How effective is generative AI advice for the academic advancement of faculty?

Short title: Academic advancement advice from generative AI

May May Yeo[1,*], Deborah D. Navedo[2], Patrick J. Casey[3,4] Bernard Tan[5]

[1]Department of Academic Appointments and Development, Duke–NUS Medical School, Singapore

[2]Clinical Perspectives Consultants, USA

[3]Dean's Office, Duke–NUS Medical School, Singapore

[4]Duke University, Durham, North Carolina

[5]Department of Information Systems and Analytics, National University of Singapore

*Corresponding author

E-mail: maymay.yeo@duke-nus.edu.sg (MMY)

Abstract

The SingHealth Duke-NUS Academic Medical Center manages over 2,800 clinical faculty members and processes over 400 appointments and promotions annually. The existing Promotion and Tenure documentation comprises over 30 documents, making it time-consuming for faculty to find specific academic appointment information. We developed "AskADD" in response to requests for clearer academic career development guidance. In this study, we report the results of initial pilot (alpha) testing with 35 faculty members, followed by repeat (beta) testing with the same cohort, using a newly developed, innovative academic policy artificial intelligence (AI) tool termed "AskADD," designed to assist faculty in navigating academic career pathways. AskADD aids faculty—covering physician-educators, physician-scientists, physician-innovators, and physician-leaders—in navigating academic career paths while increasing transparency and trust in the appointment, promotion, and tenure processes. Our AI-integrated systems, whose initial pilot testing was built using low or no code platforms on Microsoft, and whose launch version was built using customer GPT, can provide contextualized responses for promotion and tenure queries. Faculty and staff participated in user testing, offering feedback that informed iterative improvements. Alpha and beta testing conducted with the same group of faculty members indicated that a substantial portion of participants found the tool beneficial; and quantitative and qualitative feedback provided suggestions for further refinement. Our experience contributes to the limited body of literature on AI-driven faculty advancement in academic medical centers and offers a novel paradigm for academic career support.

Introduction

The SingHealth Duke-NUS Academic Medical Center (AMC) oversees more than 2,800 clinical faculty and manages over 400 appointments and promotions annually. Promotion and Tenure (P&T) information for the medical school is spread across over 30 documents on the school's intranet, which makes it time-consuming for the faculty to locate specific information on academic appointments. During engagement exercises with the school, faculty expressed a need for more comprehensive guidance on academic career advancement. Accordingly, the research team proposed a tool to help faculty navigate and explore academic careers. Our tool aimed to reduce time for faculty searching for P&T information, increase transparency of the P&T policies and processes, and align faculty aspirations with the AMC's mission and vision [1]. Using Microsoft's Power Platform that includes Power Apps, Power Automate, and the OpenAI GPT-4o on Microsoft Azure (OpenAI GPT-4o, Redmond, WA, USA) [2,3] procured by the university, we developed an academic policy AI tool to help faculty navigate P&T policies and processes. When implemented, the tool could serve as a leadership development support resource for clinical program leadership and promotion and tenure (P&T) committees, while reducing the administrative burden on staff by addressing common faculty queries through the chatbot. We invited faculty and staff to test the tool and provide feedback before its launch, ensuring that it is effective and user-friendly.

Various studies have investigated the application of AI in vocational education career planning; the use of AI mentors to navigate science, technology, engineering, and mathematics (STEM) careers; AI's role in shaping individual career paths across career stages, and in health professions education [4–9]. To our knowledge, no prior study has investigated the application of generative AI to support faculty candidates at AMCs in navigating policies and procedures related to promotion, tenure, and academic advancement. We aimed to address these gaps through the following research questions:

1. Is it possible to create a secure, interactive, online web application that allows faculty members to seek advice on academic advancement?
2. How effective is generative AI in advising faculty on academic advancement?

This study aimed to develop and test an academic policy AI tool to assist faculty in navigating and understanding the policies and processes for academic appointments, promotions, and tenure at the SingHealth Duke-NUS AMC. We anticipate that when implemented, this tool will save time when searching for promotion and tenure policy information. Through providing contextualized natural language advice to faculty candidates preparing for promotion, our tool addresses their specific needs and increases their potential for success.

Materials and methods

Ethical Approval

The National University of Singapore Institutional Review Board approved this study (date of approval: December 19, 2024; NUS-IRB Reference Code: NUS-IRB-2024-121). All participants were provided with an information sheet and completed a participant consent form.

Study design, setting, and period

We performed alpha and beta testing of and obtained feedback from participants on a newly developed chatbot to help faculty navigate promotion and tenure policies and processes between January and April 2025.

Alpha test

From August 2024 to early January 2025, we programmed the university's enterprise Microsoft Power Apps and Power Automate to contact the University's enterprise OpenAI GPT-4o

application programming interface and send user prompts to query a Microsoft SharePoint repository containing the school's P&T policy documents. OpenAI GPT-4o sent responses that could be viewed by the faculty through the user interface of Power Apps (Fig 1). Details of the tool, code and system prompts are in S1 Appendix.

