

TERMO DE REFERÊNCIA – CONSULTORIA EM SOLUÇÕES PARA SISTEMAS, APRENDIZADO DE MÁQUINA E DADOS

Título do Posto: Consultor sênior em soluções para Sistemas, Aprendizado de Máquina e Dados

Tipo de Contrato: Consultoria Individual (CI)

Local de trabalho: Remoto

1. Escritório Contratante	FUNDO DE POPULAÇÃO DAS NAÇÕES UNIDAS (UNFPA)
2. Contexto e Objeto(s) da Consultoria	<p>O IBGE, que completará 90 anos em 2026, faz grande parte do seu trabalho a partir do uso de novas tecnologias. No ano de 2010, o Censo foi coletado em dispositivo móvel, sem uso de papel. O IBGE também foi o primeiro órgão do Brasil a receber um computador. Nesse sentido, a fronteira tecnológica do país sempre esteve próxima ao IBGE.</p> <p>Em plena Era Digital e no ápice da revolução informática, novos desafios são colocados para a Instituição seguir se diferenciando e entregando dados e informações com qualidade, segurança, garantindo sigilo e as boas práticas geocientíficas e estatísticas. No IBGE, há ainda o desafio de modernização tanto no campo das estatísticas quanto no campo das geociências. As chamadas big data servem para as imagens de satélite e também para os dados estatísticos.</p> <p>Frente a um volume de dados gigantesco, o desafio é organizar e sistematizar o acúmulo de informações. Nesse sentido, técnicas de pareamento entre as pesquisas e entre pesquisas e registros administrativos tornam-se cada vez mais valiosas e necessárias. Não obstante à tecnologia existente na Era Digital, há a necessidade de qualificação de mão de obra e de aprendizado desses novos recursos, como é o caso do uso da máquina de aprendizado para dados.</p> <p>Nesse sentido, há muitos anos a Instituição produz diversas informações que, na parte de análise e crítica, por falta de pessoas e de atualização tecnológica, acabam não sendo trabalhadas e, portanto, o potencial de informação que elas podem proporcionar à população e aos gestores não é alcançado.</p> <p>Pelo âmbito das pesquisas estatísticas, o pareamento se coloca como uma estratégia alvissareira, permitindo correções e críticas entre os Censos e entre o Censo e a PNADC. Ademais, pode servir de conferência entre pesquisas do IBGE e registros administrativos. Permite, por exemplo, a confecção de dados longitudinais.</p>

Assim, frente à crescente recusa da população, com elevada resistência de se responder às pesquisas domiciliares por outras vias, telefone e internet, as estratégias do pareamento tornam-se imprescindíveis, permitindo melhorar o processo de crítica e entendimento dos dados.

Como Instituição nacional de estatísticas oficiais (INE) e coordenadora do Sistema Estatístico Nacional, o IBGE tem como meta manter suas pesquisas modernas, buscando sempre melhoria da qualidade de seus dados, bom desenvolvimento de questionários, soluções para os desafios de cobrança, bem como liberação rápida de dados e extensa disseminação dos resultados de suas pesquisas. Para isso, deve seguir as recomendações das melhores práticas estatísticas e se manter atualizada sobre a experiência e práticas estatísticas de outros países.

A construção nacional de um cadastro de endereços para fins estatísticos é parte essencial de qualquer pesquisa estatística domiciliar. A atualização de endereços, em um país tão desigual, grande e dinâmico como o Brasil, é fundamental. No Censo Demográfico de 2022, o setor de cadastro de endereços coletou 111 milhões de endereços com as suas respectivas latitudes e longitudes. Nesse contexto, o IBGE divulgou a informação geral de que existiam cerca de 600 mil estabelecimentos religiosos, 280 mil estabelecimentos de saúde e 300 mil estabelecimentos de ensino. Além desses, foram identificados mais de 11 milhões de estabelecimentos categorizados como de outros tipos, que incluem academia, bar, salão de beleza etc. Apesar da valiosa informação disponibilizada pelo Instituto, para usos mais práticos, esses estabelecimentos devem ser refinados em subtipos. Isto é, devem ser subclassificados.

Do ponto de vista metodológico, é fundamental planejar a construção de um processo claro que permita a incorporação do uso de máquinas de aprendizado não supervisionada para a automatização das classificações de dados, eliminando ou mitigando as lacunas deixadas pela falta de mão de obra para responder à demanda social pelos dados referidos.

Nas últimas décadas, ocorreram alterações e aprimoramentos em variados aspectos técnicos, tecnológicos e operacionais das pesquisas, que lhes conferiram um salto de qualidade. No contexto desta consultoria, serão trabalhadas bases de dados em conjunto com dados do Censo Demográfico de 2022 para aprimorar e modernizar as operações estatísticas.

