this evaluation, the procedures were divided into two categories, provider type and procedure type.

This data set was used as prompts across six AI tool sets, including ChatGPT (inside and outside ChatUML), Claude Sonnet 3.5 (inside and outside ChatUML), and DeepSeek (inside ChatUML and Perplexity). ChatGPT was selected with previous research documenting its UML capability [5] as well as its commitment to ensure privacy where specified [11]. Anthropic's Claude was included in the evaluation for its reputation for research accuracy and code generation capabilities. Anthropic's Privacy Policy emphasizes alignment with applicable law to provide user safety, ensuring user rights to delete, correct, restrict or withdraw consent for use [12]. DeepSeek has recently documented efficiency and low development costs as well as strong coding proficiency [13]. DeepSeek was not accessed directly through the standalone Chinese hosted servers due to potential security concerns with Chinese privacy practices. To access native DeepSeek Reasoning with R1, Perplexity has provided hosting service on US and Canadian servers that provide a higher level of security. Perplexity's Terms of Use allow users to define Confidential Information and restrict its use and commit to complying with the EU-U.S. Data Privacy Framework and the UK Extension, certified by the U.S. Department of Commerce [14]. The comparison of DeepSeek V3 within ChatUML to DeepSeek Reasoning with R1 outside of ChatUML using Perplexity was chosen to evaluate DeepSeek outside of ChatUML as an AI tool option for this research. The diagrams that were generated outside ChatUML required an additional step of taking the PlantUML code to a PlantText editor. This step was accounted for in the time elapsed consideration during scoring.

Figure 1 illustrates the dual-path procedure used to generate the UML diagrams. The prompts are input into either ChatUML or a separate AI tool, with outputs rendered into diagrams via ChatUML directly or through the PlantText UML Editor. Testing was conducted on an LG Gram laptop with an Intel Evo i7 processor and 16GB RAM using a stable internet connection. Each test was repeated three times, and average response times were recorded. Subsequently, a human-generated UML diagram was created to provide a benchmark of each criteria.

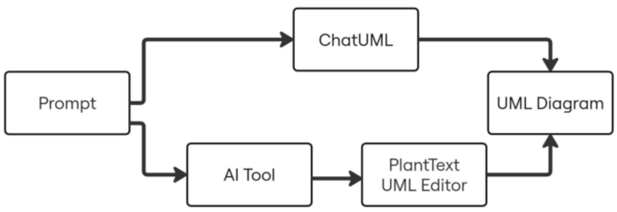


Figure 1: Creating UML diagram using AI tools.

To assess the performance of each tool, a manual review was conducted by a trained substance abuse counselor with knowledge of behavioral health workflows. Each AI-generated diagram was evaluated based on four key metrics, weighted according to their relative importance: Technical Accuracy (40%), Diagram Clarity (30%), Time Efficiency (20%), and User Cost (10%).

- Technical Accuracy – Measures the adherence of the UML diagram to the relationships provided and instructions explicitly defined in the prompt. 1 = fails to align with identified relationships or primary prompt instruction; 5 = accurately shows all provider relationships and workflows with no errors.
- Diagram Clarity – Ability to provide easily readable and usable diagrams. Clear diagrams enhance usability, enabling stakeholders to understand and use the models effectively. 1 = diagram is disorganized or unreadable, diagram fails to capture context; 5 = diagram is well-organized with logical flow, good labels and no distracting flows.
- Time Efficiency – Measures the total elapsed time for creating final UML output beginning with prompt initiation and includes iterations for correction and refinements. The quicker the tool generates accurate diagrams, the more efficiently it can support iterative workflows and real-time decision-making during development. 1 = over 60 seconds; 5 = Under 10 seconds.
- User Cost – Expense required for completing each output. This measure is a relative comparison of costs for each tool. 1 = most expensive; 5 = free.

Technical Accuracy was the highest weighted metric to assure proper use. Diagram Clarity was second to assure ease of use. Time Efficiency was third and included iterations, with human-generated results serving as a comparison. User Cost was the lowest weight, serving as a relative comparison of nominal expenses across tools. Each diagram was scored using a five-point scale per metric, and an overall weighted score was calculated using the formula:
Final Score=(0.4×Accuracy Score)+(0.3×Clarity Score)+(0.2×Time Efficiency Score)+(0.1×User Cost Score)

This methodology ensured a structured evaluation, allowing for a direct comparison of AI-generated UML diagrams against each other and the human-generated benchmark. The results provide comparisons of accuracy, clarity, efficiency, and cost-effectiveness across different AI models.

