

Conjuntos de dados inclusivos para aprendizagem automática

AMR

Webinar para candidatos

10 de julho de 2024

Altifalantes



MERCY AYUB, ACTS

ADELA BRIANSÓ JUNQUERA, LACUNA FUND

SAMUEL WANJAU, ACTS

DR. OJENGE WINSTON, ACTS

**Dra. ELIZABETH MUTUA, CONSULTORA DE QUALIDADE
DE DADOS DA ACTS**

Objectivos da reunião

- Proporcionar aos potenciais candidatos uma compreensão do Fundo Lacuna e dos requisitos da proposta.
- Dar tempo para que os candidatos possam fazer perguntas sobre o Pedido de Propostas (RFP).

Agenda

- 00:00** **Boas-vindas, apresentação,** revisão da ordem de trabalhos
- 00:10** **Apresentação: Visão Geral do Fundo Lacuna e Requisitos para RFPs, Elementos de uma Proposta Forte**
- 00:40** **Apresentação: Considerações sobre a qualidade dos dados e o alojamento**
- 00:50** **Apresentação: Oportunidade de tutoria/compartilhamento**
- 01:00** **Perguntas e respostas**
- 01:25** **Próximos passos**
- 01:30** **Adiar**

Uma lacuna é uma lacuna ou uma parte ausente de um manuscrito

O Lacuna Fund apoia a criação, expansão e manutenção de conjuntos de dados de formação e avaliação equitativos que permitem que as ferramentas de aprendizagem automática resolvam melhor os problemas urgentes em contextos de baixo e médio rendimento a nível mundial.

Princípios que orientam o Fundo Lacuna:

- Acessibilidade
- Património
- Ética
- Abordagem participativa
- Qualidade
- Impacto transformacional



Financiadores



Outros financiadores



Fundo Lacuna

Estrutura de governação

Comité de Direção

Orientação estratégica e governação global do Fundo

Colaboração com o financiador

Fornece informações sobre as principais prioridades ao Comité de Direção e proporciona um fórum para uma colaboração mais profunda entre os financiadores

Painéis de aconselhamento técnico

Abranger os processos de financiamento e selecionar propostas, contribuir para as estratégias e necessidades específicas do domínio

Secretariado

Gere as subvenções e a elaboração de relatórios, facilita a comunicação e presta



Domínios para conjuntos de dados



O Fundo Lacuna fornece recursos para a criação, expansão ou manutenção de conjuntos de dados.

Convites à apresentação de propostas em curso nestes domínios:

- **AMR**
- **PNL**

Convites anteriores à apresentação de propostas nestes domínios:

- Agricultura
- Língua
- Saúde
- Clima

Futuros convites à apresentação de propostas:

- Outros domínios em que o Comité de Direção identifique uma necessidade

Financiador para AMR

O convite à apresentação de propostas AMR é possível graças ao generoso apoio da WELLCOME.





Requisitos para a RFP

Resistência antimicrobiana (AMR)

AMR RFP - Objetivo

Objetivo declarado:

O objetivo deste convite à apresentação de propostas é apoiar os esforços de desenvolvimento de conjuntos de dados abertos e acessíveis para aplicações de aprendizagem automática relacionadas com a compreensão da resistência antimicrobiana (RAM) em países de baixo e médio rendimento (PBMR) em África, na América Latina e no Sul e Sudeste Asiático.

AMR RFP - Necessidade

Vemos a necessidade de conjuntos de dados de formação e avaliação que ajudem a compreender melhor os impactos da RAM e a desenvolver intervenções para mitigar esses impactos. Procuramos conjuntos de dados identificados por peritos locais, concebidos para responder a necessidades identificadas a nível local, pelo que os seguintes são apenas exemplos ilustrativos.

