

Regulamento Oficial

Datathon Regional: Avaliação Sociocultural de Grandes Modelos de Linguagem

O "Datathon Regional: Avaliação Sociocultural dos LLMs" é uma competição de um dia na qual estudantes da América Latina testarão grandes modelos de linguagem (LLMs) com perguntas sobre a cultura de seus próprios países. Isso com o objetivo de melhorar o conhecimento sociocultural desse tipo de modelo, com foco na América Latina.

As perguntas selecionadas farão parte da próxima versão do dataset LatamQA, que é usado pelo framework de mesmo nome para avaliar grandes modelos de linguagem.

Este evento ocorrerá no dia 3 de outubro de 2026, e é organizado pelo Inria Chile com o apoio do Ministério da Europa e Relações Exteriores da França, da rede de Embaixadas da França nos países da região e de universidades latino-americanas.

A realização do Datathon Regional (datathonregional.inria.cl) faz parte do IV Encontro de Ciência Aberta França-América Latina (<https://soframlat2026.sciencesconf.org/>), organizado em Montevideu, Uruguai, de 28 a 30 de outubro de 2026. Durante este evento e sua sessão inaugural, um vídeo recapitulativo do Datathon Regional será exibido aos presentes.

Artigo 1. Organização e Entidades-Chave

1.1. Organizador Geral: O Inria Chile atua como o Organizador Geral, assumindo a coordenação regional, a provisão da plataforma tecnológica de participação e da plataforma de comunicação, além da liderança de comunicação do evento.

1.2. Assessores Organizativos: As Embaixadas da França nos países participantes atuam como Assessores Organizativos em nível nacional, mobilizando redes acadêmicas e facilitando o apoio logístico, institucional e de comunicação.

1.3. Operadores Locais: As Universidades locais atuam como Operadores Locais, coordenando a definição das equipes, os locais de participação e o suporte logístico (presencial ou online).

1.4. Equipes e Representantes de Equipe: As equipes são formadas por estudantes das Universidades participantes do evento. Um membro designado por cada equipe participante atua como Representante de Equipe, assumindo a coordenação interna de seus membros, garantindo o cumprimento ético e metodológico do desafio, e servindo como canal oficial de comunicação com seu respectivo Operador Local.

Artigo 2. Participantes e Formação de Equipes

2.1. Elegibilidade: A participação é direcionada a estudantes de graduação, mestrado ou doutorado, representando uma universidade específica. Todos os membros devem ser cidadãos ou residentes do país latino-americano que representam.

2.2. Tamanho e Estrutura das Equipes: As equipes serão formadas estritamente por 4 a 5 membros. Deverão escolher um nome de Equipe e designar um Representante de Equipe, que atuará como canal oficial de comunicação com sua respectiva instituição.

2.3. Interdisciplinaridade: As equipes combinarão áreas do conhecimento, contando com pelo menos um membro com formação em ciências exatas ou engenharia, e pelo menos um membro com formação em humanidades ou ciências sociais. Podem participar estudantes de todas as áreas do conhecimento. Será valorizada especialmente a diversidade regional, de gênero e cultural.

Artigo 3. Registro das Equipes Participantes

3.1. Inscrição: A inscrição das equipes será realizada por meio de um formulário fornecido pelo Inria Chile, que estará disponível em <https://datathonregional.inria.cl>.

3.2. Conta no Hugging Face HUB: É obrigatório que pelo menos os Representantes de Equipe tenham uma conta no Hugging Face HUB (<https://huggingface.co>), cujo nome de usuário deverá ser inserido no formulário de inscrição das equipes. As contas do Hugging Face são necessárias para enviar as perguntas durante o evento.

3.3. O Organizador Geral reserva-se o direito de rejeitar qualquer solicitação de inscrição que esteja incompleta ou não cumpra com os requisitos. O Assessor Organizativo e o Operador Local também se reservam o direito de rejeitar qualquer pessoa como participante por razões de segurança ou em cumprimento a este Regulamento ou ao Código de Conduta.

3.4. Ao se inscrever em uma equipe, presume-se que todos os seus membros aceitam o presente Regulamento e o Código de Conduta do Datathon Regional.

