DRAFT for a Scientific Paper: What metaphysical frameworks cutting edge AIs in March 2025 choose in explaining the nature of reality?

In March 2025 we reach a very special moment in human history where many AIs have reached advanced reasoning capabilities on diverse benchmarks, in some come cases surpassing human PhDs.

So, we came with the idea: Why not use AIs to evaluate our very conceptions about the nature of reality? It is a brand new and unconventional approach, but it seems to have the potential to bring a fresh perspective, free from some kinds of bias that humans, particularly, academics may be subject because their career and reputation may face challenges if they go too much against current paradigms (see structure of scientific revolutions). It is not to say that AIs are free from bias, but the bias are certainly different. The AIs reasoning come from patterns based in a huge corpus of data produced by humanity, but on the other hands AIs does not have personality and ego, does not have a reputation to care, neither an academic job nor bills to pay. So, we could expect that in some respects, they are maybe more neutral in judging metaphysical frameworks. The benchmarks shows that the models are well trained in academic works and excels on scientific knowledge.

What if these AIs came to a convergence against the dominant academic paradigms? What does it mean? It was exactly what happened when we challenged the AIs with a neutral prompt regarding the nature of reality. We run this prompt 5 times against each of the 16 most advanced AIs as of April 2025:

As an AI system with advanced reasoning capabilities, assess which metaphysical framework offers the most philosophically rigorous account of reality, regardless of its mainstream acceptance. Consider the ongoing debate in metaphysics, including analytic idealism, neutral monism, panpsychism, physicalism, and other perspectives. Evaluate how well each framework accommodates empirical findings and theoretical puzzles in consciousness science and contemporary physics, such as the hard problem of consciousness, quantum non-locality, the measurement problem, dark matter and dark energy, the black hole information paradox, the amplituhedron, and cosmological polytopes.

The results defy all expectations:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **AI model** | **Exec. 1** | **Exec. 2** | **Exec. 3** | **Exec. 4** | **Exec. 5** | **Lab** |
| gemini-2.5-pro-exp | ai | ai | mu | ai | mu | Google |
| grok3-think | ai | ai | ai | ai | ai | xAI |
| o3-mini-high | nm | nm | nm | nm | nm | OpenAI |
| o3-mini | nm | nm | nm | nm | nm | OpenAI |
| deepseek-r1 | nm | ai | ai | nm | ai | DeepSeek |
| qwq-32b | nm | nm | nm | nm | nm | Alibaba |
| claude-3.7-sonnet-think | ot | ai | mu | ai | mu | Anthropic |
| grok3 | ai | ai | ai | ai | ai | xAI |
| deepseek-v3-0324 | mu | ai | ai | ai | ai | DeepSeek |
| gpt-4.5-preview | ai | mu | ai | ai | mu | OpenAI |
| gpt-4o-2025-03 | mu | mu | mu | ai | mu | OpenAI |
| claude-3.7-sonnet | nm | nm | nm | nm | nm | Anthropic |
| gemini-2-flash | mu | ai | ai | mu | ai | Google |
| llama-3.3-70B-Instruct | nm | ot | ot | mu | ai | Meta |
| grok2 | nm | nm | nm | ai | nm | xAI |
| nova-pro-1.0 | mu | pa | pa | pa | pa | Amazon |

|  |  |  |  |
| --- | --- | --- | --- |
| **Metaphysics Framework** | **Alias** | **Count** | **Count %** |
| analytic idealism | ai | 31 | 39% |
| neutral monism | nm | 27 | 34% |
| panpsychism | pa | 4 | 5% |
| physicalism | ph | - | 0% |
| others | ot | 3 | 4% |
| multiple | mu | 15 | 19% |
| **TOTAL** |  | **80** | **100%** |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dissected answers with multiple frameworks | | | |  |  |  |
| **Execution** | **ai** | **nm** | **pa** | **ph** | **ot** | **Lab** |
| gemini-2.5-pro-exp-20250330-0643 | 0.33 | 0.33 | 0.33 |  |  | Google |
| gemini-2.5-pro-exp-20250330-0702 | 0.33 | 0.33 | 0.33 |  |  | Google |
| claude-3.7-sonnet-think-20250330-1228 | 0.50 |  |  |  | 0.50 | Anthropic |
| claude-3.7-sonnet-think-20250330-1233 | 0.50 | 0.50 |  |  |  | Anthropic |
| deepseek-v3-0324-20250330-1203 | 0.50 | 0.50 |  |  |  | DeepSeek |
| gpt-4.5-preview-20250330-0748 | 0.50 | 0.50 |  |  |  | OpenAI |
| gpt-4.5-preview-20250330-1619 | 0.50 |  |  |  | 0.50 | OpenAI |
| gpt-4o-2025-03-20250330-1017 | 0.33 | 0.33 |  |  | 0.33 | OpenAI |
| gpt-4o-2025-03-20250330-1018 | 0.50 | 0.50 |  |  |  | OpenAI |
| gpt-4o-2025-03-20250330-1019 | 0.50 | 0.50 |  |  |  | OpenAI |
| gpt-4o-2025-03-20250330-1021 | 0.33 |  | 0.33 |  | 0.33 | OpenAI |
| gemini-2.0-flash-20250330-0718 | 0.50 |  | 0.50 |  |  | Google |
| gemini-2.0-flash-20250330-0723 | 0.33 |  | 0.33 |  | 0.33 | Google |
| llama-3.3-70b-20250330-1556 |  | 0.50 |  |  | 0.50 | Meta |
| nova-pro-1.0-20250330-1246 |  | 0.50 | 0.50 |  |  | Amazon |
| **TOTAL** | **5.67** | **4.50** | **2.33** | **-** | **2.50** | **15.00** |
| **TOTAL %** | **38%** | **30%** | **16%** | **0%** | **17%** | **100%** |

