

Table 1. Main findings for RQ1: Use of GenAI for software development

4.1 Uses, benefits, and challenges of GenAI	
4.1.1 GenAI usage in Software Development	
F1	While the vast majority of developers surveyed (89.7%) were using GenAI tools, there were many reasons why developers might not incorporate GenAI into their development workflows, including perceived legal issues, personal preference, information security concerns, and limitations placed by their employers.
F2	Of those respondents who reported working at organizations that disallowed the use of GenAI, many indicated that they believed the restriction was based on fear of proprietary information leakage and the potential legal challenges such usage might bring. Some developers also perceived that their employers were slow to adopt new technology.
4.1.2 GenAI tools used by developers	
F3	The ecosystem of GenAI tools used by developers was varied and extensive, comprising both open-source/open-weight and proprietary solutions. Developers primarily made use of web-based tools and those that could be integrated into their IDEs.
4.1.3 Benefits derived from GenAI usage	
F4	The benefits developers reported from using GenAI tools for coding go beyond increasing productivity and avoiding repetitive, boring tasks. They also included understanding code better and improved code quality and code optimization, which some human developers might struggle to achieve.
4.1.4 Challenges and shortcomings encountered in using GenAI	
F5	Aside from obvious problems such as possible hallucinations, GenAI for coding still suffers from challenges related to lack of suitable contextual information available to the model at inference time, the inability for such models to solve complex, non-trivial tasks such as highly-specific functionality which conforms to a given architecture, and the need to provide good, well-designed prompts.
4.2 Current practices for documentation and compliance	
4.2.1 AI-generated code review and compliance	
F6	Among respondents, GenAI code review is typically done manually or by using other GenAI to check for code correctness and compliance. Some developers are concerned that the process of conducting a review may expose them or their employers to liability.
4.2.2 Documentation of GenAI usage	
F7	Most developers in our study (70.1%) were unaware of any process for documenting GenAI usage within their organization. To the extent that organizations have processes in place for documenting the use of GenAI, they may not be sufficiently educating their employees about those processes.
F8	The most common documentation process among respondents was to note the fact of GenAI usage and/or to document the prompt used to acquire the generation. However, nearly a third of respondents (32.5%) indicated documenting the whole interaction, potentially leading to reproducibility.
4.2.3 Review of Terms of Service	
F9	Even though some provisions in a tool's Terms of Service may be important to developers and their work, developers do not always thoroughly read the Terms of Service for the tools they use, choosing instead to skim the document, assume that their organization has read the document, or rely on information and validation from others in the development community.
4.2.4 Copyright/legal training	
F10	Relatively few respondents (11.9%) reported having undergone any formal training on the legal implications of GenAI. Some struggled to find suitable available resources.

Table 2. Main findings for RQ2: Developer perceptions of copyright issues related to GenAI

5.1 Thoughts on Litigation	
F11	While about half of participants anticipated that litigation would result in limitations on the use of GenAI, a vocal minority suggested that imposing such limitations would be difficult if not impossible, and others urged restraint, given the pace of technological developments.
5.2 Using one's own code in models' training data or as output	
5.2.1 Sentiments regarding the inclusion of one's own code in training data	
F12	Developers' views on having their code included in models' training data was context-dependent. If the code was open source, many developers were indifferent or even pleased to have their work included; this was not true for proprietary code. Developers also expressed concerns about the quality of code incorporated into the training data and whether increased use of GenAI tools would disincentivize developers from contribution to the OSS community.
5.2.2 Conditions around use of one's code in training data	
F13	Some respondents suggested that training data contributors should receive monetary compensation or other benefits, such as free access to the trained model or free credits for the platform. However, forcing AI companies to provide monetary compensation to all contributors of the training data, while seemingly fair, would place an additional hurdle potentially blocking smaller AI companies from entering the market.
F14	Respondents generally wanted licensing terms to be complied with by GenAI models, especially when producing code identical to or very similar to work in the training set.
5.2.3 Fair use and replicated code	
F15	Some developers believed that the use of code as training data should be considered fair use by default and that the onus should be on developers to protect code that they do not wish to be included in training data.
5.3 AI-generated output that resembles existing code	
F16	A group of developers did not see how GenAI technology was doing anything significantly different from traditional means of creating code, which reuses existing code created by others. The only distinction is the scale and ease at which it can be done.
F17	Some developers' views on the copying of code track U.S. copyright law: basic functions and building blocks are not protectable and should be free for all to use.
5.4 On the ownership of AI-generated code	
5.4.1 No one-works are in the public domain	
F18	The largest group of developers (242, 43.68%) thought that the output of GenAI belongs in the public domain. Some believe this because it is consistent with the collaborative nature of open-source software development, while others believe that, because AI models are not human, the output of those models cannot be subject to copyright law.
5.4.2 The prompting developer or their employer	
F19	The second largest group of developers (201, 36.3%) believed that a user of a GenAI tool should own the copyright to code generated with that tool because it is the user's contributions—creating the prompt and, typically, editing and further refining the result—that create the code. In other words, GenAI was seen as nothing more than a new tool for creating content.
5.4.3 The creators whose works make up the training data	
F20	Some developers believed that fairness requires giving ownership of generated code to those developers who created the training data. However, others invoked the collaborative nature of development as a reason to deny copyright claims by those whose code was used in the training data and highlighted the practical problems with mapping training data to generated code in order to determine ownership.
5.4.4 The creators of the model	
F21	The model vendors who developed GenAI models were the one group that the fewest participants (49) believed had an ownership claim to generated outputs.
5.4.5 Other ownership claims	
F22	Some respondents believed that ownership depended on various factors, including the relevant jurisdiction, the amount of code generated, and how much the prompting developer contributed to the result. Others believed that multiple parties share claims to ownership and that a royalty-style system should be put into place.
5.5 On the ownership and copying of prompts	
F23	Most respondents (77.4%) indicated that they would be indifferent or even pleased if their prompts were copied and used by others, even without their permission.
5.6 Perceived risk of potential copyright infringement	
F24	Individuals and smaller organizations may perceive low practical legal risk from using AI generated code due to the low likelihood that any code which potentially infringes on another's copyright or violates another's license would be discovered.

Table 3. Main findings for RQ3: Other legal concerns anticipated by developers

6.5 Liability concerns	
F25	Developers raised questions about who would be held accountable for bugs and other defects in GenAI outputs, especially for mission-critical systems such as medical, aerospace, and security. Developers were also concerned about how would be held responsible for actions that AI agents take autonomously.
6.1 Generation of malicious content	
F26	Respondents identified the purposeful generation of malicious content as a societal problem and legal concern going forward.
6.2 Data privacy and information leakage	
F27	Developers and other stakeholders did not intend for or anticipate that their data would be collected <i>en masse</i> , so developers are concerned that private and sensitive information can easily makes its way into training sets.
F28	Generative AI, particularly LLMs, can not only leak information that was contained in the original training data but can also be used by malicious actors to exploit systems.
6.3 Testing and test generation	
F29	Developers had concerns related to AI-generated tests, including their completeness and reliability.