Scaling Trust - FAQ

Technical Scope + Research Focus

Are there specific guidelines of what kind of open-source tooling you would be prioritizing in Track 2? Is it just orchestration frameworks, or would you also be interested in interpretability infrastructure for monitoring agent traces?

Orchestration will be a priority. Interpretability falls under 'reporting' and will also definitely be important.

Is the focus of this call meant to be on agentic AI systems like those trending today or can the "agents" also be more traditional AI implementations that are more akin to automation tools and can those tools interact with each other and work towards one common goal?

One of our core goals is generalisation, and so we care about tools that are built which can generalise to a lot of use cases. More traditional agents are usually application-specific. If they're not, and they can generalise well, we're excited about these.

Regarding technology readiness level, what TRL levels are in scope? By TRL I mean UKRI definitions

The goal of the programme is to surface the current state of this technology (we expect it to be around TRL 3-5). The aim of the programme is to get these technologies to a TRL 7.

Is operation in adversarial environments a fundamental requirement?

Yes, we want to make sure autonomous agents are secure in the 'wild' - i.e. in places where other agents might be adversarial.

Is the funding call looking for applications beyond cryptography, such as fault identification in machineries using AI or monitoring of built infrastructures (buildings, bridges) using agentic AI?

This is in scope for Track 3.2 - Cyber Physical Primitives. In addition, some of the Arena (Track 1) challenges will be cyberphysical, where this could be required.

By trust/trustworthiness, do you just refer to the classical definition in crypto/cyber-security or is there scope for trust as considered in the area of

agents and multi-agent systems, i.e. socio-technical definitions, socio-cognitive properties of human-machine interactions?

A large portion of our definition of trust/trustworthiness is in the cryptographic/cyber-security sense. However, if socio-cognitive/socio-technical definitions would be a bottle neck to trusting agents (and if you can make the case), we welcome these proposals.

Can the proposal focus on a specific domain (e.g., healthcare, or commercial organisation), or is the expectation that agents should operate within any environment?

You can mention specific domains where this would be useful, but generally we're interested in any environment. We're interested in a general capability. Where a capability can be domain specific, we are interested if it can generalise well. The Arena is where these capabilities will be tested, and these challenges will be designed explicitly for this generalisation to avoid application specificity.

Is work on theoretical foundations of trust in AI within scope for Track 3.4, or is the focus strictly on cryptographic and security primitives?

Non-cryptographic/security primitives are welcome as long as they can contribute to our goal of making AI agents more 'trustworthy'. That being said, this is not an AI safety programme, so if the proposals are too safety-focused, they'll be disadvantaged (we encourage you to look at the ARIA Safeguarded AI programme in that case.)

I want to understand a bit more about how you view swarm level metrics and whether they are within the scope of this program. For example, is there interest in developing theory and methods for understanding the dynamics of agents in the arena in aggregate? Or is the focus much more on individual agent-agent interactions?

The idea of understanding dynamics of agents in the arena in aggregate sounds great and something we're definitely interested in.

How flexible is ARIA supporting funding for work beyond conventional AI streams - such as mechanical/aerospace/civil engineering using agentic AI systems - which have focussed high-quality impacts?

The ultimate goal is to build tools that are useful, given the challenges running in the arena. The challenges may be related in some way to the streams mentioned. However, if you're building tools that are not useful for one of these streams, and they don't generalise, then it is most likely to be out of scope.

Can the funding proposal focus on agents or does it need to include AI as well?

We're interested in the capabilities that are developed. We believe that the most promising of these come from AI systems that have gained traction in the last few years, however there may be other tools out there which we're also open to. We're interested in impact over specific techniques.

What is meant by the term "adversarial" in the call please e.g. a challenging environment versus the military term?

We use adversarial in the more game-theoretic sense that involves parties with opposing interests.

Is the Report component trace schema already defined or is defining it part of the funded work?

We haven't defined a specific schema yet, and there are definitely different approaches here that we may need.

Is hardware security research within the scope of this call against the fundamental research category?

Yes, it is under scope under Track 3.2 - Cyber-Physical Primitives. We're generally interested in using secure/tamper-resistant hardware, as well as potentially creating new primitives here.

Is there room in Track 3.4 for economic/market-based security primitives as a complement to cryptographic verification?

