



Delft University of Technology

**Document Version**

Final published version

**Licence**

CC BY

**Citation (APA)**

Sergeyuk, A., Koshchenko, E., Zakharov, I., Bryksin, T., & Izadi, M. (2026). Developer Needs and Feasible Features for AI Assistants in IDEs. In *Proceedings - 2026 IEEE/ACM 48th International Conference on Software Engineering: Software Engineering in Practice, ICSE-SEIP 2026* (pp. 406-416). Association for Computing Machinery (ACM). <https://doi.org/10.1145/3786583.3786883>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.  
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

**Sharing and reuse**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

*This work is downloaded from Delft University of Technology.*

# Developer Needs and Feasible Features for AI Assistants in IDEs

Agnia Sergeyuk\*  
JetBrains Research  
Belgrade, Serbia  
Delft University of Technology  
Delft, Netherlands  
agnia.sergeyuk@jetbrains.com

Ekaterina Koshchenko\*  
JetBrains Research  
Amsterdam, Netherlands  
ekaterina.koshchenko@jetbrains.com

Ilya Zakharov  
JetBrains Research  
Belgrade, Serbia  
ilia.zaharov@jetbrains.com

Timofey Bryksin  
JetBrains Research  
Limassol, Cyprus  
timofey.bryksin@jetbrains.com

Maliheh Izadi  
Delft University of Technology  
Delft, Netherlands  
m.izadi@tudelft.nl

## Abstract

Despite the increasing presence of AI assistants in Integrated Development Environments (IDEs), it remains unclear what different groups of developers actually need from these tools and which features are likely to be implemented in practice. To investigate this gap, we conducted a two-phase study. First, we interviewed 35 professional developers from three user groups (Adopters, Churners, and Non-Users) to uncover unmet needs and expectations. Our analysis revealed five key areas of need distinctly distributed across practitioners' groups: *Technology Improvement*, *Interaction*, and *Customization*, as well as *Simplifying Skill Building*, and *Programming Tasks*. We then examined the feasibility of addressing selected needs through an internal prediction market involving 102 practitioners. The results demonstrate a strong alignment between the developers' needs and the practitioners' judgment for features focused on implementation and context awareness. However, features related to proactivity and maintenance remain both underestimated and technically unaddressed. Our findings reveal gaps in current AI support and provide practical directions for developing more effective and sustainable in-IDE AI systems.

## CCS Concepts

• **Human-centered computing** → **User studies**; • **Computing methodologies** → *Artificial intelligence*; • **Software and its engineering** → *Integrated and visual development environments*.

## Keywords

Human-AI Collaboration, Artificial Intelligence, Generative AI, LLMs, Intelligent Assistants, Integrated Development Environment

### ACM Reference Format:

Agnia Sergeyuk, Ekaterina Koshchenko, Ilya Zakharov, Timofey Bryksin, and Maliheh Izadi. 2026. Developer Needs and Feasible Features for AI Assistants in IDEs. In *2026 IEEE/ACM 48th International Conference on Software Engineering (ICSE-SEIP '26)*, April 12–18, 2026, Rio de Janeiro, Brazil. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3786583.3786883>

\*Both authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution 4.0 International License. *ICSE-SEIP '26, Rio de Janeiro, Brazil*

© 2026 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-2426-8/26/04  
<https://doi.org/10.1145/3786583.3786883>

*Engineering (ICSE-SEIP '26)*, April 12–18, 2026, Rio de Janeiro, Brazil. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3786583.3786883>

## 1 INTRODUCTION

The increasing integration of Artificial Intelligence (AI) tools in software development is transforming the way developers interact with their Integrated Development Environments (IDEs) [17, 22, 26]. These AI-powered assistants facilitate tasks like code generation, debugging, and optimization with the goal of boosting productivity and reducing the cognitive load associated with repetitive programming tasks [1, 2, 10, 28]. As AI-based systems evolve rapidly, teams that build tools for tech creators face two key challenges: understanding how coding specialists actually use AI inside their IDEs and deciding which features are worth investing in.

However, actionable insights into in-IDE human-AI experience (HAX) are still limited. While recent studies have explored developers' general expectations for AI-assisted tools, few focus specifically on in-IDE interactions, particularly the successes and challenges developers face and their needs with these tools [28]. Even fewer studies consider differences across user groups: those who adopted AI into their workflow, those who abandoned it, and those who have never tried it. These distinctions are important for informing development and marketing strategies [25]. Moreover, these studies often focus on usage patterns or user perceptions without connecting them to concrete feature development decisions. As a result, we know little about how specific user needs translate into buildable and valuable AI features.

To address existing knowledge gaps, we conducted a multi-method study in two phases: qualitative semi-structured interview [19] and prediction market [32]. First, we interviewed 35 tech creators across three usage groups (Adopters, Churners, and Non-users) to gain qualitative insights into their experiences, challenges, and unmet needs with current AI features. This phase reveals unmet needs and expectations across five areas: *Technology Improvement*, *Interaction*, *Customization*, *Simplifying Skill Building*, and *Programming Tasks*. Yet, understanding user needs alone does not directly translate into feasible product development strategies. Prioritizing feature development requires additional insight into both technical feasibility and organizational readiness. To achieve this, we complemented interviews with a prediction market. Participants predicted which features would be implemented, using collective judgment

to estimate outcomes. We asked 102 internal practitioners to predict the likelihood of implementing 12 user-desired features. This allows us to integrate user-driven insights with practical considerations from the very people who shape our product development. Our results reveal that internal practitioners correctly anticipate the implementation of features around AI in Implementation and Context-aware Technology. In contrast, features requiring proactive behavior and maintenance support remain undervalued and underdelivered. These findings highlight a misalignment between developer needs and current feasibility assessments, pointing to concrete opportunities for more helpful and forward-looking AI assistance in IDEs.

With this work, we aim to provide industry practitioners and researchers with structured insights to guide product decisions. Specifically, offer:

- A categorization of user needs for in-IDE HAX to outline requirements to make AI tools more useful;
- A comparison of perspectives across Adopters, Churners, and Non-Users, revealing group-specific concerns that influence adoption and retention, such as trust and technical issues among Churners and ethical reservations and skill development concerns among Non-Users.
- A mapping of the externally voiced needs to internal perceptions of feasibility, identifying areas of alignment (e.g., implementation and context-awareness) and areas of mismatch (e.g., proactivity and maintenance support).

We are the first to triangulate detailed user needs with internal feasibility estimates using prediction markets. This approach enables organizations to better prioritize AI features grounded in what users want and what teams can build.

## 2 BACKGROUND

AI-powered programming assistants have become increasingly common, with tools such as GitHub Copilot, Cursor, and JetBrains AI integrated into developer workflows [1, 2, 10]. While their technical capabilities continue to evolve, research into how developers experience and evaluate such tools has also grown.

Recent studies have investigated what professional developers need and expect from AI-powered programming assistants [20, 21, 23, 27, 29, 31, 33]. Across contexts, developers consistently seek tools that reduce repetitive effort and provide relevant suggestions without introducing new debugging burdens [20, 29, 33]. Effective assistance is expected to align with the current task, understand code context, and support natural language interaction [20, 21, 30]. Transparency and control are also emphasized, including the ability to steer and correct outputs and control context [20, 23]. Several studies highlight that experienced users tend to expect more than raw generation [21, 23]. For example, tools are valued when they adapt to user intent and surface relevant, task-specific information [11, 20, 21]. At the same time, integration challenges persist. Developers often report friction when tools disrupt their workflow or fail to fit specific environments [27, 29]. Moreover, non-functional requirements such as performance and scalability may be harder to maintain [27]. These insights point to the need for assistants that can flexibly support diverse roles, offer customization, and operate within existing development processes [20, 23, 27].

However, most studies examine developer interaction with AI in broad settings, without addressing the unique characteristics of IDE-based workflows. This distinction is important. According to large-scale industry reports, up to 97% of developers rely on IDEs for software tasks [3]. In-IDE AI assistants operate within a structured environment that offers richer contextual signals such as file scope, project layout, and cursor history. These tools have the potential to deliver more context-aware support while minimizing context switching. Yet, their design introduces new constraints, including the need to remain non-intrusive and responsive in real time. Therefore, in-IDE AI tools shape distinct user needs, expectations, and challenges that warrant focused investigation.