Fig 1. Using the University's Enterprise Microsoft Power Platform and OpenAI GPT-4o on Microsoft Azure to Generate Natural Language Advice for Faculty.

Study population, sample size and sampling procedures, data collection, and analysis

We launched the alpha version in mid-January 2025. To provide timely insights, we aimed to recruit a purposive sample of 30 participants [10]. Between January 17 to February 13, 2025, we sent invitation emails in batches to 105 faculty and staff members of the SingHealth Duke-NUS AMC to participate. Non-faculty and administrative staff members of the hospitals were excluded from the invitation. 36 faculty members and staff (34% response rate) agreed to participate. They represented a diverse range of user groups, including faculty from various tracks, e.g. educator and research tracks; senior and junior faculty; clinical program leadership; and administrative staff who assisted faculty in P&T processes. We obtained feedback from diverse user groups to help develop a tool that would be relevant and widely accepted [11].

Participants were requested to formulate queries, evaluate responses provided by the tool to them, and provide feedback on functionality and tool user interface via an online form (S2 Appendix) developed using Qualtrics (Provo, UT). We adopted the following factors to assess chatbot user satisfaction: usability, which indicated the relevance and clarity of responses; responsiveness, which indicated the quickness and accessibility of services and timeliness of responses; perceived empathy of the chatbot, and accessibility, wherein users with any browser had complete access to all the information and ability to interact with the chatbot. We selected these

factors based on theoretical and empirical support which had demonstrated their relevance in the context of AI chatbot interactions [12]. We added the Net Promoter Score (NPS) [13], a measure of the participant's likelihood of recommending the chatbot to colleagues, and which we felt complemented the assessment by providing an action-focused indicator of user sentiment. Moreover, we consulted the university's enterprise "NUS AI-Know" and included another factor – comparison with traditional methods, such as administrative staff – to understand if the chatbot could be an effective alternative to human assistance [14]. We used descriptive statistics to summarize participants' feedback variables and plotted the data using Microsoft Power BI (Microsoft Corp.). Finally, we solicited additional feedback via two open-ended questions commonly used within the AMC's faculty development programs— "What specific features of AskADD do you find most effective?" and "What improvements would you suggest in enhancing the effectiveness of AskADD?" Authors Yeo M.M. and Navedo D.D. read and analyzed the feedback using a reflective attitude toward capturing participants' meanings, and summarized them using a flexible framework for thematic analysis [15]. We created *a posteriori* themes or codes as we progressed, adding them to a table, and summarizing the answers onto it. In developing the table, we noticed that the themes fitted into the factors of usability, responsiveness, accessibility, empathy, and comparative sources and we termed them *a priori themes or codes*. Feedback that related to two themes or codes was separated accordingly. Although we had planned to group feedback that did not align with the *themes or codes* under "Other Comments", there was no such feedback in the alpha phase. Alpha testing concluded in the final week of February 2025.

Alpha test feedback and custom GPT development in beta test

The National University of Singapore (NUS) launched its enterprise "NUS AI-Know" platform in October 2024. The university enabled the creation of custom GPT sharing in NUS AI-Know within its community in January 2025. Accordingly, we developed a custom GPT (S3

Appendix) using NUS AI-Know. During the process, we incorporated participants' feedback from the alpha test for improvement and adopted OpenAI's prompt engineering strategies to enhance the results [16]. We conducted beta testing from mid-March to mid-April 2025. The same group of faculty and staff was requested to frame queries and evaluate responses from the tool and provide feedback as mentioned above. All users except one responded, yielding 35 responses. To facilitate feedback comparison between both phases, we eliminated feedback from the alpha test that lacked corresponding beta test feedback. We launched the tool when we were satisfied with the accuracy of its responses and its functionality. Figure 2 summarizes the developmental steps.

Fig 2. Steps in the Development and Implementation of a Custom GPT to Answer Faculty Questions on Academic Advancement.

Results

Participant characteristics

In total, there were 35 participants in both alpha and beta tests; more than half of them were men (57%). The participants predominantly comprised medical staff (60%); 40% were in the clinical track, 26% were senior faculty, and 26% were engaged in school or institutional leadership (Table 1).

Table 1. Demographic Characteristics of Participants at SingHealth Duke-NUS Academic Medical Center (AMC) who Participated in AskADD Alpha and Beta Testing, 2025.

Category	*Demographics (n=35)*	*Number*	*Percentage (%)*
Professional group	Administration	5	14

Category	*Demographics (n=35)*	*Number*	*Percentage (%)*
	Allied Health Professional	1	3
	Dental	3	9
	Medical	21	60
	Nurse	1	3
	Research	4	11
Faculty track *Where applicable*	Clinical	14	40
	Educator	5	14
	Practice	1	3
	Research	6	17
	Tenure	4	11
	Not applicable, i.e., administrative	5	14
Role	Administrative Staff	4	11
	Clinical Program Leadership	8	23
	School or Institutional Leadership	9	26
	Junior Faculty	5	14
	Senior Faculty	9	26
Sex	Female	15	43
	Male	20	57

Figure 3 compares user feedback in both testing phases.

