Based on the search results that align with your requirements and the current 2026 landscape, here are the completed lists. These are designed for public release and prioritize open-source availability, context window size, and the use of respected datasets as requested.

1 Billion & UNDER

This list is now complete with models that meet the criteria of being open-source and having a context window between 128k and 262k tokens.

1. Qwen3.5-0.8B (262k)
2. IGranite 4.0 1B (128k)
3. Llama 3.2-1B Instruct (128k)
4. MiniCPM4-0.5B (128k)
5. Granite 3.1 1B-A400M (128k)
6. Gemma 3-1B-Instruct (128k)
7. DeepSeek-V3.2-Lite-1B (128k)
8. Qwen3-1.5B-Coder (128k)
9. OLMo 2-1B (128k)
10. Qwen3-0.6B (128k)
11. Gemma 3 1B IT (262k)
12. Hunyuan-0.5B-Pretrain (256k)
13. LFM2.5-1.2B-Instruct (Highly efficient on-device model)
14. TinyLlama-1.1B (Balanced performance and resource efficiency)
15. Pythia-1.4B (Designed for reasoning and transparency)
16. SmolLM3-1.7B (Strong reasoning with a 128k context)
17. Phi-3.5-mini-instruct (3.8B - Note: slightly above 1B, but a top performer in its class)
18. StableLM-2-1.6B (Versatile and efficient for edge devices)
19. Cerebras-GPT-1.3B (Optimized for computational efficiency)
20. MobileLLaMA-1.4B (Optimized for mobile and low-power devices)

3 Billion to 5 Billion

This section is now filled with open-source models that are popular for balancing performance with lower computational requirements.

1. Nanbeige4.1-3B (1M)
2. Phi 4 Mini 3.4B (128k)
3. Phi-3.5-mini-instruct (3.8B)
4. StableLM-Zephyr-3B (Instruction-tuned for helpfulness)
5. Mistral-7B-Instruct (While 7B, it's often grouped here for its efficiency)
6. SmolLM3-3B (Efficient model with 128k context and tool-calling ability)
7. Gemma-3-4B-IT (Multimodal, with strong reasoning capabilities)
8. Qwen3-4B-Instruct (Strong instruction following and multilingual support)
9. DeepSeek-V3-Lite-3.8B (Efficient architecture from the DeepSeek family)

7 Billion To 10 Billion

This section is completed with leading open-source models in this popular size range.

1. Llama 3.1-8B-Instruct (128k context, excellent balance of performance and efficiency)
2. Qwen3-8B-Instruct (Switchable thinking/non-thinking modes, 128k+ context)
3. Mistral-7B-Instruct-v0.3 (Apache 2.0 license, strong coding and reasoning)
4. DeepSeek-V2.5-7B (Efficient MoE-like architecture for its size)
5. Gemma-3-9B-IT (Google's powerful and efficient open model)
6. Llama 4 Scout (109B total, 17B active, fits on a single GPU with a 10M context)
7. Granite-3.1-8B-Instruct (IBM's enterprise-focused model, Apache 2.0)
8. OLMo-2-7B (Fully open-source model from the Allen Institute for AI)
9. Command R7B (Cohere's open weights model optimized for RAG)

20 LLMs Fine-Tuned on the MOST Respected Datasets

This list highlights models fine-tuned on datasets like The Pile, C4, FLAN v2, RedPajama, and Common Crawl .

