Notes

- If you intend to work on a doc as main author, don't forget to register in this table, and as a comment on the article doc, that you're doing so.
- If you write articles in the course of the write-a-thon, and if you write on Stampy in general, your contributions are released under a permissive copyright.
- "Write-a-thon" and "distillation hackathon" and "hackathon" all refer to the same thing in this context.
- "??" means it's not yet definite.

How to Participate

- Pick a **question** to work on. (If you happen to have a particular subject related to AI safety that you know a lot about, we encourage you to write about that; you can ask on Discord if articles on it already exist.) Your options are:
    - Take one from the short list of suggested questions
    - Use a keyword to find another question already in the system. For example, if you want to see if there's a question in progress about Cartesian frames:
        i. Go to aisafety.info
        ii. Search for "cartesian frames"
        iii. Click "I'm asking something else" when no live articles show up
        iv. Click on the "edit" button next to one of the in-progress articles that show up to go to the doc
    - Take one from the full list of questions on Coda, as long as it has a status of "Not started", "In progress", or "Bulletpoint sketch"
    - Think of a new question entirely (but first ask in the #distillation-write-a-thon channel on Discord if it's a good fit for the site)
- Make sure there's a **doc** for the question. (There's an editing guide we'll be using for the usual Stampy workflow during the fellowship; some of it doesn't apply to the hackathon, but feel free to optionally look at it for more context.)
    - If there's an existing doc in the Stampy system (linked from the suggested questions doc above, or from Coda), use it. Make sure the doc has not been claimed by other participants, or seen significant work in the last ?? month (if there's an old draft, feel free to move that draft below a line that says "Scratchpad" and start over)
    - If not, ask on Discord for a doc to be created in the Stampy system.
    - If going through this doc acceptance process is a blocker, or if you prefer to work freeform and to worry about fitting your answer into the system later, there's a hackathon directory where participants can create and work on docs (participants can ask to be given directory-wide access).
- Write an **answer** to your chosen question, usually a few paragraphs with links to outside resources (and internal links to other Stampy docs, if you want, though we can add them

later instead). Take a look at existing articles for an idea of what to aim for, and see the style guidelines for more details on what text to write.
    - You're allowed to use an LLM to write a first draft of your article (we suggest this one; be aware that it logs questions where some of us can see them), as long as you mark this with a comment and you take responsibility for any hallucinations that the LLM generates (don't do this unless you know enough about the subject to be able to judge whether it's hallucinating).
- During the hackathon, articles should usually have **one main author** who submits it and has the final say[1] over the contents. Probably we'll give the main author of each document edit access to the document, and others will have the ability to use suggestion and comment mode.
- At the same time, we encourage people to **collaborate** on their articles. Hang out in the #distillation-write-a-thon channel on Discord and in gather.town (the Schelling point is the "alignment ecosystem development" room, but people can spread out into smaller working groups) to see what articles are in progress, and get on calls about them or just make comments and suggestions on Discord or the doc. Helping people edit their articles is an important part of your fellowship application, if you're making one. We'll also take it into account as part of your entry for prizes.
    - It's probably a good idea to create a thread on the article you're writing in the #editing channel, to serve as a central point for discussion on the article (along with the Google doc). You can tag @feedback in the Discord thread if you're looking for feedback, and you can tag @reviewer if you're highly confident in the article after getting a lot of feedback. (Tagging these roles is the usual process for getting an article live on site, but we mostly aren't trying to get articles all the way to live on site during the event, so it's very optional here.)
- **Repeat** the above for as many questions as you feel like. You can apply for the fellowship while submitting only one article or even none, if you were helpful to others. More articles do help your application and increase your chances to win a prize. The application will ask which other participants were helpful in editing your articles, so please try to keep track of who contributed.
- Collect the articles you've been the main author on and **submit** them. The same submission form will also let you **apply** for the fellowship, if you'd like to participate.
    - The form asks for links to docs you were the main author on, your name and contact information, and an ordered list of who was most helpful editing your articles. If you're applying for the fellowship, it also asks for some details about your availability and an ordered list of who you'd most like to work with.
    - Participants in the first fellowship are welcome to apply for the second fellowship without participating in the hackathon.
    - Everyone is welcome to participate in the hackathon (and submit articles for a chance at winning prizes) without applying for the fellowship.
    - The exact deadline is yet to be determined, but we will encourage participants to make their submissions by the end of the event, and then give them a few more days to finalize their applications if they need that time.

[1] "Final" for hackathon purposes; if we later publish it on the site, we may change it further.

    - **If you can't make it to the hackathon**, you can submit articles you wrote during one contiguous three-day period between the announcement and the hackathon. If you're taking this option, probably talk to StevenK#3458 on Discord for details.
- Your entry may win a **prize**. The best four entries will win $1000, $600, $300, and $100. A random entry (out of other serious entries) will win $200.
    - ?? Tentatively: This will be based on the collection of all the documents you were main author on, plus the judgments of other main authors of who was helpful in collaboration. Prizes are by participant, not by article, so you can win only one. ??

Schedule

- Friday, June 16th, 7am UTC: people can start writing if they want
- Friday, June 16th, 5pm UTC: introductory Discord/GatherTown group call
- ?? Saturday, June 17th, 5pm UTC: introductory Discord/GatherTown group call for people who couldn't make it to the other one
- ?? Sunday, June 18th, 11pm UTC: wrap-up Discord/GatherTown group call
- Monday, June 19th, 7am UTC: people have to stop working on their articles, and ideally will submit their articles and fellowship applications
- Thursday, June 22nd, 7am UTC: deadline for submissions and applications (form closes)
- ?? Thursday, June 29th, 7am UTC: deadline for us to accept or reject applications
- ?? Monday, July 3rd: deadline for us to announce prizes
- Monday, July 3rd: first day of fellowship
- Monday, October 2nd: last day of fellowship

Questions

If you have any questions, don't hesitate to ask in the #distillation-write-a-thon channel, or to message StevenK#3458, on Discord.