



Manual de Orientações Técnicas - Hackathon Online 2020

Prezado participante,

esse é o Manual de Orientações Técnicas do Hackathon Serpro. Nele você vai encontrar as informações necessárias ao correto desenvolvimento da solução, além de orientações sobre o funcionamento e a dinâmica do evento.

Leia com atenção e programe-se para dispor de todos os recursos necessários para a execução do desafio. Durante o evento, teremos uma vasta programação com palestras, sessões de mentoria e momentos dedicados, exclusivamente, ao desenvolvimento da solução.

Vamos juntos nessa jornada para ajudar a reduzir os impactos da pandemia de Covid-19!

Live de abertura

Como o desafio agora é online, toda nossa interação será realizada por meio de ferramentas de comunicação multimídia como o Whatsapp e o Google Meet. No primeiro dia do evento, dia 15 de junho, às 10h, pelo menos um membro da equipe precisará estar presente na live de abertura e realizar o check-in no evento informando seu nome e o nome da equipe no chat do Google Meet.

Esse evento exige a participação de pelo menos um componente da equipe e é essencial para o entendimento de toda a dinâmica do Hackathon. Neste momento, falaremos sobre a dinâmica da maratona, sobre como tirar o máximo proveito das mentorias e sobre as ferramentas de comunicação, de desenvolvimento e bancos de dados disponíveis.

Nesta live também entregaremos as chaves de acesso ao Ambiente de Inovação do Serpro, com ferramentas de desenvolvimento, APIs Serpro e bancos de dados.

Palestras

Durante o evento teremos mais três palestras com profissionais renomados, trazendo temas emergentes nas áreas de tecnologia e negócios, considerando

nosso cenário atual.

Nessas palestras, você poderá tirar valiosos insights para o desenvolvimento da sua solução. Portanto, prestigie este momento! Os links de acesso serão passados oportunamente pela organização.

Papel dos Mentores

Os mentores são especialistas que estão no evento para orientar, dar suporte e passar conhecimentos para os participantes. É um time formado de pessoas com habilidades tanto na área de tecnologia da informação, como marketing, empreendedorismo e negócios. Apesar de poderem contribuir bastante com o sucesso do projeto da sua equipe, não fique esperando instruções dos mentores o tempo todo. Os mentores oferecem dicas durante o processo, mas a validação e a construção da solução ficam por conta dos integrantes da sua equipe. Fechado?!

Comunicação com a organização e mentores

Toda comunicação durante o HackSerpro será realizada de forma remota. Os líderes das equipes participarão de um grupo de Whatsapp com a comissão organizadora e os mentores.

A mentoria será realizada pelo Meet. Será disponibilizado um [ambiente online](#) com uma sala central para interação com os mentores onde as equipes terão sessões de mentoria em grupo e poderão solicitar mentorias individuais. Neste ambiente, cada equipe terá também uma sala específica para receber as mentorias individuais.

Cada mentoria tem duração de 15 minutos e os mentores poderão fazer visitas proativas às suas salas.

Não esqueça de manter a sala do Google Meet sempre aberta.

Pitch - Demonstração das soluções

O pitch é o momento de coroar o bom trabalho executado ao longo da competição. São 6 minutos para você “vender” da melhor maneira a sua solução e convencer os jurados que a sua equipe merece ser a vencedora. Coloque muita energia nesse momento e dedique um tempo na construção dessa apresentação, pois ela é decisiva para o resultado final do seu time.

E lembre-se, a demonstração da solução ocorrerá de forma virtual, por meio de publicação de vídeo no YouTube, até 12h (meio dia) do dia 22 de junho. Recomendamos a utilização de recursos de edição para adequação do tempo.

De acordo com o edital, essa apresentação deverá conter os seguintes elementos: 1- Público-alvo; 2 – Problema; 3 – Modelo de Negócio; e 4 – Diferencial Competitivo.

Crerios da Banca Julgadora

A análise da banca julgadora será feita levando em consideração os elementos descritos no art. 35 do Regulamento do Hackathon, disponível no site. Serão avaliados a inovação do projeto, sua usabilidade e design, completude funcional,

tecnologia empregada e relevância do problema resolvido.

Além desses itens, fiquem atentos aos critérios descritos no parágrafo único do art. 3º, eles são essenciais para o sucesso do projeto da sua equipe.

Pré-seleção das soluções

O comitê organizador do evento, formado por profissionais, mentores e especialistas multidisciplinares realizará a pré-seleção das equipes finalistas a serem avaliadas pela banca julgadora.

Serão submetidas à banca julgadora as 10 (dez) soluções melhor avaliadas pelo comitê organizador, que utilizará os mesmos critérios da banca para pré-classificação das equipes.

As equipes não selecionadas para a fase de avaliação final não terão qualquer classificação na competição.

ORIENTAÇÕES TÉCNICAS

Desenvolvimento da Solução

Antes da Dinâmica:

Serão fornecidas as seguintes ferramentas:

Jupyter: ambiente computacional web para compilação e execução de algoritmos em linguagem python;

RStudio: ambiente de desenvolvimento integrado para R, uma linguagem de programação para gráficos e cálculos estatísticos;

GitLab: ambiente para controle de versionamento de códigos. Todos os códigos e documentos gerados durante o evento **obrigatoriamente deverá ser disponibilizado no ambiente;**

- Cada equipe será responsável por incluir o arquivo e a URL do vídeo do Pitch (Youtube) no ambiente do Git;

Bases de Dados em PostgreSQL, CSV e TXT: serão disponibilizadas bases de dados de acesso anonimizada e não anotada (sem indicação de fraude) somente leitura;

- Cada equipe terá um banco de dados limitado a 30GB de armazenamento;
- Sua equipe também poderá realizar a extensão da base de dados fornecida (criação de novos registros e colunas) localmente;
- Também será fornecido um dicionário de dados descrevendo os campos;

- A utilização da base de dados fornecida será permitida somente durante o evento;

API Serpro: Cada equipe receberá uma chave de acesso para exploração de APIs de governo em suas soluções:

API CPF, API CNPJ, Datavalid, BioValid (Biodata), IntegraComex.