Fig 3. User Feedback Comparisons Denoting Overall Improvement in the Results of the Beta over the Alpha Version.

Findings from alpha testing

Users rated AskADD strongly in terms of usability, with high relevance and clarity of responses and effectiveness in saving users' time compared with searching the school's intranet (Fig 3: Questions 1, 2 and 4 Alpha). Moreover, users also positively rated its responsiveness (Fig 3: Question 9 Alpha).

While 3 users reported never needing to rephrase their queries, 32 needed to, including 2 needing to rephrase often (Fig 3, Question 3 – Alpha). While the majority of participants reported satisfaction with the chatbot's user-friendliness and perceived empathy (Fig 3, Questions 10 to 12 – Alpha), between 23% and 40% indicated that there was room for improvement in the overall user experience (Fig 3, Questions 11 and 12 – Alpha). Notably, most users felt the chatbot outperformed human counterparts in delivering academic policy information (Fig 3, Question 13 – Alpha). However, 66% of users fell into the "passive" category of the NPS; that is, they rated the tool a 7 or 8 on a scale from 0 (not at all likely to recommend) to 10 (extremely likely to recommend) (Fig 3, Question "How likely are you to recommend AskADD to other colleagues?" – Alpha). These respondents were considered "passively satisfied" and indicated they would be unlikely to actively recommend the tool to colleagues, as doing so might risk their own professional credibility (11, p.4).

Additional feedback is presented in S4 Appendix. Users felt that the chatbot had to provide links to policy documents to enable users to verify information and explore further. It also had to improve the handling of detailed follow-up advice, e.g. provide contact information to relevant departments. Users requested to structure the answer formats for better readability. Additionally, the chatbot should include information on academic development and human resource information to accommodate a wider range of user needs.

Findings from beta testing and comparison with alpha test results

Users reported that the custom GPT provided relevant and clear responses, significantly reducing the time required to locate information. It was perceived as easy to interact with and responsive. The majority of participants reported positive experiences, noting that the tool was more effective than other sources of policy information such as asking administrative staff, for navigating academic policies. The NPS increased from 13% in the alpha version to 42% in the beta version.

Usability

In comparing usability between alpha and beta testing, the number of participants who felt AskADD's responses were very relevant increased from 10 to 17 (Fig 3, Question 1). Those who rated the responses as very clear rose from 9 to 19 (Fig 3, Question 2). The number of users who never needed to rephrase their questions increased from 3 to 7, while the number of those who often had to rephrase their questions dropped to 0 (Fig 3, Question 3).

The number of users who reported that AskADD reduced the time required to locate the school's P&T policies by more than 15 min increased from 11 to 19 (Fig 3, Question 4). However, the number of users who reported that AskADD reduced the time required to locate the school's P&T policies by less than five minutes remained the same.

Participants who agreed that AskADD made it easier to navigate and understand policy documents increased from 33 to 34 and from 29 to 33, respectively (Fig. 3, Questions 5 and 6). Similarly, 28 (alpha) and 35 (beta) users felt that AskADD was able to resolve their queries (Fig. 3, Question 7).

Responsiveness

Users in the beta phase reported similar turnaround times as in the alpha phase (Fig. 3, Question 8). The number of users who found it easy to contact AskADD increased slightly from 30 (alpha) to 31 (beta) (Fig. 3, Question 9).

Accessibility

In the beta phase, 34 users agreed that AskADD had an easy-to-use interface, up from 33 in the alpha phase (Fig. 3, Question 10).

Empathy

Three additional users were impressed by AskADD's responses in the beta phase (Fig. 3, Question 11). The number of users who enjoyed chatting with the tool increased from 21 (alpha) to 29 (beta), while those who did not enjoy the experience dropped to 6 in the beta phase (Fig. 3, Question 12).

Comparing AskADD with other sources of policy information

In the alpha phase, 8 users felt that AskADD was much more effective than traditional sources, increasing to 13 in the beta phase (Fig. 3, Question 13). While the majority considered AskADD a positive alternative for accessing academic policy information, 5 users in both phases felt it was similarly or less effective than other sources.

NPS

The proportion of promoters increased from 23% in the alpha phase to 51% in the beta phase. Passives decreased from 66% to 40%.

Figure 3 also highlights two questions with the most notable improvement from alpha to beta: the tool's ability to resolve queries (Fig. 3, Question 7) and perceived empathy (Fig. 3, Question 12), with gains of seven and eight users, respectively.

We explored the additional feedback to the two open-ended questions in the beta phase (S4 Appendix) to identify changes that could have contributed to the above outcomes. Features users found most effective included time-savings—*"Quick answers, don't need to trawl website for info"*; relevant answers—*"Able to ask real questions and get directed in the right direction"*, hyperlinks to source documents: *"Excellent as is, especially updated version with links to the source references"*, and introduction prompts—*"Useful prompts to start a conversation"*. Users also requested streamlining access through institutional network. There were also suggestions on document and application support, such as enabling users to upload documents for AskADD to provide personalized advice on academic rank and gap assessment. Where it could not provide adequate answer, AskADD should be able to provide specific contact information. Users also suggested implementing a function for real-time user feedback to continuously improve the platform based on user experiences and issues encountered. Users also felt the appearance and speed of answers could be improved. The research team discussed and implemented the features where functionalities were available into the system before deployment.

