

# LLM Recommendations for Programming and Development

Based on the capability tags and descriptions provided in your model list, here is a breakdown of which Large Language Models (LLMs) are best suited for your specific programming and general tasks.

The key capabilities we looked for were **Tools**, **Thinking/Reasoning**, **Code**, **Agentic workflows**, and **Vision**.

## 1. Refactoring, New Code Creation, and Debugging

For tasks involving specific programming languages (py, html, xcode) and complex frameworks like Apple Vision Pro, you need models with strong reasoning, tool-use, and dedicated coding proficiency.

Model	Key Tags/Description	Rationale
qwen3-coder	tools, cloud, long context	Dedicated coding model with tool-use capability, ideal for complex tasks and handling large codebases (long context).
glm-4.6	Advanced agentic, reasoning and coding capabilities	Explicitly built for coding and agentic (automated task) capabilities, making it excellent for debugging and refactoring.
minimax-m2	built for coding and agentic workflows	High-efficiency model focused on automated coding workflows, very useful for large-scale refactoring and generating new blocks of code.
kimi-k2	real-world coding agent tasks	Mixture-of-Experts (MoE) model showing strong performance on agentic and coding tasks, beneficial for complex development.
deepseek-r1	tools thinking	Strong focus on reasoning and tool usage, which is critical for complex debugging and structural refactoring.
firefunction-v2	function calling model	Excellent for scenarios where the code needs to interact with

		external systems or APIs, such as integrating code functionality.
--	--	---

## 2. General Questions and Conversational Use

These models are optimized for natural language conversation, summarization, and answering broad inquiries, making them suitable for research, documentation, or quick Q&A.

Model	Key Tags/Description	Rationale
<b>alfred</b>	A robust conversational model	Explicitly designed for general chat and instruct use cases, making it a reliable choice for everyday questions.
<b>notus</b>	A 7B chat model	A focused chat model fine-tuned for high-quality conversational interaction.
<b>qwen3</b>	tools thinking	A versatile model family that can handle both general Q&A (due to strong reasoning) and more advanced tasks.
<b>goliath</b>	combining two fine-tuned Llama 2 70B models	The large parameter size suggests high capability for general knowledge and complex reasoning typically required for conversational models.

## 3. Manipulating and Understanding Images

For tasks involving images, such as analyzing existing images, creating visual content, or interpreting image data for Apple Vision Pro applications, you need a model with **vision** capabilities.

Model	Key Tags/Description	Rationale
<b>qwen3-vl</b>	The most powerful vision-language model	This model is the top choice, as it's specifically described as a powerful Vision-Language (VL) model for understanding and potentially manipulating image content.
<b>gemma3</b>	vision	Explicitly labeled with the vision tag, indicating it has multimodal capabilities

		necessary for image understanding tasks.
--	--	--

This analysis should give you a good starting point for matching your tasks to the right LLM!