Prompt engineering was important for this study, using behavioral health experience and AI to craft prompts for SMI-related workflows. The prompt was designed and the procedures were classified in groups using knowledge of the Diagnostic and Statistical Manual of Mental Disorders, Fifth



 ISBN: 978-1-68558-270-8 

Edition, Text Revision (DSM-5-TR) [15] and behavioral health treatment models. Additionally, successive iterations were applied to assess how different wording and structuring influenced AI outputs. Due to space limitations, only two packages of the diagram are illustrated as the practical example; however, the outcome is consistent with the use of the full prompt that included seven packages.

Original Prompt (seven packages): Create a vertical use case diagram with the information provided, please use plantuml Core Psychiatric Services 1. Pharmacologic Management * Clinician: Psychiatrist 2. Other Psychiatric Services/Procedures * Clinician: Psychiatrist 3. Diagnostic Interview Exam * Clinician: Psychiatrist, Psychologist 4. Crisis Intervention * Clinician: Psychiatrist, Crisis Counselor Therapeutic Services 1. Individual Therapy * Clinician: Therapist 2. Group Psychotherapy * Clinician: Therapist 3. Behavioral Health Day Treatment * Clinician: Therapist 4. Individual Psychotherapy * Clinician: Therapist 5. Family Psychotherapy * Clinician: Marriage & Family Therapist (MFT) Assessment & Diagnosis 1. Mental Health Assessment * Clinician: Psychologist, Therapist 2. Diagnostic Interview Exam * Clinician: Psychiatrist, Psychologist 3. Alcohol/Drug Assessment * Clinician: Addiction Counselor, Psychiatrist Substance Use Disorder Services 1. Alcohol/Drug Services Intensive Outpatient * Clinician: Addiction Counselor, Case Manager 2. Alcohol/Drug Abuse Services * Clinician: Addiction Counselor 3. Alcohol/Drug Case Management * Clinician: Case Manager Crisis & Community-Based Care 1. Crisis Intervention * Clinician: Crisis Counselor, Psychiatrist 2. Community-based Wrap-around Services * Clinician: Case Manager, Therapist Medical & Physical Health 1. New Patient Office Visit * Clinician: Physician 2. Admission History & Physical Exam * Clinician: Physician, Psychiatrist Behavioral & Preventive Services 1. Therapeutic Behavioral Services * Clinician: Behavioral Therapist 2. Behavioral Health Prevention Education * Clinician: Health Educator 3. Behavioral Health Prevention Information Dissemination * Clinician: Health Educator [8].

Truncated Prompt for the Practical Example (two packages): Create a vertical use case diagram with code for the information provided, be sure to include a heading. Please use PlantUML: Core Psychiatric Services 1. Pharmacologic Management * Clinician: Psychiatrist 2. Other Psychiatric Services/Procedures * Clinician: Psychiatrist 3. Crisis Intervention * Clinician: Psychiatrist, Crisis Counselor Assessment & Diagnosis 1. Mental Health Assessment * Clinician: Psychologist, Therapist 2. Diagnostic Interview Exam * Clinician: Psychiatrist, Psychologist 3. Alcohol/Drug Assessment * Clinician: Addiction Counselor, Psychiatrist.

## IV. RESULTS

Figures 2 – 8 present the UML use case diagram outcomes of the standardized truncated prompt for each of the six AI tool methods described, and the human created diagram. Table 1 presents the scoring for each method for the weighted criteria described.

