(A) Some questions I'm curious about include:

1. How much would each of the "risk factors for stable totalitarianism" reviewed by Caplan increase the risk of stable totalitarianism (if at all)? How likely is each factor's occurrence? What other risk or security factors should we focus on? How can we best influence these factors?
2. What can we learn from fields like political science, history, and economics regarding questions relevant to a "long reflection", such as:
    - How can we prevent any actors "defecting" from a long reflection?
    - Who should actively work on the long reflection? How? How can we incentivise this?
3. How does the resolution and calibration of forecasts vary by the forecasts' "range" (i.e., the length of time between the forecast and the event it's about)? How does this vary between topics, types of forecasters, forecasting approaches, etc.? Does above-average calibration or resolution for short-range forecasts generalise to longer-range forecasts?
4. Which forms of world government, or of movements towards world government, are most likely? Which are best for the long-term future? What can we do to influence these things?
5. What predictions can we presently make about aspects of space colonisation (e.g., its likelihood, timing, key actors, key technologies, likelihood of involving biological humans and/or terraforming, and impacts on our ability to coordinate and "change course")?
6. How does telling people about the idea of a long reflection influence their longtermism-relevant views or behaviours (e.g., their inclination to support existential-risk-reducing policies or charities)?
7. Were changes in the offence-defence balance a major contributor to historical declines in levels of armed conflict? If so, does that mean we should be less confident that levels of armed conflict will continue declining?

(B) Some projects I think might be valuable include:

1. Research into any of the above questions
2. Research or writing assistance for more senior researchers at FHI (or GPI or Forethought)
    a. This might allow them to complete additional valuable projects
    b. I'd be happy to provide this assistance because:
        i. I think this would be impactful
        ii. I think it would be a good way to build useful knowledge and skills
        iii. I have a suitably broad range of skills and interests
        iv. I'm as happy facilitating others' impact as I am "having an impact myself"
3. Further summer research fellowships, "Early Career Conference Programmes", or similar at FHI, CEA, GPI, and/or Forethought
    a. I'd be happy to assist or lead in coordinating this sort of thing, given my interest in roles at the intersection of research, research facilitation, and operations
    b. I'm unsure what things of this kind are already planned, what would be worthwhile, and exactly what form these things should take
    c. It could possibly be worth integrating or synchronising fellowships/programmes across more than one of those organisations
4. More seminars, other events, or one-on-ones to further connect FHI staff (including RSP participants) with university students at Oxford (or elsewhere)
    a. But perhaps enough is already being done on this front
5. Optional forecasting tournaments for FHI, CEA, GPI, and/or Forethought staff, perhaps with questions tailored to their areas of interest
6. Opportunities for RSP participants (and perhaps summer research fellows) to assist with, test fit for, and/or gain skills in grant-making, such as
    a. An internal FHI donation lottery
    b. Something analogous to the Harvard EA Philanthropy Advisory Fellowship

(C1) It could be valuable to develop one or more models of the likelihood of various types of recovery given various types of collapse scenarios, following roughly the following steps:

1. Think about how to carve up the possible causes of collapse, types of collapse (e.g., loss of population, industry), and types of recovery (e.g., recovery of GWP, political institutions), and what to focus on.
2. Begin to construct in Guesstimate one or more models of the odds of various types of recovery, given various collapse scenarios.
3. Begin to estimate the parameters of the model(s).
4. Iteratively refine the model(s), estimates, and conceptual underpinnings.
5. Present the results in EA Forum posts and possibly papers.
6. Perhaps iterate further based on new feedback from the GCR and academic communities.

Steps 1-4 could involve armchair reasoning, reading relevant writings, and getting ideas and input from longtermists and academics.

Why this might be valuable

Some experts and EAs appear to think there's a nontrivial chance of civilizational collapse. Some further argue that reducing the risk of collapse, or increasing the chance of recovery, should be a top priority. Others argue that that wouldn't be worth prioritising *even if* there's a nontrivial chance of collapse, based on the premise that the chance of recovery is already high. This difference in views on the odds of recovery may be influencing, or *should* influence, large and growing pools of money and talent. This project seems like a tractable and neglected way to improve our views on the odds of recovery.

Key concerns

- Perhaps someone else is already doing something similar
- Perhaps the model would end up extremely elaborate and yet still inadequate
- Perhaps it would be better to first model in more detail one specific collapse and recovery scenario
- Perhaps it would be better to first investigate in more detail some key parameters or causal steps
- Perhaps it would be better to first develop models of the odds of various collapse scenarios occurring in the first place
- Perhaps this project should be conducted by someone with a different background, knowledge set, or skill set to me

(C2) It could be valuable to investigate the extent to which moral circles expanding towards inclusion of one being affects the likelihood of inclusion of another being, and how this varies between different pairs of beings (e.g., farmed animals and wild animals vs. farmed animals and nonbiological minds).

Why this might be valuable

Some longtermists believe roughly the following argument:

1. The vast majority of all valenced experience that occurs may be experienced by beings towards which humans would, "by default", exhibit little or no moral concern.
2. It may thus be important to expand moral circles such that they're more likely to later include those beings.
3. This could be done by expanding moral circles to include farm animals, and/or ending factory farming.
4. We should thus allocate some longtermist resources to factory-farming-related work.

I think that that argument is plausible but speculative. This project would inform us about the third step of that argument. It could also indicate which alternative approaches to moral circle expansion may be better, and which activities could increase moral concern for future generations (which could influence existential-risk-related views and behaviours).

Concrete approaches I could take

- Reviews of relevant literatures
- Expert interviews
- Surveys investigating the extent to which moral concern for one being correlates with moral concern for another being
    - Ideally, these surveys would be longitudinal, to see how *changes* in concern are correlated
- Experiments in which (some) participants are shown messages intended to increase concern for one being, and participants are asked questions relevant to their concern for those and other beings
- Historical case studies investigating whether increases in concern for one being led to or co-occurred with increased concern for other beings
- Quantitative research on collections of such historical cases

Key concerns

- Perhaps others have done or will do similar research, or would do it if I simply encouraged them
- Perhaps we can already confidently dismiss the four-step argument above
- Perhaps this research would give "pessimistic" results even if that argument was true, because the relevant effects would only occur at a point such as the end of factory farming
- It will likely be hard to determine which correlations are causal
- Perhaps people's current, self-reported attitudes on (e.g.) nonbiological minds would bear little resemble to their later, revealed attitudes