Moreover, existing studies tend to treat developers as a single user group, overlooking differences across adoption stages [28]. However, industry practice suggests that design and feature prioritization must account for the heterogeneous needs of distinct groups: Adopters, Churners, and Non-Users [4, 25]. These groups not only engage (or not) with AI tools differently, but also hold divergent mental models, trust levels, and expectations, which may lead to different product requirements or marketing strategies. Recognizing these “personas” is necessary to design AI assistants that support retention, re-engagement, or onboarding.

At the same time, even when user needs are clearly articulated, translating them into product decisions remains a challenge. Feature roadmaps must also reflect internal constraints such as technical feasibility, resource allocation, and infrastructure readiness. Traditional prioritization techniques, such as expert panels and surveys, often fall short by providing limited collective insights or biased individual opinions [6]. An alternative approach, prediction markets, has been shown to effectively aggregate collective expert judgment to forecast the likelihood of specific outcomes [32].

A prediction market is a structured method wherein participants place bets on future events, effectively leveraging the “wisdom of the crowd” to estimate probabilities accurately [5, 32]. Previously, prediction markets have been successfully used to forecast a broad range of outcomes from presidential votes [7], to replicability of scientific research [12], and the future of real products on the market [8]. Google has been a pioneer in using internal prediction markets for software engineering and product development decisions. Research by Cowgill and Zitzewitz [9] analyzed internal prediction markets of several companies including Google, finding they produced more accurate forecasts than experts about company-specific events like product launch dates and business metrics. While prediction markets have shown promise in corporate technology settings, academic research on their application to software engineering research and practice remains limited. This gap represents a significant opportunity for HCI and software engineering research to examine how prediction markets can be adapted for the specific technical and social contexts of modern software development.

Thus, this paper combines qualitative interviews and prediction market, integrating rich user insights with realistic assessments from internal stakeholders. This dual perspective enables a grounded understanding of both user needs and internal feasibility, offering a practical contribution to the design and deployment of in-IDE AI assistants. We address the following research questions:

- **RQ1.** What are developers' needs when using in-IDE AI tools?
- **RQ2.** How do these needs differ between Adopters, Churners, and Non-Users?
- **RQ3.** To what extent do practitioners acknowledge and correctly anticipate which AI features inspired by user needs are likely to be implemented?

### 3 UNDERSTANDING DEVELOPER NEEDS

In this section, we examine how developers experience AI inside their IDEs when coding. Our goal is to provide a structured understanding of what developers need from AI assistance, what challenges they face, and how these vary across adoption groups. We address **RQ1** and **RQ2** through qualitative interviews.

#### 3.1 METHOD

The study was conducted in May-June 2024 in line with our institutional ethical standards, adhering to the values and guidelines outlined in the ICC/ESOMAR International Code [16].

**3.1.1 Recruitment.** We recruited participants from users of JetBrains, a software engineering company that also develops AI-based tools for coding, to ensure a consistent experience with development environments. Note that with that, participants could use any in-IDE AI assistant. Invitations were sent via email to consenting subscribers and promoted through the authors' social media.

Respondents completed a qualification survey. It was designed to assign them to one of three groups based on a question about AI tool experience: current users were classified as *Adopters*, former users as *Churners*, and those who never used such tools as *Non-users*. We chose these three groups to capture various experiences with AI tools in IDEs [25]. These groups are referred to using the prefixes A, C, and N, respectively, in participant identifiers throughout the paper. In the survey, we collected data on job roles, experience levels, IDEs used, AI tools tried for coding, and duration of AI usage. Out of 380 responses, we iteratively invited participants until data saturation was reached [15], which means that no new themes or insights emerged from the last interviews. As a result, we conducted 35 interviews: 15 Adopters, 12 Churners, and 8 Non-Users. Most of the participants were developers (32) with overlapping roles: 9 team leads, 8 architects, and 5 DevOps engineers. The sample included 28 seniors, 5 mid-level practitioners, and 2 juniors, with experience ranging from under 5 years (9) to over 11 years (16). Participants chose a reward between a USD 100 Amazon gift card or an equivalent-value pack of all JetBrains products.

**3.1.2 Data Collection.** The interviews followed a semi-structured script developed by the first two authors, reviewed by certified UX researchers at JetBrains, and piloted with four internal participants [19]. The script included open-ended questions on AI experience, utilized tools, challenges and successes with in-IDE HAX, unmet needs, desired features, and visions for AI in coding.

All interviews were conducted remotely by one of the first two authors, lasting 71 minutes on average (range: 46-92). Each interview followed a common structure covering challenges, successes, needs, and future expectations, with question framing adapted to

each group's AI experience while maintaining cross-group comparability. For Adopters, the questions focused on their ongoing use of AI tools, exploring typical workflows, challenges, and successes. Churners, those who have previously used AI tools but then stopped using them, reflected on their initial adoption and further disengagement, identifying reasons for it and missed opportunities. Non-Users were asked about their reasons for not adopting AI, any potential barriers and possible motivations that could encourage them to begin using these tools. This enabled group-specific depth while maintaining cross-group comparability during analysis.

**3.1.3 Data Analysis.** All interviews were transcribed using Condens<sup>1</sup>. We applied a qualitative content analysis [13, 14] to identify user-reported successes, challenges, unmet needs, and future expectations for in-IDE AI. The first two authors reviewed the first transcript together to agree on the analysis and the code-deriving process. They then went into an iterative inductive process, where they independently assigned *Codes* to quotes from the transcripts. Eventually, they compiled a list of 161 codes, met again to add or remove codes, and resolved any disagreements until full consensus was reached. The final versions of the codes were then assigned to the corresponding quotes. The codes were then grouped into *Thematic groups* representing key areas of need. These were organized into broader *Topics* that describe the categories of in-IDE HAX needs. For example, the following quote:

*"The most useful thing you could do with AI would be a summary of a Pull Request."* — N13

was coded as "Code Review Support", assigned to the group "Code Optimization", under the topic "Simplifying Programming Tasks". Following established practices, we identified emergent patterns across user groups, recognizing that in qualitative research, the presence and salience of themes matter more than frequency counts.

All materials used for data gathering, together with a full demography table of participants and a summarized codebook, are included in Supplementary Materials [18]. Full interview transcripts cannot be shared due to the company's confidentiality policy.

#### 3.2 FINDINGS

**3.2.1 Categorization of Developer Needs.** Figure 1, presents our categorization of developer needs regarding Human-AI Experience within IDEs, based on a thematic analysis of 35 interviews. It includes five overarching topics: Technology Improvements, Simplifying Programming Tasks, Technology Interaction, Technology Customization, and Simplifying Skill Building.

**Technology Improvements.** Participants identified several requirements for improvements of AI tooling for coding, beginning with *accessibility* across infrastructural, legal, and experiential boundaries. They proposed specific solutions, including natural language programming interfaces, universal language support, and modular pricing models, to reduce barriers in regions with limited connectivity or restrictive regulations. Beyond basic accessibility,