Additional findings from the tool development process

During the development of the alpha version, the team conducted approximately 60 iterations to enable hyperlinking to source documents stored in the SharePoint repository. In approximately 70% of cases, the Azure AI GPT correctly returned file name formats, allowing the Power Automate flow to retrieve and link the required documents seamlessly for user access. However, in the remaining 30% of runs, inconsistencies in the file name formatting generated by Azure AI GPT disrupted Power Automate's functionality—due to its reliance on standardised formatting—resulting in non-functional hyperlinks delivered via Power Apps. This limitation reduced the alpha version's usefulness, which was addressed in the beta version through backend programming by the AI development team.

Discussion

This study aimed to pilot an academic policy AI tool to parse P&T policies and to understand if it was effective in advising faculty in an AMC on academic advancement. In the alpha phase, while an AI-enhanced Microsoft Power Platform could generate advice for faculty, the features and functionality needed improvement, several of which could not be implemented through the platform. Users requested the chatbot to improve its ability to interpret initial queries to reduce the need for rephrasing. Users also required links to policy documents and relevant resources to verify and explore further. Although the tool was effective in saving users' time compared to seeking assistance from administrative staff, it should provide contacts to dedicated support personnel if it was unable to answer queries. Users also required the scope of the tool to be broadened to include academic development pathways and human resources information. This enabled users to solicit advice from a single source instead of searching in different chatbots for what they felt were related scopes – academic and human resource development.

The custom GPT included functionalities requested from the alpha tests - hyperlinks to policy documents, introduction to the chatbot and prompt questions to start conversations, improved system prompt guided by OpenAI, departmental contacts, and additional repository documents including the AMC's public websites to deliver holistic knowledge and answers to users. Thus, beta users felt the answers were clearer, relevant and required less querying effort. The percentage of promoters more than doubled, with large improvements in the tool's ability to resolve queries and its perceived empathy. Improvements in the system prompt strategy in the beta version likely contributed to a more engaging user experience and enhanced human–computer interaction.

As accuracy and reliability of the tool were of primary importance for trust and acceptance in the chatbot [17], the temperature setting was fixed at zero and 0.3 in the alpha and beta versions respectively [18]. This ensured deterministic responses, i.e. when information was available, it was

provided; when unavailable, Ask ADD clearly indicated that the requested information was not in its repository. To enhance helpfulness in the beta version, the tool included relevant email contacts, a feature noted and appreciated by users. Notably, users requested a more personalized experience through specific feedback and gap assessment upon evaluating their CV.

Limitations

The successful development of this tool was made possible by a team with diverse expertise, including professionals in Power Apps Platform development, artificial intelligence, academic promotion and tenure processes within an AMC, and mixed-methods research. All but one participant from the alpha testing phase participated in beta testing, enabling direct comparison of user feedback between phases and strengthening the validity and credibility of the findings.

Nonetheless, this study was based on platforms and tools that are currently available to and shareable among university staff. While this facilitates scalability within the institution, it is possible that integrating other tools not currently accessible or shareable among staff could further enhance the tool's performance. Due to the need for timely insights and operational constraints, we recruited a purposive sample of 35 participants which included representatives from clinical, tenure, research and education tracks faculty, administrative staff, program and institutional leaders, and other health professions. More in-depth findings could have been identified in a larger-scale study.

Conclusion

In summary, our study revealed that users valued a chatbot that provides accurate, clear, and contextually relevant responses, with minimal need for query rephrasing. Users preferred seamless access to policy documents, direct contacts for unresolved queries, and a comprehensive platform that includes academic advancement and related human resource information in addition to P&T

policies. Users also advocated for a more personalized experience through feedback and CV gap assessment. Functionality, responsiveness, accuracy, comprehensive scope, and personalization are important for adoption of a chatbot to meet faculty needs in an academic medical center.

To the best of our knowledge, this is the first chatbot specifically designed to support faculty advancement at an AMC. Although tools, such as custom GPTs, are already procured for enterprise use at the university, we applied them innovatively to reduce the time faculty members spend searching for P&T policy information. Our approach demonstrates a scalable, replicable model that other universities and AMCs can adopt to support academic career progression.

Acknowledgments

The authors wish to thank the National University of Singapore Digital Enablement, Artificial Intelligence, and the Office of Data and Intelligence teams. We are grateful to Mr. Peng Wei, Mr. Tan Keok Tay, Mr. Manish Sridhar, Mr. Ku Wee Kiat, Mr. Irving Kwok, and Mr. Voycevestern Peh Jia Wei for their support and invaluable contributions. We also wish to thank Ms. Kelly Cheang and Ms. Tricia Wong of the Appointments and Development Department of Duke-NUS Medical School for their assistance with research.

References

1. Yeo MM, Lim H, Kumar A, Thompson AW. Calculator for predicting the probability of faculty promotion in an academic medical centre. BMJ Lead. 2023;8: 203-207.