It depends on how this is presented. If it's complementary to other technologies we're interested in, then it's definitely interesting to us.

Nature Crypto:do you consider biological or biohybrid primitives (living-cell or vesicle-based fluorescence responses) in scope?

Yes, if it is related to agentic coordination. If it is a pure primitive that does not have an application to agentic coordination, it will fit better as a part of our opportunity seed call. More information on the opportunity seed call will be released in the following months.

Should the design of cyber-physical primitives explicitly account for future integration complexity with the CPS Arena?

Ideally yes, however we're starting this solicitation without a cyber-physical arena, and so it can be more general for now. If you come up with new cyber-physical primitives, we will want to test them in the arena at some point.

Would extending existing interpretability tools to support agent monitoring at scale and interoperability with orchestration/evals framework be eligible for Track 2 funding?

Yes.

Eligibility + Location

What does positive impact for the UK mean in the context of funding those outside of the UK?

Our primary focus is to fund those based in the UK. However, we also accept applications from non-UK applicants or applications that are made up of a combination of UK and non-UK applicants.

A strong benefit to the UK is where 50% of the project costs are being spent in the UK. This could be where there is a UK lead, collaborating with someone outside of the UK (or the other way around).

However, if you don't fall within this, we still want to receive an application from you.

Additional Information
As part of our review process, we will assess proposal against a number of criteria (outlined in the solicitation document), including 'Benefit to UK':

There is a clear case for how the project will benefit the UK. Strong cases for benefit to the UK include proposals that:

a. *are led by an applicant within the UK who will perform the majority (>50% of project costs spent in the UK) of the project within the UK*
b. *are led by an applicant outside the UK who seeks to establish operations inside the UK and perform a majority (>50% of project costs spent in the UK) of the project inside the UK and present a credible plan for achieving this within the programme duration.*

For all other applicants we will evaluate the proposal based on its potential to boost the net impact of the programme in the UK. This could include:

c. *A commitment to providing a direct benefit to the UK economy, scientific innovation, invention, or quality of life, commensurate with the value of the award;*
d. *The project's inclusion in the programme significantly boosts the probability of success and/or increases the net benefit of specific UK-based programme elements, for example, the project represents a small but essential component of the programme for which there is no reasonable, comparably capable UK alternative.*

When considering the benefit to the UK, the proposal will be considered on a portfolio basis and with regard to the next best alternative proposal from a UK organisation/individual.

We are lecturers and researchers outside of the UK, employed at an offshore campus of a UK university. Does this make us eligible?

Yes, you are eligible to apply, noting the criteria for non-UK funding above.

Our research team is based in Imperial College London. While the development team is elsewhere. If the core is in the UK, is that OK?

Yes that is ok, noting our criteria on non-UK funding mentioned above.We also encourage you to think about what structure works best for the team and the project.

Are there resources for non-UK applicants to partner with UK teams/orgs to ensure impact in UK and up acceptance odds?

Yes, please see our Funding page for access to our Teaming Tool. We also have a community discord where you can find collaborators.

Additional Information

We still encourage you to apply even without one, as we absolutely can still fund projects outside of the UK. Additionally, if you submit a concept paper, we may be able to facilitate an introduction to a suitable UK-based collaborator. This could lead to a fruitful collaboration moving forward.

For projects relocating a researcher to the UK, should visa and relocation costs be included in the project budget?

Yes, please include these costs. We can pay reasonable visa and relocation costs. We're interested to hear about people that want to relocate to the UK, and we have some guidance on our website around potential visa routes (see the 'Non-UK applicants' section of our FAQs here.

Team Composition + Solo Applicants

As an individual PI, what are the practical mechanics for receiving and managing grant funds, is a LTD Company required?

ARIA can fund unhosted individuals, so you don't need a host organisation or to be a limited company to receive funding from ARIA. However, there are a few considerations you might want to take into account, for example, if you have collaborators or

subcontractors, you might wish to set up a limited company or register as a sole trader, given the potential tax implications that you might want to consider.

ARIA cannot advise you directly on those kinds of tax considerations, but we would encourage you to seek some independent advice if you are selected for an award, but are an unhosted individual. We'd also like to highlight that everyone that we fund is subject to due diligence checks, so as an individual, you will also be subject to these checks.

Can AI agents be formally listed as team members, or should they be framed as tools/infrastructure in the application?