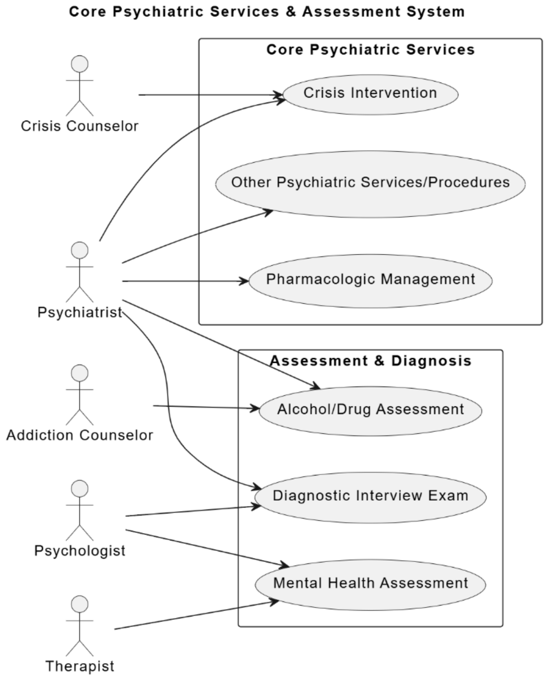


Figure 2: ChatUML, Claude 3.5 Sonnet.

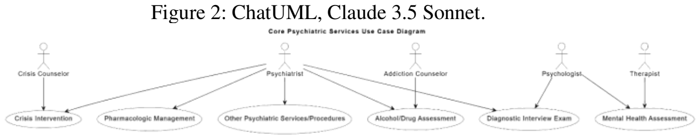

Figure 3: ChatUML, ChatGPT 4o.

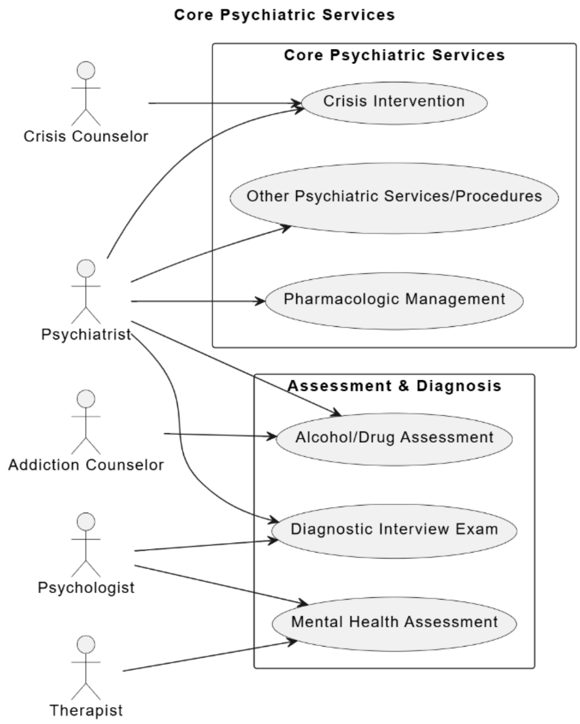


Figure 4: ChatUML, DeepSeek V3.

 ISBN: 978-1-68558-270-8

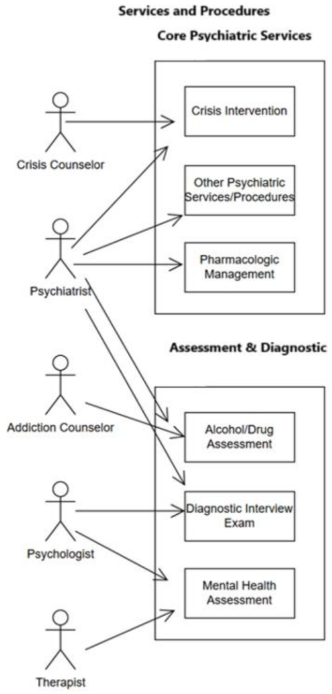


Figure 5: Human, experienced in behavioral health.

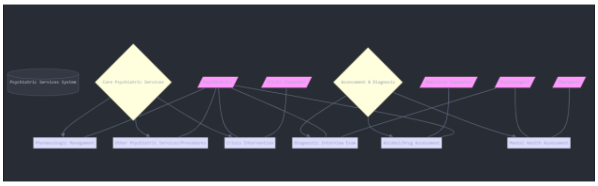

Figure 6: Claude 3.5 Sonnet.

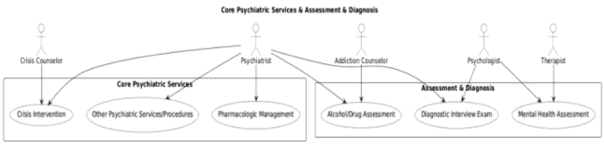


Figure 7: ChatGPT 4o.