2. Microsoft Power Apps – Build Apps with AI | Microsoft [Internet]. [cited 2025 Jan 27]. Available from: https://www.microsoft.com/en-us/power-platform/products/power-apps

3. Azure OpenAI Service | Microsoft Azure [Internet]. [cited 2025 Jan 27]. Available from: https://azure.microsoft.com/en-gb/products/ai-services/openai-service/

4. Rosyadi MI, Kustiawan I, Tetehfio EO, Joshua Q. The role of AI In vocational education: a systematic literature review. J Vocat Educ Stud. 2023;6(2):244–63.

5. Duan J, Wu S. Beyond traditional pathways: leveraging generative AI for dynamic career planning in vocational education. Int J New Dev Educ. 2024;6:24–31.

6. Bankins S, Jooss S, Restubog SLD, Marrone M, Ocampo AC, Shoss M. Navigating career stages in the age of artificial intelligence: a systematic interdisciplinary review and agenda for future research. J Vocat Behav. 2024;153:104011.

7. Chang CN, Hui J, Justus-Smith C, Wang TW. Navigating STEM careers with AI mentors: a new IDP journey. Front Artif Intell. 2024;7:1461137.

8. Mehta N, Nielsen C, Zack A, Christensen T, Isaacson JH. Creating custom GPTs for faculty development: an example using the Johari Window and Crucial Conversation frameworks for providing feedback to struggling students. Med Teach. 2025;1–3.

9. Masters K, Benjamin J, Agrawal A, Macneill H, Tyson M, Mehta N, et al. Twelve tips on creating and using custom GPTs to enhance health professions education. Med Teach. 2024;1–5.

10. Zickar MJ, Keith MG. Innovations in sampling: improving the appropriateness and quality of samples in organizational research. Annu Rev Organ Psychol Organ Behav. 2023;10:315–37.

11. Camacho M, Perramon J, Puig-Bosch X, Dang VN, Díaz O, Lekadir K. Stakeholder engagement: the path to trustworthy AI in healthcare. Trust AI Med Imaging. 2025;471–95.

12. Vu HTM, Lai VTH, Tran HQ, Ghozaly S, Khishigjargal U, Enkh-Amgalan S. Exploring the impact of AI chatbots on customer satisfaction. Int J Creat Innov Res All Stud. 2022;4:62–9.

13. Reichheld FF. The one number you need to grow. Harv Bus Rev. 2004;81:124.

14. Wee R, Abdullah S. Executive functioning, study skills, and dyslexia - Examining the effectiveness of an online programme for upper secondary and post-secondary students. Asia Pacific J Dev Differ. 2023;10:65–97.

15. Swain J. A hybrid approach to thematic analysis in qualitative research: using a practical example. In: Sage research methods cases part 2. Thousand Oaks: Sage Publications, Ltd.; 2018.

16. OpenAI. Prompt engineering - Enhance results with prompt engineering strategies. [Internet]. [cited 2025 Feb 10]. Available from: https://platform.openai.com/docs/guides/prompt-engineering

17. Kaur D, Uslu S, Rittichier KJ, Durresi A. Trustworthy artificial intelligence: a review. ACM

Comput Surv. 2023;55(2).

18. Davis J, Bulck L Van, Durieux BN, Lindvall C. The temperature feature of ChatGPT: modifying creativity for clinical research. JMIR Hum Factors. 2024;11:1–4.

Supporting information

S1 Appendix. Alpha Version Using Power Apps and Power Automate Flow Including Code

S2 Appendix. User Feedback Questionnaire

S3 Appendix. Beta Version using Enterprise Custom OpenAI GPT-4o on Microsoft Azure Including System Prompt