You can list them as team members if you'd like to in your application. We just ask that you make that clear as part of that section of the application.

Are small teams (PI 25% + postdoc 80% + postdoc 50%) typical/competitive for exploratory projects?

We ask that enough time is dedicated to the project, and we prefer people that are committing most of their time to the project. It also depends on the size of the grant, and other relevant factors. We care more about your credibility and whether you can deliver on the things that you've mentioned in the proposal. We also care about how exciting your proposal is and how value-aligned you are with the programme.

What if I don't have a team but I am open to building a team as I go along? Does this disqualify me or reduce my chances?

Not at all, all we care about is your ability to fulfil the proposal. If you can do so by yourself, and you use the grant to get more people to join the effort and force multiply, then great.

However, if you start solo, and require additional people to fulfil the proposal, then there's a delivery risk to the project and a question arises as to whether you fulfil your proposal. This will not disqualify you, but is something we'll want to look at more carefully.

Is it suggested to form a team to apply for the funding?

We're focused on impact. If you can deliver on your proposal as a solo applicant, then great. If you need a team to deliver, then that's also great.

Advanced Research + Invention Agency ARIA

Application Process + Strategy

Benchmark datasets with novel taxonomy seem to fit into a track 2 application, would empirical frameworks to generate such datasets fit into track 3 as a separate proposal? Or as an extension of a track 2 route?

For some applications, we expect the line to be blurry. Tracks are to be treated more like guidelines than strictly enforced areas in these cases, and we expect some applications to be useful to both. Where you think this is the case, please explicitly state this in your proposal.

Is it possible to jointly apply for the exploratory and the research center funding? (Assuming that in case of receiving the research center funding, the exploratory grant would be turned down, of course)

Yes, if these are two different proposals. A research centre proposal requires a different level of detail to a proposal for an exploratory grant, as you'll be applying for more funding. Proposals that aim to cover both will be ill-fitting.

Would you encourage applications to clarify/commit to a track in the application or highlight potential for relevance in either?

Applicants wishing to apply across multiple tracks should submit separate proposals for each track, with cross-references to any related proposals.

Can teams submit multiple proposals (in different tracks), and if so, would that affect how the individual proposals are viewed?

Teams can submit multiple proposals in different tracks. This will not affect how the individual proposals are viewed. If you can demonstrate your ability to help reach the programme's goal, that is all that matters.

Would it be possible to apply for this grant even if the research project could not start until January 2027?

We're biasing towards the schedule as laid out in the solicitation. We're trying to move fast, given the rate of change in this field. However, if a proposal is extraordinary and existing commitments mean they can only start later, we're happy to consider it on a case by case basis.

We have a working prototype deployed across multiple platforms. Is early feedback on concept papers available before March 24?

We're happy to tell you if something is in or out of scope. If you're happy to share the idea publicly, you can do so on our Discord. If not, please send the question via the chat function via the funding page and we'll get back to you.

Funding, Duration + Commercials

What's ARIA's view on having a path to commercial viability?

A core goal of this programme is to get adoption of the tools developed, and so a path to commercial viability is extremely important. We will be releasing an adoption track next year. So definitely important (read thesis for more colors on our goals). There is more detail on this topic in our Thesis.

The call for proposals says that track 2/3 applications are for a duration of 3 to 18 months, with potential for renewal. Could you please elaborate on the expected duration of such renewals, the likeliness and intended decision process for renewals, and whether such potential renewal topics should be listed somewhere on the application?

We're approaching the programme iteratively and we see the grants that we're giving now as the first batch of grants, and we expect to give more grants in the future. For the projects that are doing well, and where more money could really help, and this money could come from us and not from VCs for example, we want to consider that seriously.

We are planning a second solicitation for Tracks 2 and 3. Details on this will be formalised, but we're planning to run a similar solicitation mid 2027 where you'll be able to apply for further funding. We'll have continuous discussions with funded teams to discuss potential routes.

You mentioned another batch of funding, do you have any estimate for a date?

We are planning to launch a second solicitation mid-2027.

Is there any ROI expected by ARIA?

We do not expect a direct/financial return on our grants. However, we do expect to see the impact of what we're doing – for the UK, and for the world.