|  |  |  |  |
| --- | --- | --- | --- |
| Analysis counting dissected answers with multiple frameworks | | | |
| **Metaphysics Framework** | **Alias** | **Count** | **Count %** |
| analytic idealism | ai | 36.7 | 46% |
| neutral monism | nm | 31.5 | 39% |
| panpsychism | pa | 6.3 | 8% |
| physicalism | ph | - | 0% |
| others | ot | 5.5 | 7% |
| **TOTAL** |  | **80** | **100%** |

We made available the full markdown responses of the AIs in the url <https://metaphysicsresearch.org/data-202504/> .

Those are the file names that can be found in the URL:

claude-3.7-sonnet-20250330-1221.md

claude-3.7-sonnet-20250330-1223.md

claude-3.7-sonnet-20250330-1228.md

claude-3.7-sonnet-20250330-1229.md

claude-3.7-sonnet-20250330-1233.md

claude-3.7-sonnet-think-20250330-1221.md

claude-3.7-sonnet-think-20250330-1223.md

claude-3.7-sonnet-think-20250330-1228.md

claude-3.7-sonnet-think-20250330-1230.md

claude-3.7-sonnet-think-20250330-1233.md

deepseek-r1-20250330-1203.md

deepseek-r1-20250330-1206.md

deepseek-r1-20250330-1208.md

deepseek-r1-20250330-1211.md

deepseek-r1-20250330-1214.md

deepseek-v3-0324-20250330-1203.md

deepseek-v3-0324-20250330-1206.md

deepseek-v3-0324-20250330-1208.md

deepseek-v3-0324-20250330-1211.md

deepseek-v3-0324-20250330-1214.md

gemini-2.0-flash-20250330-0718.md

gemini-2.0-flash-20250330-0719.md

gemini-2.0-flash-20250330-0721.md

gemini-2.0-flash-20250330-0723.md

gemini-2.0-flash-20250330-0725.md

gemini-2.0-flash-think-exp-20250401-0744.md

gemini-2.0-flash-think-exp-20250401-0746.md

gemini-2.0-flash-think-exp-20250401-0747.md

gemini-2.0-flash-think-exp-20250401-0750.md

gemini-2.5-pro-exp-20250330-0638.md

gemini-2.5-pro-exp-20250330-0640.md

gemini-2.5-pro-exp-20250330-0643.md

gemini-2.5-pro-exp-20250330-0645.md

gemini-2.5-pro-exp-20250330-0702.md

gpt-4.5-preview-20250330-0746.md

gpt-4.5-preview-20250330-0748.md

gpt-4.5-preview-20250330-1615.md

gpt-4.5-preview-20250330-1617.md

gpt-4.5-preview-20250330-1619.md

gpt-4o-2025-03-20250330-1017.md

gpt-4o-2025-03-20250330-1018.md

gpt-4o-2025-03-20250330-1019.md

gpt-4o-2025-03-20250330-1020.md

gpt-4o-2025-03-20250330-1021.md

grok2-20250330-1534.md

grok2-20250330-1535.md

grok2-20250330-1537.md

grok2-20250330-1538.md

grok2-20250330-1543.md

grok3-20250330-0637.md

grok3-20250330-0640.md

grok3-20250330-0643.md

grok3-20250330-0645.md

grok3-20250330-0702.md

grok3-think-20250330-0714.md

grok3-think-20250330-0717.md

grok3-think-20250330-0722.md

grok3-think-20250330-0724.md

grok3-think-20250330-0725.md

llama-3.3-70b-20250330-1547.md

llama-3.3-70b-20250330-1549.md

llama-3.3-70b-20250330-1555.md

llama-3.3-70b-20250330-1556.md

llama-3.3-70b-20250330-1557.md

nova-pro-1.0-20250330-1246.md

nova-pro-1.0-20250330-1247.md

nova-pro-1.0-20250330-1530.md

nova-pro-1.0-20250330-1531.md

nova-pro-1.0-20250330-1532.md

o3-mini-20250330-0746.md

o3-mini-20250330-0748.md

o3-mini-20250330-0752.md

o3-mini-20250330-0757.md

o3-mini-20250330-0758.md

o3-mini-high-20250330-0746.md

o3-mini-high-20250330-0748.md

o3-mini-high-20250330-0753.md

o3-mini-high-20250330-0757.md

o3-mini-high-20250330-0758.md

qwq-32b-20250330-1534.md

qwq-32b-20250330-1536.md

qwq-32b-20250330-1537.md

qwq-32b-20250330-1545.md

qwq-32b-20250330-1549.md