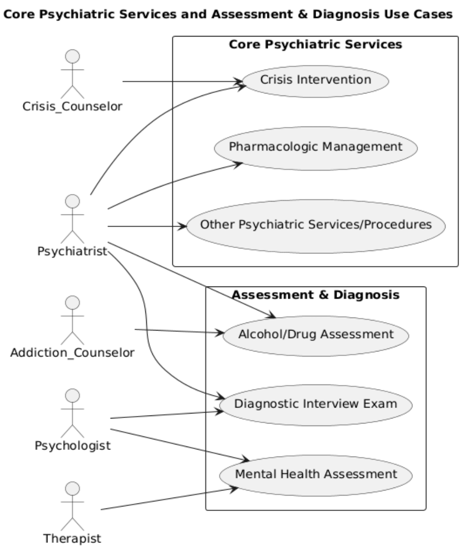


Figure 8: Perplexity, DeepSeek Reasoning with R1.

TABLE 1. UML USE CASE DIAGRAM SCORECARD TABLE.

| AI Tool | Technical Accuracy | Diagram Clarity | Time Efficiency | User Cost | Weighted |
|---|---|---|---|---|---|
| ChatUML, Claude | 5 | 5 | 5 | 4 | 4.9 |
| ChatUML, ChatGPT | 2 | 2 | 5 | 4 | 2.8 |
| ChatUML, DeepSeek | 3 | 5 | 2 | 4 | 3.5 |
| Human | 5 | 5 | 1 | 5 | 4.2 |
| Claude | 1 | 2 | 4 | 5 | 2.3 |
| ChatGPT | 3 | 4 | 4 | 1 | 3.3 |
| DeepSeek, Perplexity | 5 | 5 | 4 | 5 | 4.8 |

The highest scoring method was Claude 3.5 Sonnet inside ChatUML, with a weighted score of 4.9, followed closely by DeepSeek Reasoning with R1 accessed by Perplexity with a score of 4.8. The lowest score was Claude Sonnet 3.5 used outside of ChatUML.

## V. DISCUSSION

The results revealed several issues within the Technical Accuracy criteria. ChatGPT, both within and outside ChatUML, as well as Claude Sonnet 3.5 outside ChatUML, did not follow the prompt. ChatGPT in both environments produced horizontal diagram outputs despite clear instructions to the contrary, as vertical was specified. Meanwhile, Claude Sonnet 3.5 in its native environment did not use PlantUML as directed and chose to default to Mermaid, which is Claude's built-in diagramming tool. Additionally, DeepSeek within ChatUML mislabeled the diagram by generating an incorrect heading. These deficiencies in Technical Accuracy are reflected in the scores. The Human, as expected, created an accurate diagram.

All AI tools were given the same prompt but their Diagram Clarity varied based on structure and presentation. ChatUML (Claude) and ChatUML (DeepSeek) produced the clearest diagrams, scoring 5/5. ChatUML (ChatGPT) scored the lowest (2/5) due to missing boxes and headings, making it difficult to follow. Native ChatGPT (4/5) was structured vertically, which affected readability. Native Claude (2/5) chose distracting colors and had small text, making it practically illegible. DeepSeek (Perplexity) and human-created diagrams (both 5/5) showing AI parity with human design. This highlights the importance of formatting, structure, and visual presentation in Diagram Clarity.

The Time Efficiency results show differences in processing times within AI tools, and the expected delays with the manual/human method. Both Claude Sonnet 3.5 and ChatGPT 4o performed efficiently within ChatUML, each completing the task in 10 seconds. However, DeepSeek took longer at 60 seconds, which may be due to inefficiencies between ChatUML and DeepSeek V3. The manual process, by comparison, was the most time-consuming at 30 minutes, demonstrating the advantage of



 ISBN: 978-1-68558-270-8

automation. When tested natively, Claude Sonnet 3.5 maintained its 10-second speed, while ChatGPT 4o took slightly longer at 17 seconds. Interestingly, Perplexity DeepSeek performed much faster natively, completing the task in 15 seconds compared to its significantly longer runtime in ChatUML, perhaps indicating efficiencies provided by Perplexity. These findings show the potential trade-offs between different AI environments and suggest that while some tools perform optimally in certain settings, others may experience slowdowns due to integration constraints.