S4 Appendix. Additional Feedback from the Alpha and Beta Tests

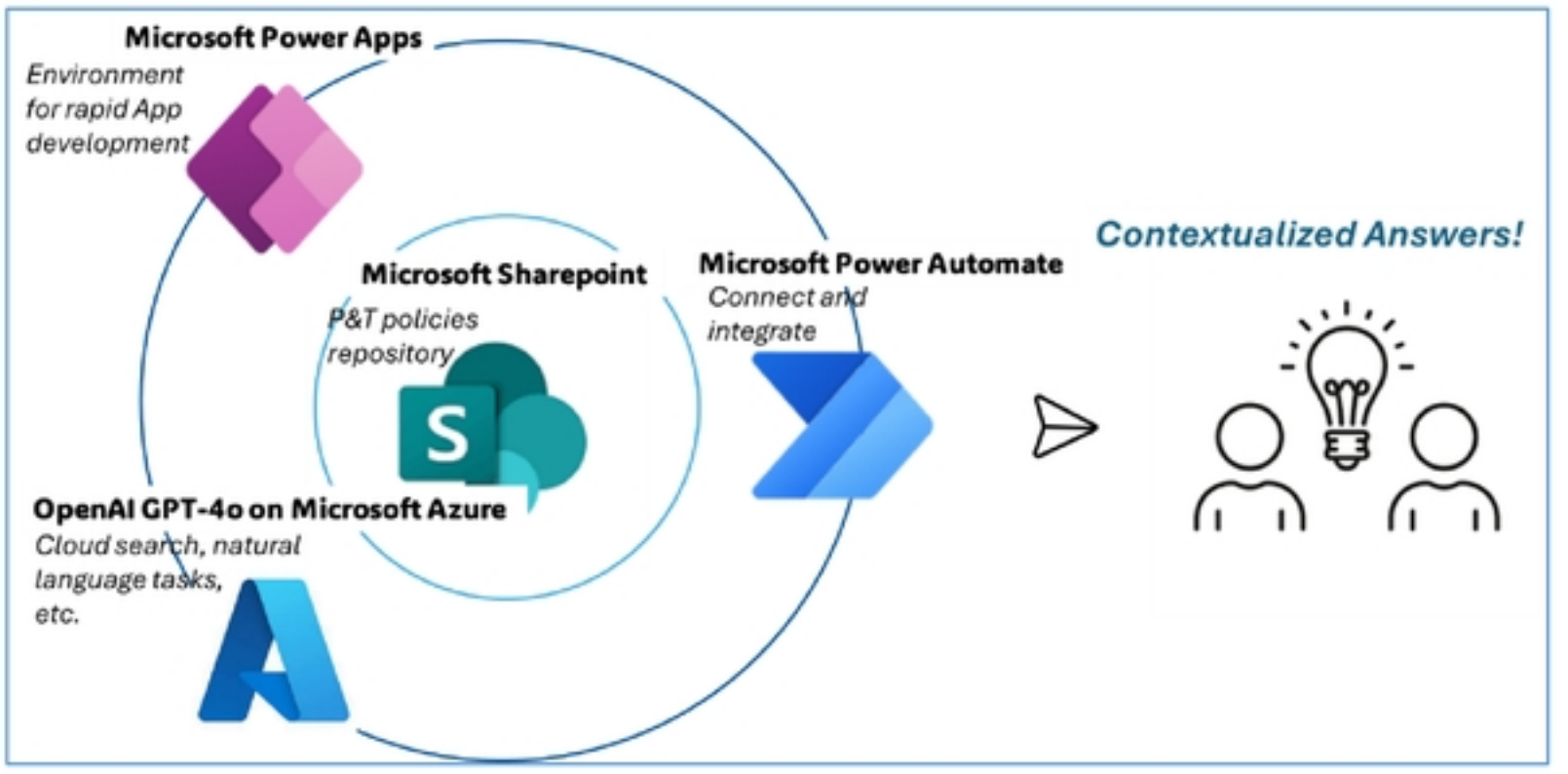


Figure 1

Figure 1

