# Article information:

AI and compute  
<https://openai.com/research/ai-and-compute>

# Article summary:

1. The amount of compute used in AI training has been increasing exponentially, doubling every 3.4 months since 2012.

2. The increase in compute has been a key driver of progress in AI, allowing for the development of more complex models and achieving state-of-the-art results.

3. While massive amounts of compute are not always necessary to produce important results, the trend towards increased compute usage is likely to continue as researchers push the boundaries of what is possible with AI.

# Article rating:

Appears moderately imbalanced: The article provides some useful information, but is missing several important points or pieces of evidence that would be required to present the discussed topics in a balanced and reliable way. You are encouraged to seek a more balanced perspective on the presented issues by exploring the provided research topics and looking at different information sources.

# Article analysis:

The article titled "AI and compute" provides a detailed analysis of the relationship between artificial intelligence (AI) and computing power. The article presents two methodologies used to generate data points, which include counting the number of floating-point operations (FLOPs) in the described architecture per training example and looking at GPU training time and total number of GPUs used. The calculations are not intended to be precise but aim to be correct within a factor 2-3.

One potential bias in the article is that it focuses solely on the relationship between AI and computing power, without considering other factors that may impact AI development, such as data quality or algorithmic advancements. Additionally, the article only presents examples of papers that have used significant amounts of computing power, without exploring papers that have achieved notable results with modest amounts of compute.

The article also includes unsupported claims, such as assuming a utilization efficiency of 0.33 for GPUs and 0.17 for CPUs based on the author's own experience. This assumption may not hold true for all cases and could lead to inaccurate estimations.

Furthermore, while the article acknowledges that its calculations are not precise, it does not explore potential risks associated with relying on imprecise estimations when making decisions about resource allocation or investment in AI research.

Overall, while the article provides valuable insights into the relationship between AI and computing power, it would benefit from a more comprehensive analysis that considers other factors impacting AI development and explores both sides of the argument regarding the importance of computing power in AI research.

# Topics for further research:

* Factors impacting AI development beyond computing power
* Algorithmic advancements in AI research
* Data quality and its impact on AI development
* Examples of AI research with modest amounts of compute
* Risks associated with imprecise estimations in AI research
* Importance of considering multiple factors in AI research and resource allocation

# Report location:

<https://www.fullpicture.app/item/21702adc18cfcea38efc342cd46e69bf>