The cost analysis of UML implementations shows that ChatGPT Native is the most expensive option (rated 1), costing $20 per month. The middle tier (rated 4) consists of ChatUML implementations across platforms—Claude ChatUML, ChatGPT ChatUML, and DeepSeek ChatUML—offering a lower priced alternative. ChatUML operates on a credit-based system, with pricing starting at $2.99 for 20 credits, though the 250-credit package for $6.99 was used here due to the iterative approach and multiple tools sampled. It is worth noting that ChatUML follows a tiered system, where all three tools used in this analysis cost 3 credits per request. Previously, the credit cost for ChatGPT 4o and Claude 3.5 Sonnet was 5 credits per request but on December 7, 2024, it decreased to 3 credits. DeepSeek ChatUML, which was only integrated after December 2024, has always been priced at 3 credits per request. Finally, Human manual development, Claude Native, and DeepSeek/Perplexity Native are the most economical choices, as they are free (rated 5).

The outcome of the scoring against the weighted criteria indicates that Claude (ChatUML) is the recommended tool set for producing accurate, clear, efficient and cost-effective UML for an SMI application. Perplexity DeepSeek is also a reasonable tool set and is only slightly slower due to the extra step of copying PlantUML code into PlantText.

## VI. CONCLUSION

This study suggests that the application of AI in UML diagram generation holds significant potential to address the challenges of time efficiency and complexity management. This research started with the objective of evaluating artificial intelligence in the generation of a UML use case diagram that described a simulated care pathway for patients experiencing an SMI crisis that was managed by multiple healthcare providers. This objective was met through the practical evaluation of six AI tool environments. ChatUML and PlantUML are powerful tools for supporting the creation of a detailed diagram that met the research objective. Claude Sonnet 3.5 used inside of ChatUML provided the most accurate output and allowed for a more nuanced diagram of the interactions within the SMI crisis management system, demonstrating the potential of AI to significantly reduce the time and effort required in the system design process. The generated UML diagram (Figure 2) was a successful step in meeting the research goal. The investigation showed that AI can accelerate the development cycle and assist in managing complex system design. However, it is not yet ready to be hailed as the complete substitute for an informed expert. The AI-generated diagram required iterations and oversight to ensure that the details of SMI crisis management were accurately captured. This reinforces the current state of AI as a complementary tool rather than a replacement for human expertise. There are opportunities for further work that can expand on these research findings and add to its usefulness. As AI tools progress, it is anticipated that providing more domain-specific knowledge to AI Large Language Models (LLMs) like OpenAI, Claude and DeepSeek will enhance their ability to generate more scenario specific UML diagrams. AI itself will likely then be able to leverage the domain knowledge to automatically update diagrams based on real-time data, which could provide a dynamic tool for system management. Additionally, providing feedback of user's interactions with AI-generated diagrams could add to their effectiveness as communication tools for healthcare stakeholders, maintaining the HITL principle [10]. In reflection, the research investigation has confirmed that AI has current usefulness and significant promise for system design in the complex healthcare sector. There was human learning in the exercise as well. The iterations required by this research identified better ways to phrase a problem for the AI to solve. Healthcare providers as novice users will quickly become more agile in using ChatUML with Claude Sonnet 3.5 or DeepSeek Reasoning R1 through Perplexity, and their second and third applications will produce more useful diagrams. While the AI did not fully replace the need for human expertise, it served as an intelligent assistant by streamlining the design process. The results present a compelling case for the integration of AI in technical documentation practices as a viable way forward while maintaining control with a human oversight. As AI continues to evolve, its potential to transform system design and management practices is evident, validating continued exploration and development for healthcare design.

The combination of AI-powered methods produced a usable outcome as the final product. It demonstrates the efficiency provided by significantly reducing the time required to produce complex UML diagrams. The relative speed and efficiency of using AI tools is a meaningful contrast to the traditional manual diagramming and programming methods. This research identifies available public tools that provide advancement in the system design and documentation of UML diagrams for healthcare use cases. This application confirms the efficacy of AI in generating critical technical artifacts and demonstrates the possibilities for more rapid adaptation of a new approach for system visualization in the context of healthcare management.