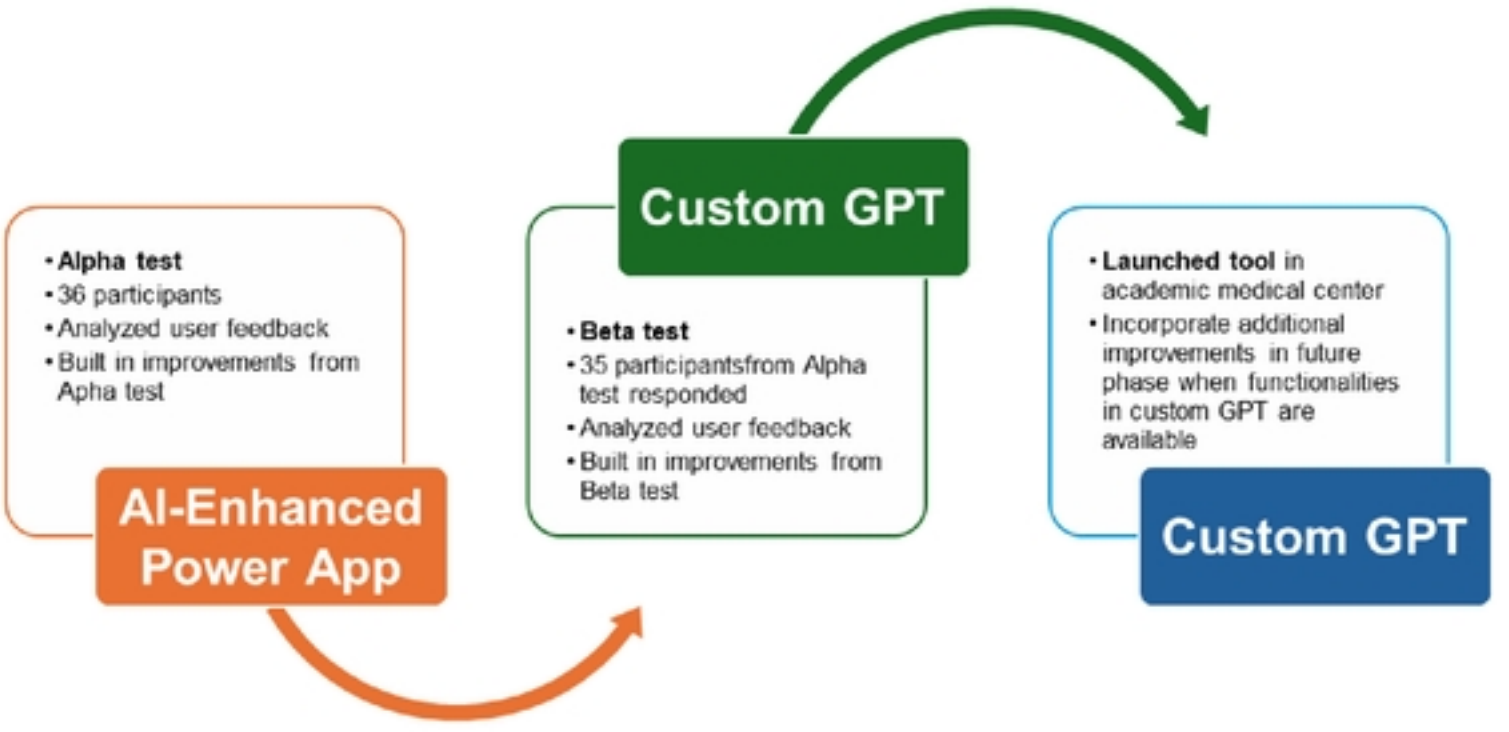
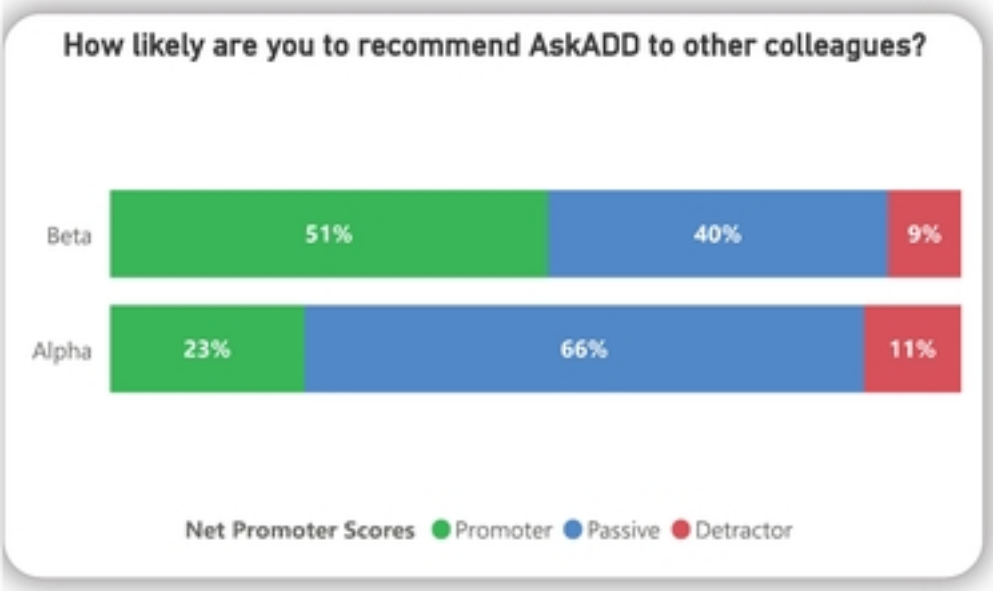
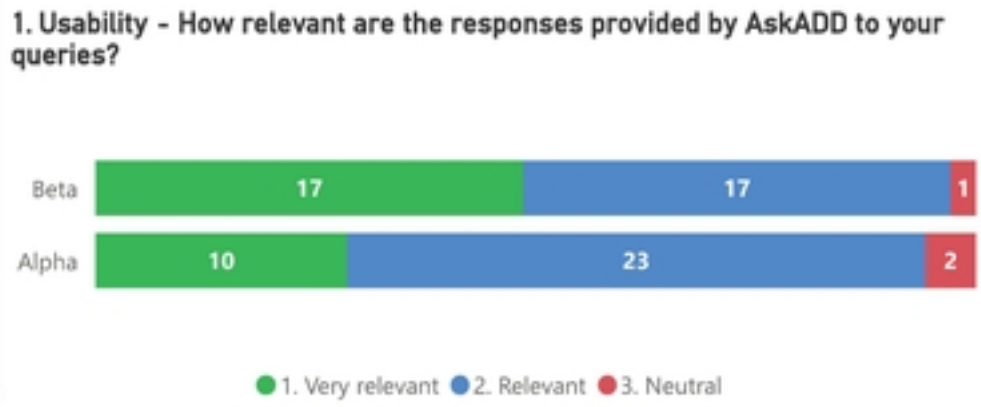
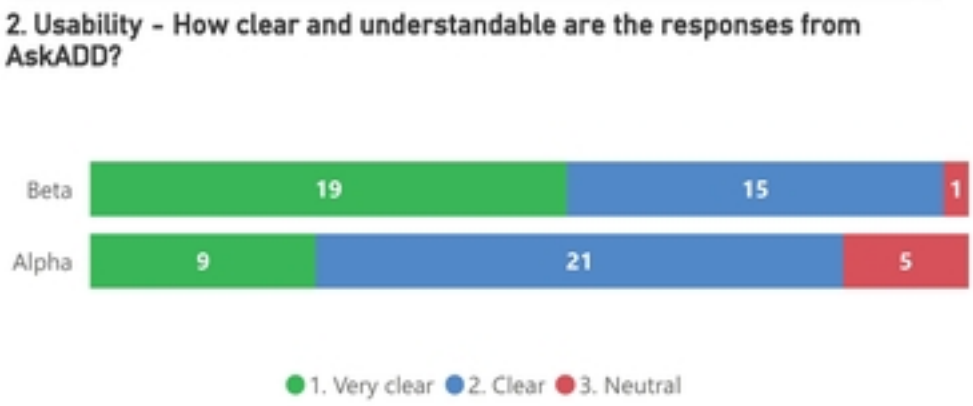


