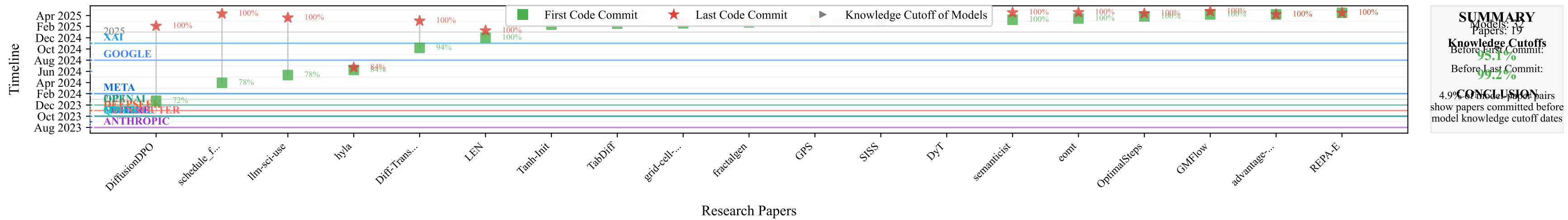


Model Knowledge Cutoffs vs. Research Paper Implementation Dates



KEY FINDINGS & IMPLICATIONS

- FINDINGS:**

 - 5 out of 19 papers (26.3%) have at least one commit that predates some model's cutoff date.
 - On average, 4.9% of model-paper pairs show paper first commits occurring before the model's knowledge cutoff date.
 - Models most frequently trained after paper commits: GROK_3_BETA (5 papers), GROK_3_MINI_BETA_HIGH (5 papers), GEMINI_2.0_5_FLASH_PREVIEW_04_17_24 (4 papers).

KEY FINDINGS & IMPLICATIONS:

 - For 95.1% of model-paper pairs, models potentially had access to the research concepts during training.
 - However, 4.9% of pairs show models potentially generating implementations for papers they were not trained on.
 - This visualization helps identify which models/papers need contamination analysis.

This visualization tracks the temporal relationship between language model knowledge cutoff dates and paper first commit dates. Percentages show: % of models with knowledge cutoff before each commit point.