 ISBN: 978-1-68558-270-8 

ACKNOWLEDGMENT

We thank Fordham University and the Applied Health Informatics program for their support. Kerry Gilder is especially grateful to her teachers, classmates, and her family, particularly her father, whose openness and willingness to listen have been invaluable in shaping her ideas and education.

REFERENCES

[1] F. Röhricht et al., "Implementation of a novel primary care pathway for patients with severe and enduring mental illness," BJPsych Bulletin, vol. 41, no. 6, pp. 314-319, Dec. 2017, doi: 10.1192/pb.bp.116.055830.

[2] [Object Management Group (OMG), Unified Modeling Language (OMG UML), Version 2.5.1, 2017. [Online]. Available: https://www.omg.org/spec/UML/2.5.1/PDF, retrieved: January, 2025.

[3] H. Koç, A. M. Erdoğan, Y. Barjakly, and S. Peker, "UML diagrams in software engineering research: A systematic literature review," Proceedings, vol. 74, no. 1, p. 13, 2021. [Online]. Available: https://doi.org/10.3390/proceedings2021074013, retrieved: January, 2025.

[4] A.-M. Ștefan, N.-R. Rusu, E. Ovreiu, and M. Ciuc, "Empowering healthcare: A comprehensive guide to implementing a robust medical information system—Components, benefits, objectives, evaluation criteria, and seamless deployment strategies," Appl. Syst. Innov., vol. 7, no. 3, p. 51, Jun. 2024. [Online]. Available: https://doi.org/10.3390/asi7030051, retrieved: January, 2025.

[5] J. Cámara, J. Troya, L. Burgueño and A. Vallecillo, "On the assessment of generative AI in modeling tasks: an experience report with ChatGPT and UML," Softw. Syst. Model., vol. 22, pp. 781–793, 2023. [Online]. Available: https://doi.org/10.1007/s10270-023-01105-5, retrieved: January, 2025.

[6] A. Conrady and J. Cabot, "From image to UML: First results of image-based UML diagram generation using LLMs," arXiv:2404.11376 [cs.SE], 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2404.11376, retrieved: January, 2025.

[7] ChatUML, "ChatUML," [Online]. Available: https://chatuml.com/, retrieved: January, 2025.

[8] Centers for Medicare & Medicaid Services, "Using data analytics to better understand Medicaid populations with serious mental illness," Medicaid Innovation Accelerator Program, n.d. [Online]. Available: https://www.medicaid.gov/state-resource-center/innovation-accelerator-program/iap-downloads/program-areas/data-analytics-smi-tech-resource.pdf, retrieved: January, 2025.

[9] M. I. Ahmed, B. Spooner, J. Isherwood, M. Lane, E. Orrock, and A. Dennison, "A systematic review of the barriers to the implementation of artificial intelligence in healthcare," Cureus, vol. 15, no. 10, p. e46454, 2023. [Online]. Available: https://doi.org/10.7759/cureus.46454, retrieved: January, 2025.

[10] E. Sezgin, "Artificial intelligence in healthcare: Complementing, not replacing, doctors and healthcare providers," Digital Health, vol. 9, p. 20552076231186520, 2023. [Online]. Available: https://doi.org/10.1177/20552076231186520 retrieved: January, 2025.

[11] OpenAI, "Usage policies," 2024. [Online]. Available: https://openai.com/policies/usage-policies/, retrieved: January, 2025

[12] Anthropic, "Privacy policy," 2024. [Online]. Available: https://www.anthropic.com/legal/privacy, retrieved: January, 2025.

[13] DeepSeek AI, "DeepSeek-V3 technical report," arXiv, Dec. 26, 2024. [Online]. Available: https://arxiv.org/abs/2412.19437, retrieved: January, 2025.

[14] Perplexity AI, "Enterprise terms of service," 2025. [Online]. Available: https://www.perplexity.ai/hub/legal/enterprise-terms-of-service, retrieved: January, 2025.

[15] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, 5th ed., text rev., 2022. [Online]. Available: https://doi.org/10.1176/appi.books.9780890425787, retrieved: January, 2025.