- Conjuntos de dados para melhorar a prevenção e o controlo das infeções (IPC)
- Conjuntos de dados para a descoberta de medicamentos
- Conjuntos de dados para a vigilância global da RAM e abordagens antecipatórias
- Conjuntos de dados que apoiam o diagnóstico
- Conjuntos de dados que contêm informações sobre o acesso aos agentes antimicrobianos e a falta dele
- Conjuntos de dados que melhoram a compreensão biológica básica dos agentes patogénicos e da resistência antimicrobiana
- Em todos os conjuntos de dados, a capacidade de resposta às questões de



O que estamos à procura

- Recolha e anotação de novos dados
- Anotação e libertação de dados existentes
- Expandir e/ou manter os conjuntos de dados existentes e aumentar a sua facilidade de utilização
- Ligação e harmonização dos conjuntos de dados existentes



Componentes da proposta

- Informações sobre o candidato
- Narrativa da proposta
- Cronograma do projeto
- Orçamento e descrição orçamental
- Formulário de auto-relatório de diligência devida



Requisitos de elegibilidade

- Os requisitos específicos do financiador são os seguintes
 - Ser uma entidade sem fins lucrativos, uma instituição de investigação, uma empresa social com fins lucrativos ou uma equipa de tais organizações. Os indivíduos devem candidatar-se através de um patrocinador institucional. As parcerias são fortemente encorajadas como forma de reforçar a colaboração e maximizar os benefícios derivados da utilização dos conjuntos de dados, mas apenas o candidato principal receberá fundos.
 - A entidade deve ter uma missão que apoie o bem da sociedade, definido em sentido lato.

Requisitos de elegibilidade

- Estar sediado no país ou na região onde os dados serão recolhidos. O candidato principal deve estar sediado num país de rendimento baixo ou médio e deve propor um trabalho no mesmo país ou num país de rendimento baixo ou médio em África, na América Latina ou no Sul ou Sudeste Asiático. Para obter uma lista dos países elegíveis, consultar <https://wellcome.org/grant-funding/guidance/low-and-middle-income-countries>.
- As instituições sediadas noutros países ou regiões podem candidatar-se como parceiros da instituição principal. Tal como acima referido, apenas o candidato principal receberá fundos.

Requisitos de elegibilidade

- Ter todas as aprovações nacionais ou outras necessárias para efetuar a investigação proposta. O processo de aprovação pode ser efectuado em paralelo com o pedido de subvenção, se necessário. Os custos de aprovação, caso existam, são da responsabilidade do candidato.
- Ter a capacidade técnica - ou a capacidade de desenvolver essa capacidade através de uma parceria descrita na proposta - para efetuar a rotulagem, criação, agregação, expansão e/ou manutenção de conjuntos de dados, incluindo a capacidade de aplicar as melhores práticas e as normas estabelecidas em matéria de resistência antimicrobiana para permitir que múltiplas entidades efectuem análises de IA/ML de elevada qualidade.

Datas importantes

27 de junho: Pedido de Propostas Aberto - [Candidate-se - Fundo Lacuna](#)

12 de julho: Prazo para a apresentação de perguntas

As perguntas devem ser enviadas para secretariat@lacunafund.org

29 de julho: Respostas publicadas

26 de julho: Última oportunidade para se

- **novembro:** O Painel Técnico Consultivo toma decisões sobre a atribuição de subsídios

- **dezembro:** Notificação dos beneficiários dos subsídios

- **maio:** Anúncios de prémios



Elementos de propuestas sólidas

Elementos de propostas sólidas

- **Equipa multidisciplinar** - experiência em ciência de dados e aprendizagem automática, AMR
- **Casos de utilização e envolvimento da comunidade**
- **Declaração** clara do **problema** e da forma como o conjunto de dados ou a agregação propostos ajudarão a resolver o problema
- Especificidade quanto à dimensão do conjunto de dados - deve ser de **dimensão e qualidade suficientes** para ser útil em aplicações futuras
- **O orçamento** é adequado à dimensão do conjunto de

Elementos de propostas sólidas

- Considerações **de equidade** (género, estatuto socioeconómico, etnia, etc.)
- Consideração e plano para potenciais questões **de privacidade e ética**
- Planear a **gestão de dados e o licenciamento**
- Trabalhar em vários domínios sempre que possível e pertinente: tempo, regiões, agentes patogénicos, resultados para os doentes...
- **O orçamento** é adequado à dimensão do conjunto de dados produzido



Envolvimento da comunidade

- Descrever as consultas anteriores e/ou a colaboração proposta com os beneficiários previstos.
 - Quando e onde se reuniu/encontrará com os parceiros
 - Como é que os parceiros serão envolvidos:
 - Identificação das necessidades de dados
 - Recolha de dados, etiquetagem
 - Governança de dados
 - Utilização do conjunto de dados
 - Como é que os parceiros beneficiarão do conjunto de dados novo/expandido



Privacidade e ética

Explique como é que a sua equipa vai abordar a questão:

- a) preocupações com a privacidade,
- b) potencial de utilização indevida a jusante,
- c) possíveis vectores de discriminação (por exemplo, género), e
- d) condições de trabalho justas e equitativas, caso o projeto envolva rotuladores remunerados.

