[Your name]'s AGI scenario

Your name (optional; please add your name to the file name if not anonymous):

Your scenario has automatically been shared with everyone in [folder]. If you prefer not to share it, feel free to remove it from that folder.

Can we print your scenario for others to look at? **Please update the file name of this file accordingly.** E.g., if you're up for it being printed, remove "no print" from the file name.

Your task: Write your own AGI scenario, depicting how AGI is eventually created in year X, and what happens after. The goal is NOT to entertain anyone or prove any sort of point; instead, the goal is to be your genuine best guess as to what the future will look like, depicted in as much detail / spelling out as much content as you can manage in 75 minutes. Err on the side of being specific and concrete. Any such specific forecast will inevitably be inaccurate, so please aim to provide the most accurate possible scenario, *given that it is highly specific*.

Why a scenario? We've found that thinking about the future in a specific, concrete scenario sparks interesting disagreements and allows us to reason more intuitively about the future.

Stage 1: Runup to the first AGI: Spend 35 minutes [to organizers: ideally use longer if you have time]

When and how will the first AGI system be created?

Suggested definition of AGI: An AI system as skilled at every *cognitive* task as the best humans, while being faster and cheaper than human labor.

Possible questions to answer:
- When will AGI be created?
- Which actor will be the first to create AGI? How far behind will other actors be?
- What training method will be used to create AGI?
- What safety techniques will be used when training AGI?
- Which actors will know about AGI soon after it is created? How will they know and how will they react?
- Will any actors successfully steal the AGI weights and/or algorithm secrets?
- What goals does the AGI have (to the extent it is correct to describe it as having goals)?

Your scenario:

Stage 2: After the first AGI: Spend 40 minutes [to organizers: ideally use longer if you have time]

What will the implications of AGI and further superhuman AI systems be?

Possible questions to answer:
- Do the AGI systems automate AI research? By how much is AI research being accelerated?
- Does AGI progress to ASI (artificial superintelligence)? Why, or why not? How long does it take? Just how capable do the AIs get?
  - *Suggested definition of ASI:* An AI system significantly better than the best humans at every cognitive task, e.g. as far away from the best humans as the best are from the median professional, while being faster and cheaper than human labor.
- How do the actors with access to AGI use it? When is it publicly released via API or product?
- How do superhuman AI systems affect the geopolitical landscape?
- Is there significant international coordination? If so, what are the specifics?
- What goals do further iterations of superhuman AI systems have?
- How does the story end? What happens to humanity on Earth and what happens with space resources?