<sup>1</sup>Condens: a user research repository for analyzing qualitative data. <https://condens.io/>

| Technology Improvements                                                                |                                                                                             |                                                                                            |                                                                                                      |                                                                                                                 |                                                                                                          |
|----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
| <b>Accessible Technology</b>                                                           | <b>Proactive Technology</b>                                                                 | <b>Reliable Technology</b>                                                                 | <b>Up-to-date Technology</b>                                                                         | <b>Ethical Technology</b>                                                                                       | <b>Private Technology</b>                                                                                |
| Craft technologies that are both affordable and adaptable to various constraints       | Enhance efficiency by reducing users' manual effort and preemptively solving issues         | Ensure the robustness and accuracy of AI systems and increase users' trust in AI solutions | Ensure AI systems are current and competitive by integrating the latest advancements and updates     | Balance innovation with conscientious practices, fostering technology that is both effective and morally sound  | Focus on safeguarding user privacy and enhancing data security                                           |
| <i>Separate feature purchase</i>                                                       | <i>Proactive optimization, AI performing actions</i>                                        | <i>Output clarifications, Providing sources</i>                                            | <i>SOTA tooling, Up-to-date data</i>                                                                 | <i>Rules and restrictions, Automated regulation</i>                                                             | <i>On-premise, Local AI</i>                                                                              |
| Technology Improvements                                                                |                                                                                             |                                                                                            | Simplifying Programming Tasks                                                                        |                                                                                                                 |                                                                                                          |
| <b>Secure Technology</b>                                                               | <b>Aligned Technology</b>                                                                   | <b>Context-aware Technology</b>                                                            | <b>AI Support in SDLC</b>                                                                            | <b>Navigating the Codebase</b>                                                                                  | <b>AI in Ideation</b>                                                                                    |
| Fortify AI systems against vulnerabilities and ensure their integrity                  | Ensure that AI-generated solutions are coherent, non-redundant and consistent with codebase | Enhance AI's ability to understand and integrate various levels of context                 | Aim for universal adoption to handle tedious and well-known tasks all over the development lifecycle | Facilitate a deeper comprehension of code and project structure to aid in navigating and understanding codebase | Enhance the creative and brainstorming process by providing effective ways to generate and explore ideas |
| <i>Annotating AI code</i>                                                              | <i>Library suggestions, Better commit messages</i>                                          | <i>Manually set context, Project-wide awareness</i>                                        | <i>Universal programming assistance</i>                                                              | <i>Onboarding support, Database analytics</i>                                                                   | <i>Creative AI, Cross-model prompting</i>                                                                |
| Simplifying Programming Tasks                                                          |                                                                                             |                                                                                            |                                                                                                      |                                                                                                                 |                                                                                                          |
| <b>AI in Implementation</b>                                                            | <b>AI in Testing</b>                                                                        | <b>AI in Maintenance</b>                                                                   | <b>Code Optimization</b>                                                                             | <b>External System Integration</b>                                                                              |                                                                                                          |
| Support development workflows smoothing the process from concept to implementation     | Enhance test generation and ensure robust, thorough evaluation of software systems          | Simplify maintenance tasks to optimize and enhance ongoing maintenance and updates         | Enhance code efficiency and quality to ensure high-performance codebases                             | Incorporate AI into broader workflows and external systems, enhancing development connectivity                  |                                                                                                          |
| <i>Multi-file generation, Project set-up assistance</i>                                | <i>Coverage report, UI testing</i>                                                          | <i>Repository, Dependency management</i>                                                   | <i>Code evaluation, Code review support</i>                                                          | <i>AI in CI/CD Pipelines, Out-IDE AI access</i>                                                                 |                                                                                                          |
| Technology Interaction                                                                 |                                                                                             | Technology Customization                                                                   |                                                                                                      | Simplifying Skill Building                                                                                      |                                                                                                          |
| <b>Non-Interruptive Technology</b>                                                     | <b>Internal System Integration</b>                                                          | <b>Adjustable Technology</b>                                                               | <b>Optimizing Workflow</b>                                                                           | <b>User Education</b>                                                                                           | <b>Learning Support</b>                                                                                  |
| Maintain a smooth workflow by reducing interruptions and addressing interaction issues | Enhance the cohesion and interaction of AI systems within development environments          | Enhance functionality through configurable AI solutions that adapt to diverse requirements | Create a cohesive work environment that supports organization and facilitates daily tasks            | Equip users with the knowledge and tools needed to effectively interact with AI systems                         | Support the acquisition of new concepts and skills, making the learning process more empowering          |
| <i>Stable AI functionality, Voice interaction</i>                                      | <i>Anywhere-in-IDE AI access</i>                                                            | <i>Configurable AI, AI action customization</i>                                            | <i>Daily ecosystem summarizer</i>                                                                    | <i>Shareable prompt library</i>                                                                                 | <i>Step-by-step output revealing</i>                                                                     |

Figure 1: The Categorization of Users' Needs in HAX Within IDEs With Descriptions

participants envisioned AI systems transitioning from reactive assistants to **proactive** collaborators. This shift would enable AI to identify potential issues during development:

*"It would be cool if it could search for problems along the way. It could say 'Did you realize you just created an out-of-bounds condition?'" — A44*

Such proactive capabilities, however, require fundamental improvements in system **reliability** and integration of the **latest advancements** in the field. Participants emphasized that unreliable systems create additional verification overhead:

*"If I felt the AI system was really reliable, it would have saved me time because I wouldn't need to double-check and could just apply the right solution right away." — C159*

These technical improvements must be accompanied by robust **ethical** frameworks. Participants recognized AI's potential and the corresponding responsibility this creates:

*"AI, like anything with the potential to destroy jobs, carries a huge responsibility. I really hope that the community and organizations invest heavily in getting it right. Otherwise, recovering might be really difficult." — A37*

Ethical considerations extend to **privacy** and **security** requirements. Participants strongly favored on-premise or local AI solutions to ensure data confidentiality and comply with organizational or national restrictions. They emphasized the need for vulnerability prevention and maintaining code integrity, including mechanisms to clearly identify AI-generated code for inspection and verification. Finally, participants stressed that all improvements must support better **alignment** with existing development practices. **Context awareness** emerged as essential for generating relevant and accurate code. They specifically requested mechanisms to manually set, view, and edit the context provided to AI assistants. Current systems' lack of transparency in context construction creates frustration and undermines user confidence in AI-generated outputs.

| Thematic Group              | Successes | Challenges | Needs | Thematic Group              | Successes | Challenges | Needs |
|-----------------------------|-----------|------------|-------|-----------------------------|-----------|------------|-------|
| Accessible Technology       |           | ● ★ ▲      | ● ★   | AI Support in SDLC          | ● ★ ▲     |            | ● ★ ▲ |
| Proactive Technology        | ●         | ● ★        | ● ★ ▲ | Navigating the Codebase     | ● ★       |            | ● ★ ▲ |
| Reliable Technology         |           | ●          | ● ★ ▲ | AI in Ideation              | ● ▲       | ● ▲        |       |
| Up-to-date Technology       |           | ● ★ ▲      | ●     | AI in Implementation        | ● ★ ▲     | ●          | ● ★ ▲ |
| Ethical Technology          |           | ★ ▲        | ★     | AI in Testing               | ● ★       | ●          | ● ★ ▲ |
| Private Technology          |           | ● ★ ▲      | ● ★ ▲ | AI in Maintenance           | ● ★ ▲     | ●          | ● ★ ▲ |
| Secure Technology           | ★         | ★ ▲        | ★ ▲   | Code Optimization           | ● ★ ▲     | ●          | ● ★ ▲ |
| Aligned Technology          | ● ★       | ● ★ ▲      | ●     | External System Integration |           |            | ● ★ ▲ |
| Context-aware Technology    | ● ★ ▲     | ● ★ ▲      | ● ★ ▲ | Adjustable Technology       | ● ★       | ★          | ● ★ ▲ |
| Non-Interruptive Technology | ● ★ ▲     | ● ★        | ● ★ ▲ | Optimizing Workflow         | ●         |            | ● ★ ▲ |
| Internal System Integration |           |            | ● ★ ▲ | User Education              |           | ● ★ ▲      | ● ★ ▲ |
|                             |           |            | ● ★ ▲ | Learning support            | ● ★ ▲     |            | ● ★ ▲ |

Note: Symbols represent the participant groups: Adopters (●), Churners (★), and Non-Users (▲). The presence of a symbol indicates that group members mentioned the corresponding theme in the associated interview section.

Figure 2: Comparative Analysis of Thematic Group Mentions Across User Types.