Descreva o processo que irá utilizar para detetar potenciais questões éticas (por exemplo, um comité de análise institucional, etc.)

Normas e partilha de dados

- Localizável - fácil de encontrar numa plataforma pública e amplamente utilizada
- Acessível - formato aberto (CCBY 4.0)
- Interoperável - formato dos dados
- Reutilizável - metadados

- Sustentável - plano de manutenção
- Partilhado - comunidade de utilizadores envolvida e plano de partilha do conjunto de dados uma vez concluído



Orçamento e custos permitidos

- Deve ser apresentado no modelo Excel.
- Em caso de dúvida, explicar!
- Os custos permitidos incluem:
 - Formação/desenvolvimento de capacidades;
 - Recolha de dados; rotulagem; QA/QC ou verificação; e pós-processamento;
 - Modelo de base;
 - Licenciamento;
 - Alojamento do conjunto de dados;
 - Publicação dos resultados em acesso livre;
 - Metadados ou declaração de dados;
 - Esforços de crowdsourcing, tais como "label-a-thons".



Qualidade dos dados

- Processamento de dados - Envolve a análise de dados para obter informações, incluindo tipos de dados, distribuições, valores em falta, valores anómalos e potenciais inconsistências.
- Processo de limpeza de dados - Envolve o tratamento de valores em falta, a correção de erros, a remoção de valores anómalos, a resolução de inconsistências e a normalização de formatos de dados.
- Exatidão dos dados - verificar a exatidão dos dados comparando-os com fontes fidedignas ou

Considerações sobre a qualidade dos dados

1. Alojamento e acessibilidade

- Como e onde alojamos os dados é realmente importante.
- A aceitação de um conjunto de dados pode depender de como e onde o conjunto de dados é alojado

Onde está atualmente alojado o seu conjunto de dados? O que é que influenciou essa decisão?



Alojamento e acessibilidade

O que deve ter em conta ao escolher uma plataforma de alojamento?

- Domínio
- Caso de utilização
- Custo

Política do Fundo Lacuna:

Flexível para permitir a propriedade local dos dados e para responder aos casos de utilização previstos para os dados. Orientado pelos princípios do Fundo Lacuna

Alojamento e acessibilidade - Princípios da Lacuna

| Accessibility



| Equity



| Ethics



| Participatory Approach



| Quality



| Transformational Impact



2. Documentação do conjunto de dados

"Na indústria eletrónica, é normal que cada componente seja acompanhado de uma ficha de dados com características de funcionamento normalizadas, resultados de testes, utilização recomendada e outras informações.

Do mesmo modo, recomendamos que cada conjunto de dados seja acompanhado de uma ficha de dados que documente a sua criação, composição, utilizações previstas, manutenção e outras propriedades."

Benefícios da documentação do conjunto de dados

- Reprodutibilidade
- Transparência
- Usabilidade

A documentação do conjunto de dados mostrará o contexto, a estrutura e o conteúdo

Documentação do conjunto de dados - Directrizes Lacuna

- Ficha de dados
- Arquivo de metadados
- Identificador de objeto digital (DOI)



3. Licenciamento - Directrizes Lacuna



- Política de Propriedade Intelectual da Lacuna. Os conjuntos de dados e a respectiva propriedade intelectual desenvolvidos com fundos de subvenções serão:
- Ser propriedade da entidade subvencionada
- Ser objeto de uma licença aberta pela entidade subvencionada, a fim de maximizar o potencial de libertação responsável e de utilização a jusante dessa

Licenciamento.. o que está coberto?



- Metadados, fichas de dados ou outras informações.
- Orientações incluídas no conjunto de dados para guiar a sua utilização.
- Modelos de amostra ou outros instrumentos informativos divulgados com os dados.

