Pre-Phase 1 Requirements

1. Briefly describe the purpose of the AI agent, including key population(s) targeted, how you intend the population(s) to use the system, estimated cost, and measurable outcomes you seek to improve:

2. Name up to 10 internal experts that need access to this tool for testing purposes, and justify their capability to judge the accuracy of the agent's responses:

3. List the accountable owner's name and title:

4. Describe the secure, vetted agentic cybersecurity environment used. Mention how the environment limits the agent's ability to communicate externally, retrieve external content, and/or conduct autonomous system actions not relevant to its purpose (this section should ideally contain stock or preapproved language from the information technology [IT] team):

5. Name and provide the title of the cybersecurity official that provided or approved the secure agentic environment (this should ideally be set by default by the IT team):

6. The accountable owner should set performance targets for the pilot according to the following metrics. These should represent the minimum values necessary for the tool to be useful enough to pass into more rigorous testing in phase 2.

    a. **Correctness: %.** The share of responses that are factually accurate and do not introduce errors.

    b. **Groundedness: %.** The share of responses that are supported by the allowed sources or context, cite sources, and do not invent unsupported claims.

    c. **Explainability and clarity: %.** The share of responses that provide evidence for the agent's output that is meaningful to the intended user, accurately reflects the system's process for generating the answer, and clearly communicates the agent's knowledge limits.

    d. **Completeness: %.** The share of responses that correctly cover all necessary elements of the request—no key steps, constraints, or required details are omitted.

    e. **Reliability under variation: percentage points.** The difference in correctness across reasonable variations in phrasing, order, multi-turn context, or scenario paths.

    f. **Cost per completed agentic task: $.** The total cost for the agent to complete all necessary actions for the given task.

    g. **Usage: # .** The total number of completed tasks.

7. Check the boxes below to affirm the following:

    ☐ No sensitive or confidential data have been used (dummy data used if necessary).

    ☐ I will make this document, apart from any sensitive security information, available to all users of the agent ahead of testing.

    ☐ I have notified pilot testers, and they understand that this is a limited-term pilot.

As the accountable owner, I affirm the above to be true and accurate as of

Accountable owner signature: _____

As the responsible agentic AI lead, I approve the pilot above for no more than six months as of

Responsible agentic AI lead signature: _____