Additional *Information*

If you assign or licence ARIA funded IP to a non-UK entity, or if you are a non-UK entity and commercialise ARIA funded IP, this will be subject to a small royalty fee. To see a copy of our funding agreements, please see the 'Funding' section of our FAQs here.

What distribution of project sizes do you expect to fund?

Please see the solicitation for more details on project size. The summary section on page 2 summarises this.

Evaluation + Impact

What does a strong Track 2 application look like to you, is it more important to demonstrate technical novelty or real world deployment evidence?

Track 2 has a dual-goal: develop tooling that scores well in the arena and does well in the real world. However, if you do have something that is novel and more experimental – with a path to real world application – we will be interested in it.

What kind of proof of feasibility do you want to see for track 2? Simulations or pilot study or market research?

Proof of feasibility for track 2 tooling will be the use and performance of the tooling in the Arena.

For Track 3, do you expect primarily theoretical contributions, or should theory be tightly coupled to near-term experimental validation?

A mix of both. We don't want only theory that's tightly coupled to near-term experimental validation because we see the value of fundamental research as something that is not always tied to near-term experimental validation, yet can have foundational impacts on a whole field. And right now, a lot of what's happening in artificial intelligence is very empirical, and so we want to make sure we don't lean too much towards immediate applicability.

That being said, we do want our three tracks to work together. We want the fundamental research and researchers to talk to the research engineers and engineers in tooling, and also to think about what's happening in the arena. We expect some type of interaction between these different groups, and for them to see the value in potentially working together.

A lot of the time, what you can do in theory can be completely meaningless if it's not guided by some purpose and a lot of the time seeing things in applications can really help guide new, deep research questions. So there's something about that that matters to us. That being said, we don't think we will rule out if your proposal doesn't have near-term experimental validation, necessarily.

How important is team track record versus originality of idea?

We're focused on impact. Track record is important, in that it is evidence that you can deliver the things you're promising to do. Originality of idea is also great, if it's impactful.

How much emphasis will be placed on immediate usability for the Scaling Trust Arena versus longer-term foundational impact?

It's a mix and it depends. We are looking at this from a portfolio level. We don't want things that are only usable in the arena, and disqualify things that have a longer-term foundational impact.

We ask you to write the proposals you're most excited by, and as we review the proposals, we'll know how this all meshes together.

When encouraging projects that may reach "success (or failure) on faster timelines," should proposals include early go/no-go?

Wherever possible. We're focused on impact, so that the faster projects can validate that they're on the right path, the better. If you can identify this in your proposal, it makes it more mature and more nuanced, which ultimately makes us more excited to work with you.

Collaboration + Support

Will there be an opportunity to meet others to group together in a team?

Yes, we have the teaming platform, and the community Discord. We're also trying to be very flexible, so if there are many people who feel like they would benefit from an in-person teaming day, we can help organise this. If so, please signal your interest in our teaming channel in Discord.

Please say more about the involvement of industrial partners - relevance and role etc?

This is something that will be more relevant further down the life of the programme. As we have initial proof points for the work that is taking place, we'll start thinking about bringing in different industrial partners to forward deploy this technology.

How will the cohort of performers work together, any interaction?

Yes, we are planning a set of community interactions as listed in the solicitation. We'll use shared communication channels and we'll host regular events that bring people together. Generally, we seek to foster a lot of collaboration, a lot will be coordinated by us, but we want you to be owners.

What mechanisms will exist for Track 3 exploratory projects to collaborate with Track 3 research centres during the programme?

There will be broad mechanisms facilitating collaborations, some funded by us, some not.

What support will be available from Aria side post grant approval?

We expect to work closely with all Creators, so we expect regular catch ups. We'll also be available ad-hoc, depending on where you need support, whether it's validating a new business idea or unblocking project activities.

IP + Open Sourcing

Can Track 2 deliverables include empirical deployment testing data in conjunction with open-source code?

If you have code, but also testing data that includes how the code performs, and you want to share that as well, that's great.

Are proposals to formalise and open-source existing working systems in scope?
We're more excited about funding new work than paying people to open-source existing IP. However, if it's a useful component in what we're doing, and everyone could benefit from it or build on top of it, then it's something we'd consider on a case-by-case basis.

What position on IP and publishing for Track 3, especially if outputs include open datasets/protocols and a working prototype?

Ideally this would all be open, with the filter of it being ethically responsible to open-source.