**Simplifying Programming Tasks**. Participants identified AI's primary value in automating routine and repetitive activities across the **software development lifecycle** (SDLC). This automation would enable developers to redirect effort from low-level tasks toward more complex and creative work spanning ideation, implementation, testing, maintenance, and optimization phases. This vision extends beyond simple task automation to AI functioning as an intelligent collaborator:

*"I would say it should be like a colleague, someone who understands what I wrote and can guide or help me where I lack knowledge." — A26*

A critical component of this collaborative relationship involves AI-supported **codebase navigation**. Participants particularly valued AI's potential to help newcomers understand existing systems and adhere to established standards:

*"If you gave the AI the company documentation and code and then let it automatically detect what's needed, this would help newcomers follow the company's standards." — C94*

Participants also envisioned AI tools that seamlessly connect IDEs with external environments, including Continuous Integration and Continuous Deployment (CI/CD) pipelines, version control platforms, and support ticketing systems.

*"Right now, so many things come through email, then there's the support ticketing system, and, of course, our Git repositories. There could be ways to integrate all these even better." — C118*

**Technology Interaction**. Participants identified seamless **internal system integration** as fundamental to an effective AI tool, emphasizing the need for AI to interact efficiently with existing IDE features. Beyond that, participants sought **uninterrupted** experiences that enable flexible AI interaction modes throughout their development workflows. However, current implementations frequently fall short of this due to technical limitations that disrupt workflow continuity. These disruptions manifest in several forms: slow response times that break the natural rhythm of development, frequent authentication failures that force repeated logins, and functional limitations including copy/paste failures and conversation

dead-ends where AI systems become unresponsive. Such technical barriers undermine the seamless integration that participants consider essential for productive AI-assisted development.

**Technology Customization**. The interviewees stressed the importance of **adjustable** AI systems that offer a high degree of personalization and flexibility, allowing users to retain control as decision-makers. Key features included configurable AI, the ability to choose between different models, and the option to customize AI actions based on specific tasks or workflows. Participants also expressed a desire for task-specific setups that allow AI to adapt to unique user preferences and working styles. To support **work-flow optimization**, participants were also interested in tools that could help manage distractions, track time, and summarize daily activities within their development ecosystem. The ability to toggle time tracking, manage task tickets, and maintain a personal prompt library was highlighted as essential for staying organized and improving productivity.

*"At my job, I have around 20 tickets pending, and sometimes I forget which one I'm currently working on and what to do next, so I need something that can help me stay organized." — N14*

**Simplifying Skill Building**. The participants noted that AI tools often have a steep learning curve. They identified the need for better **user education**, particularly around effective prompt formulation. A shareable prompt library was seen as a possible valuable resource, as many reported difficulties in crafting prompts that yield useful results. Additionally, the unfamiliar interface of some AI systems contributed to these challenges. Several Non-Users mentioned that the primary reason they had not adopted AI tools was the time required to become proficient.

*"It's a useful tool if employed correctly. But there is a learning curve. If I figure out how to work with it efficiently, how to set up the right prompts, etc., it should speed up my work." — N84*

The participants also emphasized the role of AI as a **learning enabler** and mentor, helping developers acquire new skills and understand new concepts. Features such as step-by-step output reveal

were considered helpful in fostering long-term skill development through more interactive and adaptive guidance.

*“This way, you wouldn’t have to keep requesting new hints would be more interactive, allowing you to reveal the steps as you go and adjust based on what you need.” — A58*

**Answer to RQ1.** Developers require AI tools that are advanced, well-integrated, customizable, dedicated to simplifying programming tasks and skill building.

**3.2.2 Comparison of Needs’ Themes Among Groups.** We present a comparative analysis of the Thematic groups among Adopters, Churners, and Non-Users in Figure 2, focusing on participants’ success stories (or anticipated advantages for Non-Users), challenges (or barriers for Non-Users), and needs in the context of in-IDE HAX. The presence of a symbol indicates that at least one participant within that user group mentioned the corresponding theme during their interview. This binary presence/absence approach was chosen to capture the full spectrum of experiences and concerns across adoption stages, as even singular mentions may represent important but underexplored aspects of the user experience that warrant consideration in feature development.

The **Context-aware Technology** thematic group is uniquely mentioned by all three groups of participants as a success, a challenge, and a need, indicating the most ambiguous relationship to this aspect of technology. **Privacy** and **User Education** emerge as universal concerns and points of potential and urgent improvement, with all three groups reporting challenges and needs in this regard. All three groups of participants mention **Learning support** and **AI Support in various stages of SDLC** as success and need, signaling that the evolution of AI tooling is sufficient but still relevant in these areas. Notably, **AI in the Ideation** theme is not mentioned by any of the groups as a need, signaling that people are less likely to seek AI support at this stage of SDLC. On the other hand, participants mention **Internal** and **External System integration** only as a need, implying that this is a valuable attention point for tool builders.

It is interesting to note that participants usually mention in challenges and needs **Technology Improvements** -related themes, while **Simplifying Programming Tasks** -related themes are usually mentioned in successes and needs. This trend might be interpreted as a direction for tool builders to improve technology, making it more advanced, as opposed to widening applications, which are already successful enough.

**3.2.3 Comparison of Needs’ Codes Among Groups.** Further, we explore the similarities and differences of participants’ 78 specific needs at the Codes’ level, represented in Table 1.

All three groups expressed needs for comprehensive development support by AI (universal programming assistance, higher quality output, productivity enhancement, multi-file generation, system testing), enhanced tools’ context awareness (project-wide and business-context awareness, manually set context), infrastructure flexibility (on-premise AI, CI/CD pipeline integration, external system integration), proactive AI capabilities (code generation, optimization, autonomous actions), and improved user experience (suggestion preview, configurable AI, task automation, learning