Oportunidade de mentoria: Programa de Mentoria ACTS

Centro Africano de Estudos Tecnológicos
(ACTS)



O Programa de Mentoria ACTS

- **Missão:** A missão do ACTS é reforçar a capacidade e as políticas dos países e instituições africanos para aproveitar a ciência, a tecnologia e a inovação em prol do desenvolvimento sustentável. O programa de tutoria apoiará os candidatos no desenvolvimento de candidaturas de qualidade, bem como os orientará para compreenderem e seguirem as melhores práticas de implementação dos seus projectos.
- **No** âmbito do presente convite à apresentação de propostas do Fundo Lacuna, o ACTS oferecerá aos candidatos de todo o mundo uma oportunidade especial de participarem no programa de mentores, no qual os candidatos e as equipas serão acompanhados por um mentor

O Programa de Mentoria ACTS

- Candidatar-se no sítio **Web do ACTS**
[Programa de Mentoria ACTS](#)
- Seleccione participar em "**Apply for Mentorship**" OU "**Apply to be a Mentor**"
- Ambos os formulários têm campos semelhantes onde é necessário fornecer as suas informações pessoais, o fuso horário, a sua língua, a área de especialização, a área temática da proposta
- Aconselhamo-lo a preparar um esboço detalhado ou um projeto de proposta antes de se candidatar à tutoria, o que lhe permitirá conhecer a sua área temática
- Candidatar-se, **pelo menos, 4 semanas antes da data limite, até 26 de julho de 2024**
- Os mentores serão atribuídos por ordem **de chegada**.
- Todos os candidatos devem **ler e respeitar o código de ética e de conduta do programa de tutoria**.

O Código de Conduta ACTS

- Não causar danos
- Consentimento livre, prévio e informado
- Confidencialidade
- Segurança
- Sem coação
- Envolvimento de crianças (menores de 18 anos)
- Género
- Inclusão

O Programa de Mentoria

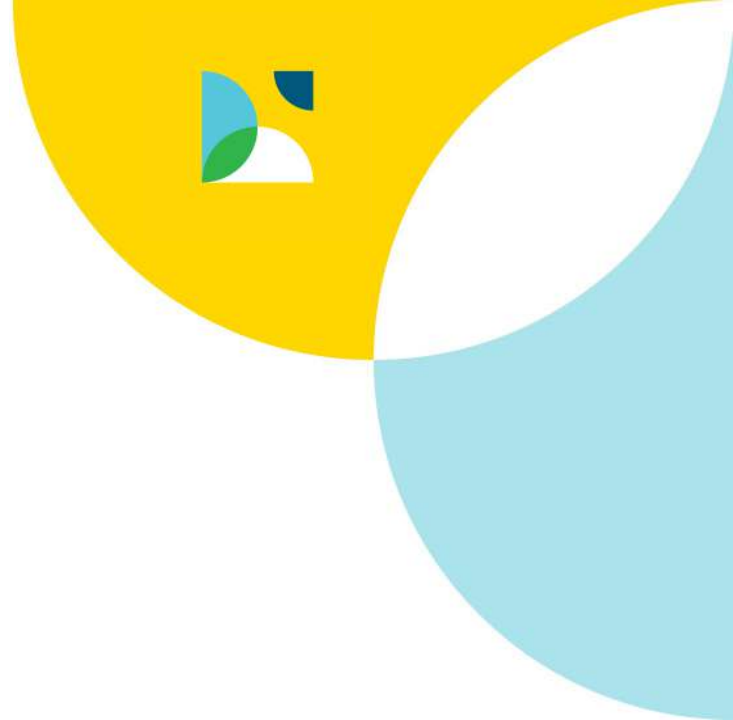
ACTS



Papéis de um Mentor

- Fornecer tutoria e orientar os candidatos para que compreendam o objetivo do convite
- Construir relações com os mentorados e compreender os seus pontos fracos e fortes
- Prestar assistência aos mentorandos através da revisão de propostas, consoante as necessidades
- Orientar os mentorandos sobre os papéis do grupo e os conflitos que possam surgir
- Participar nas sessões ou debates do grupo da proposta ou reunir-se separadamente com os investigadores principais.
- Discutir o trabalho relevante na área da proposta para garantir que a proposta é contextualizada.

Perguntas?



Próximos passos

- **Enviar perguntas adicionais para secretariat@lacunafund.org até 12 de julho**
- **Respostas publicadas** na página do Lacuna Fund Apply em **29 de julho**
- **Candidatar-se a mentor o mais rapidamente possível e até 26 de julho**
- **Propostas até 23 de agosto**