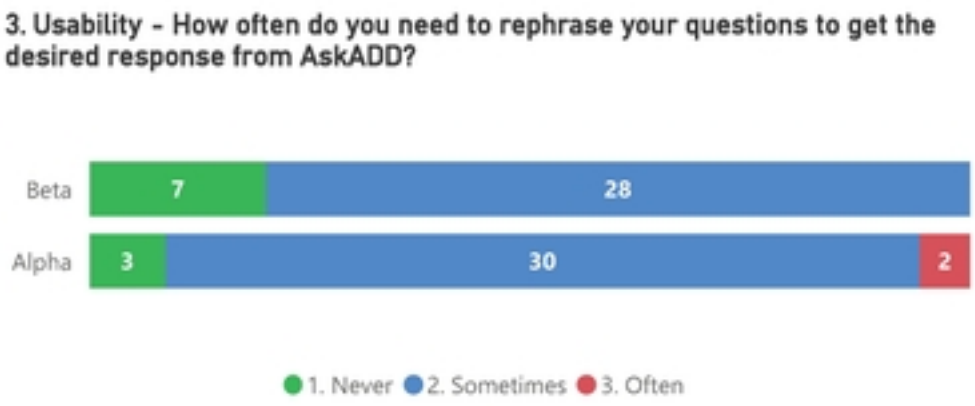
Figure 2

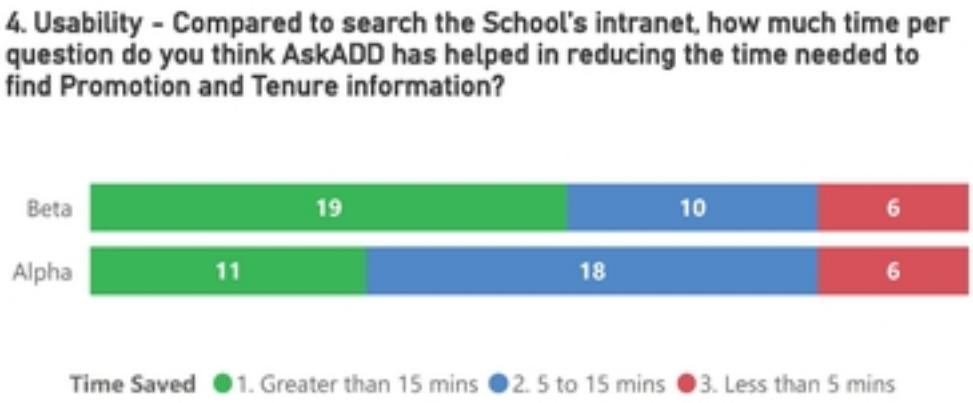
Figure 2

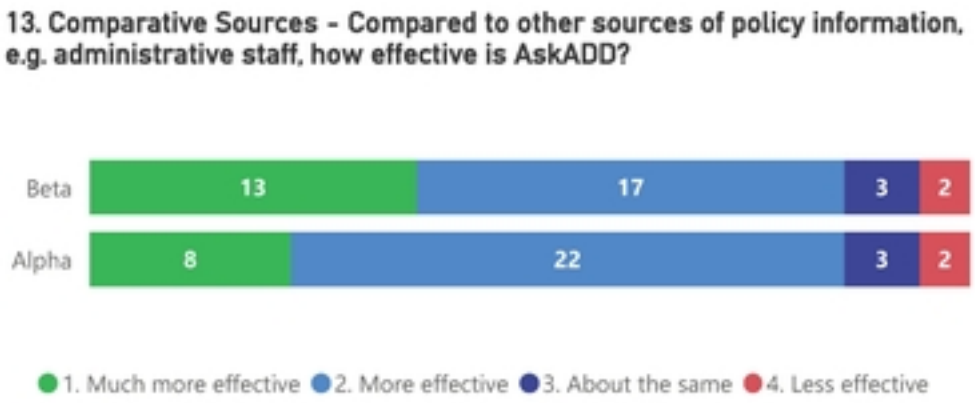












Number of Users Who Felt AskADD Performed Better in Beta than in Alpha and Why

Figure 3

Figure 3