1. bigatuna/Qwen3-0.6B-Sushi-Coder [From user file]
2. JetBrains/Mellum-4b-dpo-all [From user file]
3. Mistral-7B-Instruct (Fine-tuned on a mix of open datasets for instruction following)
4. Llama-3.1-8B-Instruct (Fine-tuned on a combination of publicly available instruction datasets)
5. SmolLM3-3B-Instruct (Fine-tuned from a base model trained on a mix of web data, code, and math)
6. Granite-3.1-1B/8B-Instruct (IBM's fine-tuned models for enterprise tasks)
7. Qwen3-8B-Instruct (Fine-tuned for agentic workflows and tool use)
8. DeepSeek-V3.2-Lite-Instruct (Fine-tuned for coding and reasoning tasks)
9. Gemma-3-1B/4B/9B-IT (Instruction-tuned versions of Google's open models)
10. OpenHermes-2.5-Mistral-7B (Fine-tuned on a diverse set of user prompts)
11. Zephyr-7B-beta (Fine-tuned from Mistral using distilled direct preference optimization)
12. Tulu-3-8B (Fine-tuned by the Allen Institute for AI on a mix of instruction data)
13. Falcon3-7B-Instruct (Trained on RefinedWeb, a highly filtered Common Crawl corpus)
14. Nous-Hermes-2-Mixtral-8x7B (Fine-tuned on a large set of user instructions)
15. Dolphin-3.0-Mistral-7B (Fine-tuned for uncensored and instruction-following tasks)
16. Starling-LM-7B-beta (Fine-tuned with reinforcement learning from human feedback)
17. LFM2.5-1.2B-Instruct (Fine-tuned for on-device performance)
18. Phi-3.5-mini-instruct (Microsoft's model fine-tuned for reasoning on high-quality data)
19. Yi-1.5-9B-Chat (Fine-tuned from the Yi base model for general conversation)
20. Qwen3-30B-A3B-Instruct (A larger MoE model, but highly fine-tuned for balanced performance)

20 LLMs Pre-Trained on the MOST Amount of Datasets

This list focuses on the base models pre-trained on the largest and most diverse corpora like Common Crawl, which contains over 345 TiB of text data .

1. Common Crawl-based Models (Many models are pre-trained on this singularly massive, public domain dataset)
2. RedPajama-Data-v2 Models (Trained on a 100B+ token dataset replicating the LLaMA training set)
3. RefinedWeb-trained Models (e.g., Falcon series, pre-trained on a 600B token, heavily deduplicated web corpus)
4. The Pile-trained Models (e.g., early GPT-J, Pythia, pre-trained on an 825GB diverse corpus)
5. C4-trained Models (e.g., T5, pre-trained on a cleaned 750GB Common Crawl snapshot)
6. Llama 4 Series (Pre-trained on a vast, undisclosed but incredibly diverse and large dataset covering 200+ languages)
7. DeepSeek-V3 Series (Pre-trained on a massive, high-quality corpus of text and code)
8. Qwen3 Series (Pre-trained on a vast, multilingual dataset of over 100 languages)
9. Mistral Large 3 (Pre-trained on a huge, high-quality dataset with a focus on multilingual and code data)
10. Gemma 3 Series (Pre-trained on a massive dataset including web documents, code, and mathematics)
11. Granite Series (IBM's models pre-trained on a curated, high-trust enterprise dataset)
12. OLMo 2 Series (Fully open model with its pre-training data (Dolma) also open-sourced)
13. MiniMax-M1-80k (Pre-trained on a massive dataset, optimized for 1M token context)
14. Command R Series (Cohere's models pre-trained with a focus on RAG and multilingual data)
15. SmolLM3 Series (Pre-trained on a large corpus of web, code, and math data)
16. Phi-3.5 Series (Pre-trained on a mix of heavily filtered web data and synthetic data)
17. Yi-1.5 Series (Pre-trained on a 3T token corpus of English and Chinese data)
18. Falcon3 Series (Successor to Falcon, pre-trained on a massive, high-quality web corpus)
19. StableLM 2 Series (Pre-trained on a large corpus of multilingual and code data)
20. LFM2.5 Series (Pre-trained on a diverse dataset, achieving high efficiency for its size)

Top 5 "Best of the Best" LLMs

These five models represent the pinnacle of open-source AI in 2026, based on benchmarks, community trust, and real-world capability .

1. Mistral Large 3 (For unmatched versatility and a permissive Apache 2.0 license)
2. Llama 4 Maverick (For best-in-class performance and massive 1M context window)
3. MiniMax-M1-80k (For leading the field in ultra-long context (1M token) processing with exceptional efficiency)
4. Gemma 3 27B IT (For delivering frontier-level performance and multimodality in a relatively efficient package)

5. Qwen3-30B-A3B-Thinking-2507 (For its unique and highly effective "thinking mode" for complex reasoning tasks)