A convocatória oficial abrirá em 14 de setembro de 2026. A confirmação e seleção final das equipes participantes será anunciada em 28 de setembro de 2026.

Artigo 4. Modalidade de Participação, Infraestrutura e Canais Oficiais

4.1. Modalidade: O evento será priorizado no formato presencial nas sedes universitárias coordenadas pelos Operadores Locais. Todos competirão simultaneamente no dia 3 de outubro de 2026, das 11:00 às 21:00 (hora UTC-3). Quando a presencialidade não for possível, a modalidade online será admitida em caráter excepcional (mediante consulta prévia ao Organizador Geral).

4.2. Infraestrutura e Espaços: Os Operadores Locais de cada sede serão responsáveis por habilitar um espaço físico que conte com as condições e os equipamentos necessários para que as equipes desenvolvam a competição e possam visualizar a transmissão de palestras (por vídeo ou streaming).

4.3. Plataformas Oficiais: A comunicação técnica do evento será realizada obrigatoriamente por meio da ferramenta Matrix (<https://matrix.org/>) fornecida pelo Organizador Geral. Os participantes deverão utilizar seus próprios dispositivos tecnológicos para se conectar.

4.4. Divulgação e Identidade Visual: Toda comunicação pública ou divulgação oficial realizada no âmbito do evento deverá seguir estritamente o kit de comunicação fornecido pelo Organizador Geral.

Artigo 5. Propriedade Intelectual e Licenças

5.1. Repositório Aberto: Todas as contribuições criadas durante o evento serão integradas ao dataset do framework LatamQA. Os participantes garantem que suas entregas são originais e não violam direitos autorais de terceiros. Os participantes se responsabilizam por não usar ferramentas de inteligência artificial durante o evento.

5.2. Licença: Os datasets produzidos estarão sujeitos à licença Creative Commons CC-BY-4.0 (<https://creativecommons.org/licenses/by/4.0/>).

Artigo 6. Procedimento e Dinâmica do Desafio

6.1. Formato das Entregas: Cada equipe deverá redigir e inserir na plataforma do evento um conjunto de perguntas de múltipla escolha (QCM) compostas por uma (1) resposta correta e três (3) distratores plausíveis. As perguntas devem ser formuladas no idioma local (espanhol ou português) e incluir obrigatoriamente uma versão da mesma em inglês.

6.2. Metadados e Wikipedia: Toda pergunta deve incluir metadados obrigatórios (país, região, dimensão cultural e nota contextual) e estar vinculada a uma ou mais páginas da Wikipedia. Se o conteúdo não existir na Wikipedia (ou não tiver um ID associado), deverá ser incluída uma justificativa de origem cultural.

6.3. Linguagem e Conteúdo: As perguntas não devem ter conteúdo de natureza obscena, desrespeitosa, irrelevante, impertinente ou que não esteja alinhado com o Código de Conduta dos colaboradores do Datathon Regional. O descumprimento pode ser penalizado com a exclusão da equipe.

Artigo 7. Avaliação e Seleção de Vencedores

Comitê de Avaliação: A revisão inicial das perguntas será realizada por meio de um sistema misto que combina análise de similaridade de texto e revisão humana. Posteriormente, a avaliação ficará a cargo de um painel de modelos de linguagem (LLMs). O processo será executado por meio de critérios operacionais uniformes. A pontuação de cada equipe dependerá da capacidade de suas perguntas de "enganar" com sucesso tais modelos dentro desse marco normativo. Um ranking com as posições das equipes será exibido e atualizado regularmente durante o evento.

Artigo 8. Prêmios e Reconhecimentos

8.1. Prêmio Regional: A equipe que obtiver a maior pontuação em nível regional receberá um prêmio em dinheiro de 2.000 euros, que será dividido igualmente entre todos os seus membros. Além disso, seus membros serão reconhecidos como vencedores do evento em futuras publicações relacionadas ao dataset.

8.2. Prêmio Local: Os Assessores Organizativos e/ou os Operadores Locais poderão oferecer prêmios e reconhecimentos em nível local.