**Table 1: User group needs across topics.**

| Topic                              | Code of Need                       | A | C | N |
|------------------------------------|------------------------------------|---|---|---|
| <b>Accessible Technology</b>       | Polyglot AI                        |   |   | ✓ |
|                                    | Separate feature purchase          | ✓ | ✓ |   |
|                                    | Reasonable pricing                 | ✓ | ✓ |   |
| <b>Proactive Technology</b>        | Proactive code generation          | ✓ | ✓ | ✓ |
|                                    | Proactive documentation            | ✓ | ✓ |   |
|                                    | Proactive optimization             | ✓ | ✓ | ✓ |
|                                    | Proactive debugging                | ✓ | ✓ |   |
|                                    | AI performing actions              | ✓ | ✓ | ✓ |
|                                    | File management                    | ✓ |   |   |
| <b>Reliable Technology</b>         | Library management                 |   |   | ✓ |
|                                    | Providing sources                  | ✓ |   |   |
|                                    | Output clarifications              | ✓ |   |   |
| <b>Up-To-Date Technology</b>       | Higher quality                     | ✓ | ✓ | ✓ |
|                                    | Sota tooling                       | ✓ |   |   |
| <b>Ethical Technology</b>          | Ethical technology                 |   |   | ✓ |
| <b>Private Technology</b>          | Privacy                            |   |   | ✓ |
|                                    | On-premise AI                      | ✓ | ✓ | ✓ |
|                                    | Local AI                           |   | ✓ | ✓ |
| <b>Secure Technology</b>           | Annotating ai-code                 |   | ✓ | ✓ |
|                                    | Security                           |   |   | ✓ |
| <b>Aligned Technology</b>          | Library suggestions                | ✓ |   |   |
|                                    | Generation by example              | ✓ |   |   |
| <b>Context-Aware Technology</b>    | Contextual awareness               | ✓ |   | ✓ |
|                                    | Project-wide context awareness     | ✓ | ✓ | ✓ |
|                                    | Business-context awareness         | ✓ | ✓ | ✓ |
|                                    | Manually set context               | ✓ | ✓ | ✓ |
|                                    | Viewable context                   | ✓ | ✓ |   |
|                                    | High level understanding           |   |   | ✓ |
| <b>AI Support In SDLC</b>          | Productivity boost                 | ✓ | ✓ | ✓ |
|                                    | Universal programming assistance   | ✓ | ✓ | ✓ |
| <b>Navigating The Codebase</b>     | Onboarding support                 | ✓ | ✓ | ✓ |
|                                    | Database analytics                 |   |   | ✓ |
|                                    | Diagram generation                 | ✓ | ✓ |   |
|                                    | Code explanations                  | ✓ |   |   |
| <b>AI In Implementation</b>        | Project set-up assistance          | ✓ | ✓ |   |
|                                    | Full stack generation              | ✓ | ✓ |   |
|                                    | Ui generation                      | ✓ | ✓ |   |
|                                    | Multi-file generation              | ✓ | ✓ | ✓ |
| <b>AI In Testing</b>               | Sketch2code                        |   |   | ✓ |
|                                    | Coverage report                    | ✓ | ✓ |   |
|                                    | System testing                     | ✓ | ✓ | ✓ |
|                                    | Ui testing                         | ✓ | ✓ | ✓ |
| <b>AI In Maintenance</b>           | Error handling                     |   |   | ✓ |
|                                    | Repository management              |   |   | ✓ |
|                                    | Database merging                   |   |   | ✓ |
|                                    | Dependency management              | ✓ | ✓ |   |
| <b>Code Optimization</b>           | Guideline enforcement              | ✓ |   |   |
|                                    | Optimization                       | ✓ |   | ✓ |
|                                    | Code review support                | ✓ | ✓ | ✓ |
| <b>External System Integration</b> | Code evaluation                    | ✓ | ✓ |   |
|                                    | Commit management                  | ✓ |   |   |
|                                    | AI in ci/cd pipelines              | ✓ | ✓ | ✓ |
|                                    | Out-side AI access                 | ✓ |   |   |
| <b>Non Interruptive Technology</b> | External system integration        | ✓ | ✓ | ✓ |
|                                    | Voice interaction                  | ✓ | ✓ |   |
|                                    | Stable AI functionality            |   |   | ✓ |
|                                    | Neuralink                          | ✓ |   |   |
| <b>Internal System Integration</b> | Suggestion preview                 | ✓ | ✓ | ✓ |
|                                    | In-line code explanation           |   | ✓ |   |
|                                    | Key mappings                       |   |   | ✓ |
|                                    | AI in code with me                 | ✓ |   |   |
| <b>Adjustable Technology</b>       | Anywhere-in-ide AI access          | ✓ | ✓ |   |
|                                    | Internal system integration        | ✓ |   | ✓ |
|                                    | Personalization                    | ✓ |   |   |
|                                    | Configurable AI                    | ✓ | ✓ | ✓ |
| <b>Optimizing Workflow</b>         | Choosing models                    | ✓ | ✓ |   |
|                                    | AI action customization            | ✓ |   |   |
|                                    | Toggleable time tracking           |   |   | ✓ |
|                                    | Daily ecosystem summarizer         | ✓ |   |   |
|                                    | Distraction management             |   |   | ✓ |
|                                    | Ticket management                  |   |   | ✓ |
| <b>User Education</b>              | Automation of repetitive tasks     | ✓ | ✓ | ✓ |
|                                    | Proactive workflow recommendations | ✓ | ✓ |   |
|                                    | Prompting support                  | ✓ | ✓ |   |
|                                    | Shareable prompt library           | ✓ |   |   |
| <b>Learning Support</b>            | User education                     |   | ✓ | ✓ |
|                                    | Step-by-step output revealing      | ✓ |   |   |
|                                    | Learning support                   | ✓ | ✓ | ✓ |

Note: A = Adopter, C = Churner, N = Non-User

support). The convergence on these needs suggests they represent core requirements for AI-powered development environments

**Adopters** demonstrated the broadest range of needs, expressing requirements across all identified categories. This comprehensive need profile reflects their active engagement with AI tools and suggests that continued innovation in advanced features will strengthen retention among this group. Drawing from their relatively extensive AI experience, uniquely identified needs are:

- Advanced tooling and management features (SOTA tooling, File management, Guideline enforcement, Code explanations),
- Enhanced alignment and reliability (Personalization, AI action customization, Library suggestions, Generation by example, Providing sources, Output clarifications),
- Deeper integration capabilities (Commit management, out-IDE AI access, AI in Code With Me, Neuralink),
- Workflow and learning optimization (Daily ecosystem summarizer, Step-by-Step Output Revealing, Shareable Prompt Library).

**Churners** expressed needs that align substantially with Adopters, sharing 36 common requirements across most topics. However, Churners uniquely emphasized:

- System stability (Stable AI functionality),
- Specific conceptual technical capabilities (Polyglot AI, High-level understanding),
- Infrastructure management (Library management, Repository management),
- Specialized interface features (In-line Explanation, Sketch2Code).

This pattern suggests that while Churners share Adopters' vision for comprehensive AI for code, their focus on stability and conceptual depth indicates that current implementations may lack the robustness and sophistication necessary to sustain engagement. Therefore, implementation quality rather than feature availability drives disengagement decisions.

**Non-Users** expressed needs in specific areas that represent potential adoption catalysts. Their unique requirements centered on foundational concerns including:

- Ethical technology, privacy protection, and security,
- Specialized functionality (Database merging, Key mappings),
- Productivity control features (Toggleable time tracking, Distraction management).

This pattern suggests that Non-Users might prioritize trust-building mechanisms as prerequisites for adoption, indicating that addressing ethical safeguards may be essential for expanding adoption beyond current users.

**Answer to RQ2.** Adopters seek advanced capabilities, Churners require conceptual improvements, and Non-Users need trust-building.

**Table 2: Prediction market results for proposed AI features.**

| Feature                  | Tool          | Resolved | Acc. | Prob. | Vol. | Liq. |
|--------------------------|---------------|----------|------|-------|------|------|
| Multi-file Generation    | Any Tool      | ✓        | ✓    | 98%   | 1946 | 570  |
|                          | JetBrains IDE | ✓        | ✓    | 92%   | 4288 | 610  |
| Project-wide Context     | Any Tool      | ✓        | ✓    | 82%   | 440  | 290  |
|                          | JetBrains IDE | ✓        | ✗    | 41%   | 2238 | 790  |
| Sketch2Code              | Any Tool      | ✓        | ✓    | 90%   | 1019 | 310  |
|                          | JetBrains IDE | ✗        | ✓    | 10%   | 500  | 270  |
| Proactive Docs           | Any Tool      | ✗        | ✓    | 33%   | 878  | 310  |
|                          | JetBrains IDE | ✗        | ✓    | 17%   | 1115 | 370  |
| Business Context         | Any Tool      | ✓        | ✗    | 22%   | 272  | 150  |
|                          | JetBrains IDE | ✗        | ✓    | 10%   | 190  | 130  |
| Real-Time Refactor       | Any Tool      | ✗        | ✓    | 19%   | 950  | 370  |
|                          | JetBrains IDE | ✗        | ✓    | 8%    | 420  | 250  |
| Proactive Optim.         | Any Tool      | ✗        | ✓    | 19%   | 650  | 310  |
|                          | JetBrains IDE | ✗        | ✓    | 8%    | 325  | 210  |
| Proactive Debug          | Any Tool      | ✗        | ✓    | 6%    | 750  | 210  |
|                          | JetBrains IDE | ✗        | ✓    | 8%    | 340  | 150  |
| Brain-computer Interface | Any Tool      | ✗        | ✓    | 4%    | 1241 | 490  |
|                          | JetBrains IDE | ✗        | ✓    | 1%    | 2150 | 510  |

Note: **Resolved**: feature launched (✓) or not (✗). **Acc.**: prediction correctness. **Prob.**: final market probability. **Vol.** and **Liq.**: participant engagement metrics.

## 4 ESTIMATING FEATURE FEASIBILITY VIA PREDICTION MARKET

Having identified developer needs across user groups (RQ1, RQ2), we now turn to RQ3: assessing which user needs and feature directions tool builders consider feasible and valuable and whether this reflects an accurate intuition about how the field is likely to evolve. To answer this, we conducted a prediction market “tournament” on JetBrains’ internal platform, where participants bet on whether specific AI features would be released within six months.

All materials used for data gathering at this phase, as well as the full demography table and a summarized report of the prediction market, are in Supplementary materials [18].

