AI Safety work Compartmentalization

- AI Comms (communications) space expansion
  - 1. AI risk arguments database in mainstream media
    - tangible work to be done: a spreadsheet of 1. argument 2. news source 3. direct quote
    - tangible work to be done: a spreadsheet of 1. AI-risk journalist name 2. journalist's email 3. special attributes of said journalist
  - 2. AI risk argument prioritization research (Which AI risk arguments should be mentioned before all others?)
    - tangible work to be done: a ranked list (1-2 pages) of present AI risks arguments, explanation (3-5 sentences) for why they should be prioritized.
  - 3. Open letter writing
    - tangible work to be done: 1. conduct literature review about how open letters work. 2. Explain how to make them effective 3. draft an open letter about AI risk
- AI Governance
  - 1. Visualization of AI Governance landscape using precedent examples (ie. creation of FDA)
    - tangible work to be done: 1. 1-5 pages of research on past government intervention on new/emergent technology 2. hypothesize how analogous AI governance policy will be comparatively using standards, collaboration, and methodology extracted from these historical precedents.
  - 2. Model-sharing policy research
    - tangible work to be done: 1. 1-2 pages or a list of items to consider when releasing a model (ie. ChatGPT, GPT-4). Some examples are race dynamics and influence of culture. 2. Include any extraneous factors like how and when to share models, how widely to share (just researchers or anyone?), motivations to increase/decrease accessibility
  - 3. Export controls research
    - tangible work to be done: 1. research about current export control policies 2. ask GPT-4 if export control is a promising intervention in the future
  - 4. Database of Behavioral science labs and Institutionalized decision making labs
    - tangible work to be done: 1. a spreadsheet of current labs working on alignment research/ intervention methods research 2. In the spreadsheet, be sure to specify these items (research focus, why you think it is beneficial, team size/key figures, how these behavioral science + institutionalized decision making labs can collaborate with AI labs)
- AI Safety Field Building
  - 1. Help people
    - tangible work to be done: 1. helping people scan through messages 2. filter important people to talk to 3. to be trained to do one-on-ones well
  - 2. Database of Information Security Professionals
    - tangible work to be done: 1. create a spreadsheet of people who are working on information security
  - 3. Education of AI Safety importance
    - Tangible work to be done: 1. A 1-3 page report with resources, similar to AGI Safety Fundamentals
  - 4. Advertisement of fellowships and programs
    - Tangible work to be done: 1. increase publicity of SERI MATS, ERA, CERI summer fellowship, and more. by whatever means possible.