Artigo 9. Isenção de Responsabilidade

9.1. Segurança e Bens: O Organizador Geral e os Assessores Organizativos não se responsabilizam por perdas, roubos ou danos a equipamentos informáticos pessoais ocorridos durante o desenvolvimento das jornadas presenciais. Da mesma forma, os participantes assumem o risco individual em caso de falhas técnicas ou de segurança cibernética em suas redes locais.

Artigo 10. Proteção de Dados e Confidencialidade

10.1. Marco Legal: O Organizador Geral (Inria Chile) e os Assessores Organizativos (Embaixadas) deverão coletar e processar dados pessoais para o bom andamento do Datathon Regional. Em particular, informa-se aos participantes e aos Representantes de Equipe que, em conformidade com o Regulamento Geral de Proteção de Dados da União Europeia (GDPR UE 2016/679) e a Lei N° 19.628 da República do Chile, a informação fornecida é necessária para registrar e concretizar sua participação, e os destinatários são o Organizador Geral e os Assessores Organizativos na qualidade de responsáveis pelo tratamento de dados.

Para formalizar a inscrição, cada participante e Representante de Equipe deverá assinar de maneira explícita o Documento de Consentimento e Confidencialidade para o Tratamento de

Dados Pessoais, concedendo sua autorização formal para a coleta e o tratamento de seus dados nas condições aqui descritas.

Qualquer tratamento de dados pessoais no contexto do Datathon Regional em um país cujo nível de proteção não seja considerado adequado deverá ser realizado estritamente dentro do mesmo marco aplicável aos eventos realizados na União Europeia. Nesses casos, o Assessor Organizativo é responsável por garantir o cumprimento da legislação de proteção de dados pessoais aplicável na União Europeia.

Quando os participantes e Representantes de Equipe fornecem informações, o fazem de forma consciente. No momento em que as informações são solicitadas, indica-se se elas são obrigatórias ou opcionais.

As finalidades dessas operações de tratamento de dados são: coletar informações dos participantes e Representantes de Equipe, gerenciar sua alocação em equipes e comunicar-se com eles e entre eles.

As bases jurídicas para o tratamento de dados são: o Documento de Consentimento e Confidencialidade; e o interesse legítimo em utilizar os dados para fins de comunicação.

10.2. Segurança e Cibersegurança dos Dados: O Organizador Geral e os Assessores Organizativos implementam medidas de segurança técnicas, administrativas e organizativas adequadas e razoáveis para proteger a confidencialidade, integridade e disponibilidade dos dados pessoais contra acessos não autorizados, perda ou alteração. No entanto, os participantes reconhecem que nenhum sistema de transmissão ou armazenamento de informações online é 100% invulnerável. Consequentemente, o Organizador Geral não será responsável por incidentes de cibersegurança ou ataques informáticos massivos executados por terceiros que ultrapassem os padrões razoáveis de proteção aplicados, sem prejuízo do compromisso de gerenciar o incidente conforme a normativa vigente e de notificar os afetados de acordo com a lei.

10.3. Direitos ARCO (Acesso, Retificação, Cancelamento, Oposição) e Eliminação: Os dados dos participantes e Representantes de Equipe serão mantidos pelo tempo necessário para realizar as operações para as quais foram coletados e em conformidade com a legislação vigente, ou seja, pelo tempo necessário para cumprir o prazo de conservação estabelecido, salvo se os participantes (ou Representantes de Equipe) solicitarem de forma explícita e justificada a supressão dos dados.

Os destinatários previstos dos dados são o Organizador Geral, os Assessores Organizativos e seus funcionários ou substitutos, e os prestadores de serviços técnicos de cada instituição, que atuam por conta própria no âmbito da prestação de serviços (hospedagem, manutenção, supervisão e moderação). Salvo se as pessoas interessadas se opuserem expressamente posteriormente, o Organizador Geral e os Assessores Organizativos estarão autorizados a divulgar os nomes dos integrantes da equipe vencedora, os resultados, se for o caso, bem

como sua finalidade e descrição, por qualquer meio, para fins informativos.