Sterna: serviço de georreferenciamento de governo que permite, entre outros recursos, a obtenção de dados georreferenciados via CEP.

A equipe não está restrita ao uso das ferramentas disponibilizadas, desde que todo o código seja disponibilizado no ambiente Git do evento;

A autenticação de usuário e senha para cada equipe será disponibilizada no dia do evento;

É proibida a construção, ainda que parcial, de qualquer parte da solução antes do início do evento.

Durante a Dinâmica:

Sua equipe poderá utilizar qualquer linguagem de programação (Python, R, Java, etc..)

Também é livre a utilização de qualquer framework ou biblioteca (Tensorflow, Keras, Pandas, etc..) além do ambiente já disponibilizado.

Será permitido utilizar qualquer banco de dados SQL ou NoSQL (Postgres, MySQL, Elasticsearch, MongoDB, etc...)

Será permitido utilizar qualquer arquitetura computacional (CPU, GPU, etc...) on premise;

Será permitido utilizar APIs ou ferramentas externas (GCP, AWS, Azure, SPSS, SAS, etc..)

Será permitido utilizar ambiente em nuvem para criação da solução (GCP, AWS, Azure, etc..).

Toda e qualquer solução, inclusive as geradas com as ferramentas disponibilizadas, deverá ter seu código disponibilizado no git do evento. Todos os códigos e procedimentos disponibilizado no git deverão ser suficientes para a criação e replicação da solução.

Algumas bases que estarão disponíveis:

CBO - Classificação Brasileira de Ocupações (CBO): uma norma de classificação enumerativa e descritiva de atividades econômicas e profissionais determinada pela Comissão Nacional de Classificação para o uso de órgãos governamentais, servindo como base para a estruturação de carreiras e ocupação de vagas no setor

público e no setor privado.

CNPJ - Cadastro Nacional da Pessoa Jurídica (CNPJ): um número único que identifica uma pessoa jurídica e outros tipos de arranjo jurídico sem personalidade jurídica (como condomínios, órgãos públicos, fundos) junto à Receita Federal do Brasil, compreendendo as informações cadastrais das entidades de interesse das administrações tributárias da União, dos Estados, do Distrito Federal e dos Municípios.

COMEX - Estatísticas de Comércio Exterior (Comex Stat): dados do comércio exterior brasileiro disponibilizados pelo Ministério da Indústria, Comércio Exterior e Serviços - MDIC, utilizados para a construção da balança comercial brasileira (exportações e importações de produtos), sendo detalhados por NCM/SH4 do produto e/ou por município do exportador/importador.

COMPRAS - Compras Governamentais: dados abertos do Sistema Integrado de Administração e Serviços Gerais (SIASG), onde se operacionalizam as compras do Governo Federal, contemplando o catálogo de materiais e serviços (CATMAT/CATSER), o cadastramento e divulgação da licitação (SIDECA/Divulgação), as intenções de registros de preços (IRP), o cadastramento dos fornecedores (SICAF), a realização das licitações (Compras Governamentais/Sessão Pública/RDC), o resultado das licitações (SISPP/SISRP), os empenhos de pagamentos (SISME) e o registro e gestão dos contratos (SICON).

ENEM - Exame Nacional do Ensino Médio (Enem): uma prova realizada pelo INEP/MEC utilizada para avaliar a qualidade do ensino médio no país, cujo resultado serve para acesso ao ensino superior em universidades públicas brasileiras.

IPEA - Instituto de Pesquisa Econômica Aplicada (Ipea): conteúdos de pesquisas que fornecem suporte técnico e institucional às ações do governo para a formulação e reformulação de políticas públicas e programas de desenvolvimento (ex: Índice de Desenvolvimento Humano Municipal - IDHM).

NCM - Nomenclatura Comum do Mercosul (NCM): uma convenção de categorização de mercadorias adotada desde 1995 pelo Uruguai, Paraguai, Brasil e Argentina e que toma por base o Sistema Harmonizado (SH), criada a fim de melhorar e facilitar o crescimento do comércio internacional, facilitando também a criação e comparação de estatísticas.

MCIDADANIA - Bases de alguns programas sociais do governo federal.

SANÇÕES - Dados referentes aos cadastros de sanções aplicadas a pessoas físicas, jurídicas e servidores públicos federais.

BASE DE DADOS DE SAÚDE - Serão sugeridos o uso de datasets de bases de dados abertos relacionados à saúde.

ATENÇÃO: Apesar de as soluções poderem ser desenvolvidas utilizando um software, plataforma ou aplicação proprietária, obrigatório que ela seja baseada em software livre e licenciada nos termos da licença livre GNU LGPL v3.0, que compreende, dentre outras coisas, a exigência de disponibilização de código fonte aberto, a liberdade de uso e a possibilidade de distribuição e modificação do software.

Sua equipe ainda tem alguma dúvida? Mande um e-mail para hackathon@serpro.gov.br que algum dos nossos mentores responde pra você.

Até breve!

Equipe Hackathon Serpro