### 4.1 METHOD

**4.1.1 Materials.** Based on insights from interviews, we distilled an initial list of 38 potential AI-based IDE features [18] grounded in ideas and suggestions expressed by participants. This list was reviewed by seven experienced researchers from industry and academia: three specialists in human–AI experience, one assistant professor in computer science, and three senior software engineering researchers. All reviewers had more than five years of professional experience. Reviewers evaluated and discussed asynchronously each feature on three dimensions: existence in current tools, clarity of description, and suitability for prediction market assessment. Through this collaborative process, we selected 9 features for the prediction market.

These nine features were selected because they are representative of thematic groups that were mentioned by all three user groups (Adopters, Churners, and Non-Users) as both successes and needs. This cross-group presence indicates broad relevance and

high potential for user benefit, which in turn justified their perceived feasibility within the short-term horizon of 6 months. In contrast, features from thematic groups such as *AI Support in SDLC* and *Learning Support*, while also appearing in both successes and needs categories, were expressed in ways that were too generic to be translated into concrete and clearly defined feature candidates. These were therefore excluded from the prediction market to ensure that only features with a precise and operational definition were included.

Based on the final list of features, we formulated 18 binary questions for the prediction market. Each question asked whether a given AI-based feature would be released within the next six months, either in JetBrains' IDEs or any other AI-enabled development tool. These questions were published on JetBrains' internal prediction market platform, where employees place virtual bets to estimate the likelihood of future events. All 18 questions were grouped into a single tournament, where cash prizes (\$1500, \$1000, \$500) were to be awarded to the top three predictors on the leader board to incentivize participation.

**4.1.2 Participants.** All participants of the prediction market were employed by JetBrains. A total of 102 participants [18] across 84 teams placed bets on at least one question during the tournament. The median tenure was 4 years and 10 months. The most common roles included Software Developer (34 participants), Team Lead (20), Product Manager (6), Analyst (6), Quality Assurance (QA) Engineer (5), and Machine Learning (ML) Engineer (4).

**4.1.3 Data collection.** The tournament was launched in December 2024. Each participant received 500 virtual coins to place bets on the outcomes of 18 binary questions. For each question, participants could allocate any number of coins to the YES or NO outcome. Before trading, all participants were informed about tournament rules, trading procedures, and strategies for making informed bets. During trading, the current market probability for each question was visible. Bidding closed on June 1, 2025, and all questions were resolved in July 2025. Participants who made correct predictions earned coins, and the top performers were awarded monetary prizes.

The final resolutions were made by a panel of four expert judges, each with over seven years of professional experience in software engineering, AI/ML research, and product development at the company. Judges did not participate in the initial preparation of the features for prediction. Before resolution, the organizing team compiled a summary report [18] on the market status of each feature based on analysis of the feature set of JetBrains' IDEs, and 20 other tools that were considered, e.g., Windsurf, Cursor, GitHub Copilot. This document served as a reference for the judges, who reached each final decision through several rounds of structured discussion.

**4.1.4 Data analysis.** We analyzed the prediction market results along three dimensions: participant engagement, accuracy of forecasts, and final probability estimates. Engagement was measured using two metrics: trading volume and market liquidity. Trading *Volume* captures the total number of virtual coins bet across both outcomes of a question, reflecting overall activity. Market *Liquidity* reflects the breadth of participation and was calculated by assigning a base value of 50 to each question, with 20 additional points added for every unique participant who placed a bet. Accuracy

was assessed by comparing the community's majority prediction with the actual outcome, as defined by the resolution criteria. The final probability estimate for each question was calculated as the proportion of coins placed on the YES outcome relative to the total number of coins bet on that question.

## 4.2 FINDINGS

To assess whether internal practitioners correctly anticipate which AI features inspired by user needs are likely to be implemented, we analyzed prediction market outcomes along three dimensions: forecast accuracy, final probability estimates, and participant engagement. The results are summarized in Table 2.

**Forecasting performance.** Ultimately, experts concluded that 4 of the 9 features have appeared on the market in any coding tool, and 2 of the 9 appeared specifically in JetBrains IDEs by the end of the prediction window (July 2025). The market produced accurate forecasts for 16 of the 18 questions (88%), showing that at least for short-term predictions the professional community demonstrates considerable precision.

Two of the incorrect forecasts were false negatives: participants assigned low probabilities to features that were actually released: *Project-wide Context Awareness* in JetBrains IDEs, and *Business-context Awareness* in any of the tools available on the market. This pattern suggests a potential **a blind spot**: while these needs were voiced by users, they were underestimated by internal practitioners, potentially due to organizational constraints or differing views on feasibility. What is interesting, no feature was overestimated, meaning there were no cases where the market predicted that a feature would be implemented, but it actually was not. This conservative bias suggests that **market signals can be trusted to filter out overly speculative ideas** (e.g., *Brain-computer Interface Coding*), but may understate the pace of external innovation.

**Alignment with user needs.** Among the 9 AI features derived from user needs in Phase 1, three were correctly forecast as likely to be implemented and were in fact released: *Multi-file Generation*, *Project-wide Context Awareness*, and *Sketch2Code*. These features respond to the needs around *AI in Implementation* and *Context-aware Technology*. Practitioners identified them as feasible and valuable, and their accurate prediction indicates **strong consensus and implementation readiness**.

In contrast, features that required deeper system integration or organizational data, such as *Proactive Documentation*, *Debugging*, *Optimization*, and *AI Real-Time Refactoring*, received low probability scores (all below 33%) and were not released. This suggests that despite user interest, **the technical or infrastructural barriers prevent realization of such features**. Here, the low probabilities assigned to these features do not imply irrelevance. On the contrary, their presence in the prioritized by experts list of interesting features from Phase 1 and the fact that practitioners did vote for them, albeit cautiously, indicates that these needs remain relevant but unaddressed by current tools. This creates a clear opportunity for future development.

Our results suggest that prediction markets can help professional communities identify feasible features and areas for innovation, as

well as guide product planning and standardization efforts around AI-assisted development.

Future work should investigate intervention strategies to close the recognition-implementation gap. Potential approaches include: (1) *capability reassessment workshops* where development teams re-evaluate the feasibility of previously dismissed features using structured technical analysis; (2) *incremental implementation pathways* that break complex features into manageable development milestones; and (3) *user-developer collaboration sessions* that allow practitioners to better understand the business value and user impact of challenging-to-implement features. Additionally, longitudinal studies could track whether features initially deemed infeasible by practitioners are successfully implemented by competitor tools or through alternative technical approaches, providing evidence for or against the accuracy of initial feasibility judgments.

**Answer to RQ3.** Practitioners correctly predicted the outcome for 16 of the 18 features, showing that collective judgment aligns with both user needs and actual realization. Features related to implementation and context awareness were accurately anticipated and released.

## 5 DISCUSSION

This study bridges qualitative insights from developers with internal feasibility assessments, offering a dual-perspective view on in-IDE AI. By integrating user voices and practitioner judgment, we extend prior literature on human-AI experience in software engineering.

While past studies have mapped general expectations for AI tools for coding, our work focuses on in-IDE use, capturing and systematizing nuanced needs shaped by the real-time and context-rich nature of coding environments. Our findings support existing calls for context-aware [27], non-intrusive [24, 31], transparent [34], ethical [21, 23], and customizable [20] AI systems. However, our study also reveals new priorities: integration into broader business workflows, support for skill building, and proactive behavior.

The inclusion of Churners and Non-Users helps surface barriers often missed in adoption-centered research [28]. In particular, there is a need for trust-building mechanisms and accessible onboarding.

Importantly, the prediction market reveals a pattern of conservatism: features that are straightforward to implement are correctly anticipated and realized. In contrast, proactive features and those tied to code maintenance are underestimated despite user interest. This points to a gap between user demand and what teams believe is feasible in the near term.

Together, our two-phase approach illustrates not only what developers want and expect from AI in IDEs, but also which of these expectations are likely to shape product development. Based on our research findings, we propose several key recommendations to guide the future development of in-IDE AI assistants to improve usability, enhance trust, and better integrate AI into developers' workflows. These recommendations could be generalized to most IDEs and editors since they represent the requests and needs of the users rather than specific features of a specific system.