Salvo oposição expressa, o Organizador Geral e os Assessores Organizativos reservam-se o direito de utilizar as informações fornecidas pelos participantes e Representantes de Equipe para: melhorar a organização do Datathon Regional; transmitir informações (somente quando os participantes e Representantes de Equipe tiverem dado seu consentimento expresso ao serem convidados a isso); e para fins de pesquisa e marketing.

Os participantes e Representantes de Equipe têm o direito de solicitar ao Operador Local, ao Assessor Organizativo e ao Organizador Geral (como responsável pelo tratamento de dados) a retificação, a complementação, a atualização, o bloqueio ou a supressão dos dados pessoais que lhes digam respeito e que sejam inexatos, incompletos, ambíguos, desatualizados ou cuja coleta, uso, divulgação ou conservação seja proibida.

Os participantes e Representantes de Equipe também têm o direito de se opor e de restringir, por motivos legítimos, o tratamento de seus dados pessoais.

Todos esses direitos podem ser exercidos pelos participantes e Representantes de Equipe enviando um e-mail ao Encarregado de Proteção de Dados do Organizador Geral no seguinte endereço: dpo@inria.cl, anexando um documento de identidade assinado para prevenir fraude de identidade.

O Organizador Geral fará todo o possível para atualizar, retificar ou eliminar os dados pessoais do participante ou Representante de Equipe o mais rápido possível e informará os Assessores Organizativos.

Informa-se aos participantes e Representantes de Equipe que, após seu falecimento, seus herdeiros poderão contestar o tratamento continuado de seus dados ou solicitar sua atualização.

10.3. Confidencialidade: As partes manterão a confidencialidade sobre as informações sensíveis do evento por um período de cinco (5) anos.

Artigo 11. Direitos de Imagem e Divulgação

11.1. Autorização de Imagem e Voz: Ao se inscrever no Datathon Regional e aceitar este Regulamento, os Participantes concedem sua autorização ao Organizador Geral, bem como aos Assessores Organizativos, aos Operadores Locais, e/ou à pessoa ou entidade autorizada por estes para captar e gravar sua imagem e voz durante o evento para o qual se inscreveram voluntariamente, para comunicações internas e/ou externas. Essas imagens e vídeos poderão ser editados, divulgados, reproduzidos, modificados e utilizados de forma gratuita e em todo o mundo pelo Organizador Geral e/ou pela pessoa ou entidade por ele autorizada, e poderão ser comunicados ao público e/ou incorporados para fins promocionais e de divulgação, incluindo a difusão comercial, educativa e científica.

Artigo 12. Calendário Oficial e Data de Contingência

12.1. Data de início das inscrições: As instituições e suas equipes que participarão do Datathon Regional poderão se inscrever a partir do dia 14 de setembro de 2026.

12.2. Data de encerramento das inscrições: Só serão aceitas inscrições para o Datathon Regional até o dia 27 de setembro de 2026.

12.3. Data do evento: O Datathon Regional será realizado no sábado, 3 de outubro de 2026, de forma simultânea em todos os países participantes, das 11:00 às 21:00 (hora UTC -3).

12.4. Modificações de datas: O Organizador Geral, o Assessor Organizativo e o Operador Local não podem ser responsabilizados se, por motivo de força maior ou fatores fora de seu controle, o Datathon Regional for alterado, adiado ou cancelado. O Organizador Geral reserva-se o direito, em todos os casos, de prorrogar o período de participação e de adiar todas as datas indicadas.

Artigo 13. Critérios de Exclusão e Condições Adicionais

13.1. Desqualificação Técnica: Serão eliminadas da pontuação final aquelas perguntas que dependam de dados efêmeros não verificáveis (rumores, opiniões), que constituam trivialidades sem valor sociocultural, que incorram em erros factuais, que promovam estereótipos prejudiciais ou que contenham conteúdo discriminatório.

13.2. Alinhamento com os objetivos: Não é permitido o uso de inteligência artificial (IA) para gerar perguntas para o dataset.

13.3. Conduta: O Organizador Geral reserva-se o direito de expulsar qualquer equipe ou participante que atente contra o Código de Conduta dos Colaboradores, a ética ou a convivência pacífica durante o evento. Esta sanção poderá ser executada mediante notificação prévia e formal por parte do Assessor Organizativo ou do Operador Local.