**Technology Improvements. Create Proactive AI.** All three participant groups expressed a need for Proactive AI, which was also rated as relevant in the prediction market. AI assistants should

take a more proactive role in development, moving beyond reactive responses to explicit prompts and recommendations. Key actions include: (1) *Autonomous suggestions*: continuously analyze the code and suggest improvements, such as identifying potential bugs, performance issues, or optimization opportunities; (2) *Automated actions*: instead of suggesting changes, the assistant should be able to perform them proactively, e.g., automatically creating new files with boilerplate code or project structure when a user starts a new project.

**Build User Trust.** Trust is critical for AI adoption in development environments. Developers should prioritize features that build transparency and reliability: (1) *On-premise AI*: offer on-premise AI to enhance security and privacy for organizations with sensitive codebases; (2) *Transparent outputs*: providing explanations and sources (e.g., Stack Overflow links) for the assistant's suggestions helps users understand and trust the output; (3) *Up-to-date data*: ensure that the AI assistant leverages current and relevant libraries, frameworks, and coding standards to provide reliable recommendations.

**Ensure Accessibility.** AI technologies should be usable by developers from various regions and backgrounds: (1) *Flexible pricing and access*: enable users to choose the features they want to pay for and provide options to access AI from various platforms; (2) *Language support*: incorporate multi-language support and localized interfaces to ensure that AI tools are usable in different countries; (3) *Inclusive access*: consider regional differences in infrastructure by offering lightweight or offline versions for areas with limited internet connectivity.

**Ensure Privacy.** Privacy and security were unresolved concerns for many participants. To support adoption under organizational and national constraints, AI systems should: (1) *Support on-premise and local AI*: allow use without sending code to external servers, enabling companies to keep data within national or internal infrastructure; (2) *Enable auditability*: ensure that AI-generated code is clearly marked to support inspection and maintain trust.

**Simplifying Programming Tasks. Look Beyond Code Writing.** To support the needs of all users' groups, AI should not only assist in code writing but also in code-base navigation, evaluation, and in improving code quality: (1) *Codebase navigation*: support developers in understanding and exploring the codebase through contextual summaries and links between related components; (2) *Project analysis*: AI assistants should be able to provide evaluations of the codebase and suggest optimizations for performance, security, or maintainability.

**Integrate AI Systems.** Participants in all three groups expressed an unmet need for in-IDE AI systems that are tightly connected with internal systems and other engineering tools. Smooth integration with external systems will boost the AI's utility throughout the SDLC: (1) *CI/CD integration*: enable the AI to support continuous integration and continuous deployment pipelines, automating build and deployment tasks; (2) *Task tracking integration*: connect with task management tools, allowing the AI to monitor issue statuses, update progress, and link IDE activities with external sources; (3) *Commit management*: automate and manage commits, generating commit messages and simplifying the version control process within IDE; (4) *Code review support*: integrate with code review

platforms to assist in reviewing code, and suggesting improvements without leaving the IDE.

**Technology Interaction. Streamline In-IDE Interaction.** Participants emphasized the importance of minimizing context switching and ensuring smooth interaction between AI and existing IDE features. To support this: (1) *Enable embedded interfaces*: keep AI interaction within the IDE through integrated chat and in-line code explanations to reduce disruption; (2) *Support multimodal input*: allow voice commands or device-based control (e.g., wearables) to increase accessibility and flexibility.

**Technology Customization. Personalize AI.** Developers, especially Churners, seek personalized AI experiences that adapt to their unique needs and workflows. This can be achieved by: (1) *Custom setups*: allow users to configure the AI assistant, e.g., selecting models, defining interaction methods, or choosing the technology stack; (2) *Learning from interactions*: let the assistant learn and adjust from user feedback (e.g., declining specific completions).

**Simplifying Skill Building. Reduce the Learning Curve.** Onboarding and education are key for users to fully utilize AI assistants. All three groups of participants expressed challenges and needs in this regard: (1) *User training*: offer tutorials and examples to help users interact effectively with the AI; (2) *Prompting assistance*: implement prompt suggestions and provide users with feedback on their prompts to improve their AI interaction skills.

## 6 THREATS TO VALIDITY

**Internal validity.** The sample size was determined to reach data saturation; therefore, conclusions are drawn based on the patterns observed within this sample, and further research with a larger or different sample may reveal additional insights or variations of existing topics.

The interview guide was developed and piloted to refine the questions and structure. However, despite these efforts, there may still be inherent biases in the questions or the way they were presented that could influence participants' responses. The iterative process of developing and refining the interview guide aimed to mitigate this risk.

The prediction market participants were JetBrains employees whose predictions may reflect insider knowledge. While this limits external generalizability, it strengthens ecological validity as these practitioners directly shape product decisions.

**External validity.** The sample was recruited from a pool of developers using JetBrains products. By focusing on JetBrains' users but allowing for different AI tools, we aimed to balance consistency with a broad range of experiences, ensuring that our findings are applicable beyond any single tool. However, it may not fully capture the diversity of experiences and perspectives of developers using different IDEs, potentially limiting the findings' generalizability. Furthermore, our interview sample consists predominantly of senior practitioners, potentially underrepresenting juniors' needs. Future research should recruit across experience levels and IDE ecosystems to better capture variation in AI tool expectations.

Additionally, the interviews and prediction market were conducted during May-June 2024 and December 2024 - June 2025, respectively, a period characterized by rapid evolution in AI-assisted development tools and changing developer attitudes toward AI

integration. The findings reflect developer experiences and expectations during this specific temporal context. Given the fast-paced nature of AI tool development and adoption, developers needs, trust levels, and feature priorities may shift as the technology matures and becomes more ubiquitous. Future research should consider how temporal factors influence the stability of the identified personas and feature preferences over time.

**Construct validity.** Our thematic analysis followed established qualitative research protocols with iterative coding and consensus-building between two researchers to ensure consistency and reliability. Our goal was to surface previously unexplored differences across adoption stages to inform product strategy, not to establish statistical significance. While our approach cannot statistically validate whether observed thematic patterns represent true group distinctions versus differential reporting rates, this limitation is inherent to exploratory qualitative research with modest sample sizes. The practical value of these insights is demonstrated by their successful translation into actionable features in our prediction market phase, where internal experts and practitioners recognized the feasibility of the features addressing the user needs we identified. Future work could complement our exploratory findings with larger-scale quantitative validation, but the qualitative patterns we uncovered provide essential foundational insights for understanding developer personas in AI tool adoption.

In the second phase, conceptual features derived from interviews were used to formulate prediction market questions. However, these features did not always align well with concrete implementation outcomes. While a structured resolution guide and expert panel were used to mitigate this issue, some interpretive variability may remain, affecting the alignment between predicted outcomes and actual market developments.

## 7 CONCLUSION

This study examined what different groups of developers want and expect from AI in IDEs, focusing on unmet needs and anticipated features. Through interviews with 35 Adopters, Churners, and Non-Users, we surfaced a structured landscape of user expectations across five areas: *Technology Improvements*, *Programming Tasks*, *Interaction*, *Customization*, and *Skill Building*. These highlight the demand for more advanced, reliable, and trustworthy AI experiences within development environments. To assess implementation feasibility, we conducted an internal prediction market with 102 practitioners betting on 9 AI features covering unmet needs. The results show alignment between user needs and practitioner judgment for core implementation and context-aware features. However, proactive behaviors and long-term code maintenance remain both underdelivered and underestimated, pointing to a systemic blind spot. Our findings bridge user expectations and product planning by showing where practitioner foresight aligns, or fails to align, with real-world developer needs. Tool builders should prioritize context-awareness and implementation support today, while investing in proactive and maintenance features for the longer term. Future work should explore how to operationalize unmet users' needs that are recognized but deprioritized by tool builders.

## Acknowledgments

This work was supported by JetBrains as part of the AI for Software Engineering (AI4SE) collaboration with Delft University of Technology.

## References

- [1] 2024. GitHub Copilot. <https://github.com/features/copilot>. Accessed: June 2024.
- [2] 2024. JetBrains AI. <https://www.jetbrains.com/ai/>. Accessed: June 2024.
- [3] 2024. Stack Overflow's 2023 Developer Survey. <https://survey.stackoverflow.co/2023/technology>. Accessed: September 2024.
- [4] Saleema Amershi, Daniel S. Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi T. Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13. doi:10.1145/3290605.3300233
- [5] Kenneth J Arrow, Robert Forsythe, Michael Gorham, Robert Hahn, Robin Hanson, John O Ledyard, Saul Levmore, Robert Litan, Paul Milgrom, Forrest D Nelson, et al. 2008. The promise of prediction markets. 877–878 pages.
- [6] Jay R Avella. 2016. Delphi panels: Research design, procedures, advantages, and challenges. *International Journal of Doctoral Studies* 11 (2016), 305.
- [7] Joyce E. Berg and Thomas A. Rietz. 2014. Market Design, Manipulation, and Accuracy in Political Prediction Markets: Lessons from the Iowa Electronic Markets. *PS: Political Science & Politics* 47, 2 (April 2014), 293–296. doi:10.1017/S1049096514000070
- [8] John Bryant and Kirk Jensen. 1994. Forecasting Inkjet Printers at Hewlett-Packard Company. *The Journal of Business Forecasting* 13, 2 (July 1994), 27.
- [9] Bo Cowgill and Eric Zitzewitz. 2015. Corporate prediction markets: Evidence from google, ford, and firm x. *The Review of Economic Studies* 82, 4 (2015), 1309–1341.
- [10] Cursor AI. 2024. Cursor AI: The AI-First Code Editor. <https://www.cursor.com/>. Accessed: May 2025.
- [11] Aral de Moor, Arie van Deursen, and Maliheh Izadi. 2024. A transformer-based approach for smart invocation of automatic code completion. In *Proceedings of the 1st ACM International Conference on AI-Powered Software*. 28–37.
- [12] Anna Dreber, Thomas Pfeiffer, Johan Almenberg, Siri Isaksson, Brad Wilson, Yiling Chen, Brian A. Nosek, and Magnus Johannesson. 2015. Using Prediction Markets to Estimate the Reproducibility of Scientific Research. *Proceedings of the National Academy of Sciences* 112, 50 (Dec. 2015), 15343–15347. doi:10.1073/pnas.1516179112
- [13] Satu Elo and Helvi Kyngäs. 2008. The qualitative content analysis process. *Journal of advanced nursing* 62, 1 (2008), 107–115.
- [14] Jane Forman and Laura Damschroder. 2007. Qualitative content analysis. In *Empirical methods for bioethics: A primer*. Emerald Group Publishing Limited, 39–62.
- [15] Monique Hennink and Bonnie N. Kaiser. 2022. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science and Medicine* 292 (2022), 114523. doi:10.1016/j.socscimed.2021.114523
- [16] ICC/ESOMAR. 2024. International Code on Market, Opinion and Social Research and Data Analytics. <https://esomar.org/uploads/attachments/ckqtawvj00uukdtrhst5sk9u-icesomar-international-code-english.pdf>. Accessed: October 2024.
- [17] Maliheh Izadi, Jonathan Katzy, Tim Van Dam, Marc Otten, Razvan Mihai Popescu, and Arie Van Deursen. 2024. Language models for code completion: A practical evaluation. In *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering*. 1–13.
- [18] JetBrains Research. 2025. Supplementary Materials. <https://doi.org/10.5281/zenodo.17976045>. Accessed: December 2025.
- [19] Hanna Kallio, Anna-Maija Pietilä, Martin Johnson, and Mari Kangasniemi. 2016. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of advanced nursing* 72, 12 (2016), 2954–2965.
- [20] Jenny T. Liang, Chenyang Yang, and Brad A. Myers. 2024. A Large-Scale Survey on the Usability of AI Programming Assistants: Successes and Challenges. In *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering* (Lisbon, Portugal) (ICSE '24). Association for Computing Machinery, New York, NY, USA, Article 52, 13 pages. doi:10.1145/3597503.3608128
- [21] Chao Liu, Xindong Zhang, Hongyu Zhang, Zhiyuan Wan, Zhan Huang, and Meng Yan. 2024. An Empirical Study of Code Search in Intelligent Coding Assistant: Perceptions, Expectations, and Directions. In *Companion Proceedings of the 32nd ACM International Conference on the Foundations of Software Engineering* (Porto de Galinhas, Brazil) (FSE 2024). Association for Computing Machinery, New York, NY, USA, 283–293. doi:10.1145/3663529.3663848
- [22] Mark Marron. 2024. A New Generation of Intelligent Development Environments. In *Proceedings of the 1st ACM/IEEE Workshop on Integrated Development Environments*. 43–46.
- [23] Andrew M McNutt, Chenglong Wang, Robert A Deline, and Steven M. Drucker. 2023. On the Design of AI-powered Code Assistants for Notebooks. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 434, 16 pages. doi:10.1145/3544548.3580940
- [24] Hussein Mozannar, Gagan Bansal, Adam Fourney, and Eric Horvitz. 2024. When to Show a Suggestion? Integrating Human Feedback in AI-Assisted Programming. *Proceedings of the AAAI Conference on Artificial Intelligence* 38, 9 (Mar. 2024), 10137–10144. doi:10.1609/aaai.v38i9.28878
- [25] Everett M Rogers, Arvind Singhal, and Margaret M Quinlan. 2014. Diffusion of innovations. In *An integrated approach to communication theory and research*. Routledge, 432–448.
- [26] Anton Semenkina, Yaroslav Sokolov, and Evgeniia Vu. 2024. Context Composing for Full Line Code Completion. In *Proceedings of the 1st ACM/IEEE Workshop on Integrated Development Environments* (Lisbon, Portugal) (IDE '24). Association for Computing Machinery, New York, NY, USA, 15–17. doi:10.1145/3643796.3648446
- [27] Agnia Sergeyuk, Yaroslav Golubev, Timofey Bryksin, and Iftekhar Ahmed. 2024. Using AI-Based Coding Assistants in Practice: State of Affairs, Perceptions, and Ways Forward. arXiv:2406.07765 [cs.SE] <https://arxiv.org/abs/2406.07765>
- [28] Agnia Sergeyuk, Ilya Zakharov, Ekaterina Koshchenko, and Maliheh Izadi. 2025. Human-AI Experience in Integrated Development Environments: A Systematic Literature Review. *arXiv preprint arXiv:2503.06195* (2025).
- [29] Priyan Vaithilingam, Tianyi Zhang, and Elena L Glassman. 2022. Expectation vs. experience: Evaluating the usability of code generation tools powered by large language models. In *Chi conference on human factors in computing systems extended abstracts*. 1–7.
- [30] Tim van Dam, Maliheh Izadi, and Arie van Deursen. 2023. Enriching source code with contextual data for code completion models: An empirical study. In *2023 IEEE/ACM 20th International Conference on Mining Software Repositories (MSR)*. IEEE, 170–182.
- [31] Chaosheng Wang, Junhao Hu, Cuiyun Gao, Yu Jin, Tao Xie, Hailiang Huang, Zhenyu Lei, and Yuetang Deng. 2023. How Practitioners Expect Code Completion?. In *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (San Francisco, CA, USA) (ESEC/FSE 2023). Association for Computing Machinery, New York, NY, USA, 1294–1306. doi:10.1145/3611643.3616280
- [32] Justin Wolfers and Eric Zitzewitz. 2004. Prediction markets. *Journal of economic perspectives* 18, 2 (2004), 107–126.
- [33] Beiqi Zhang, Peng Liang, Xiyu Zhou, Aakash Ahmad, and Muhammad Waseem. 2023. Demystifying practices, challenges and expected features of using GitHub Copilot. *International Journal of Software Engineering and Knowledge Engineering* 33, 11n12 (2023), 1653–1672.
- [34] Xiyu Zhou, Peng Liang, Beiqi Zhang, Zengyang Li, Aakash Ahmad, Mojtaba Shahrin, and Muhammad Waseem. 2024. Exploring the Problems, their Causes and Solutions of AI Pair Programming: A Study on GitHub and Stack Overflow. arXiv:2311.01020 [cs.SE] <https://arxiv.org/abs/2311.01020>