Para preencher essas lacunas, de pareamento e de subclassificação dos tipos de estabelecimentos são assaz necessários. O consultor selecionado será responsável por prover apoio à Diretoria de Pesquisas – e à Coordenação de métodos e qualidade e ao Cadastro Nacional de Endereços, atuando na prospecção de soluções tecnológicas, modelagem de dados e proposição de ferramentas computacionais no âmbito da Modernização da

	<p>Produção Estatística e Geocientífica. Em particular, atuará diretamente no escopo de soluções à incorporação do uso de base de dados internas e externas em conjuntos com práticas de máquina de aprendizado para a disseminação de operações geocientíficas e estatísticas oficiais. As atividades e resultados esperados do consultor estão descritos abaixo e serão desenvolvidos observando, principalmente, a aderência aos processos no âmbito do Modelo Genérico do Processo de Produção Estatística - GSBPM.</p>
<p>3. Escopo do Trabalho e Produtos</p>	<p>ATIVIDADES:</p> <ol style="list-style-type: none"> 1. Participar de reuniões preparatórias com a equipe da supervisão da consultoria; 2. Elaborar relatórios técnicos sobre a prospecção de soluções tecnológicas e modelagem de dados; 3. Projetar e documentar a arquitetura das soluções tecnológicas; 4. Detalhar as tecnologias a serem empregadas; 5. Propor metodologias de teste e avaliação para os modelos de aprendizado de máquina, levando em consideração os dados disponíveis; 6. Propor novas formas de levantar e validar dados para o treinamento dos modelos de aprendizado de máquina; 7. Identificar e documentar às necessidades de aperfeiçoamento das operações de campo, coleta e comunicação de dados via Dispositivo Móvel de Coleta (DMC) com foco em pareamentos futuros; 8. Parear dados de diferentes bancos de dados internos e externos; 9. Elaborar documentação técnica sobre os produtos desenvolvidos; 10. Registrar os passos - mapeamento de processos - para que a tecnologia usada possa ser mantida na Instituição, trazendo conhecimento para outros servidores do IBGE. <p>RESULTADOS ESPERADOS:</p> <ol style="list-style-type: none"> 1. Melhoria das estratégias e dados para treino dos modelos de aprendizado de máquina; 2. Aprimoramento da qualidade dos resultados obtidos, melhorando a performance dos modelos nas métricas selecionadas; 3. Bases de dados integradas e pareadas;

	<p>4. Diversificação das formas de utilização de bases de dados geocientíficas nas informações estatísticas;</p> <p>5. Incorporação de novos atributos geoespaciais a temas já pesquisados nos censos demográficos.</p> <p>PRODUTOS:</p> <p>Produto 1: Relatório contendo proposta de metodologia e classificação em subtipos dos estabelecimentos, identificados como do tipo Religioso, a partir de agrupamentos definidos pela equipe técnica do IBGE, através do uso de Machine Learning não supervisionada, permitindo uma informação mais completa a respeito da distribuição territorial dos estabelecimentos religiosos de diferentes matrizes.</p> <p>Produto 2: Relatório contendo proposta de metodologia e classificação em subtipos dos estabelecimentos identificados como do tipo Educação, a partir de agrupamentos definidos pela equipe técnica do IBGE, através do uso de Machine Learning não supervisionada, permitindo uma informação mais completa a respeito da distribuição territorial dos estabelecimentos de ensino pelos seus diferentes níveis de instrução.</p> <p>Produto 3: Relatório contendo proposta de metodologia e classificação em subtipos dos estabelecimentos identificados como do tipo Saúde, a partir de agrupamentos definidos pela equipe técnica do IBGE, através do uso de Machine Learning supervisionada, permitindo uma informação mais completa a respeito da distribuição territorial dos estabelecimentos de saúde segundo áreas de tratamento para a população.</p> <p>Produto 4: Relatório contendo a proposta de pareamento dos dados do Censo 2022 com o CadÚnico, identificando uma chave única do morador capaz de posteriormente identificar dados que foram suprimidos do Censo 2022.</p>
<p>4.Duração e Horário do Trabalho</p>	<p>Duração: 10 meses contados a partir da data de contratação.</p> <p>Horário de trabalho: jornada de trabalho a ser acordada junto ao gerente da área na qual o consultor atuará. O trabalho poderá ser desenvolvido remotamente.</p>

<p>5. Local onde os serviços devem ser entregues:</p>	<p>Os produtos deverão ser entregues à Presidência do IBGE para aprovação, depois centralizados no Diretor ou Coordenador Nacional do Projeto no IBGE que, posteriormente, enviará ao Escritório do UNFPA Brasil, por e-mail, para aprovação final e pagamento.</p>
<p>6. Datas de entrega e como o trabalho será entregue (ex. arquivo eletrônico, meio físico, etc.)</p>	<p>Produto 1: 55 dias após a assinatura do contrato – R\$ 23.000,00 (Vinte e três mil reais)</p> <p>Produto 2: 110 dias após a assinatura do contrato – R\$ 23.000,00 (Vinte e três mil reais)</p> <p>Produto 3: 200 dias após a assinatura do contrato – R\$ 34.500,00 (Trinta e quatro mil e quinhentos reais)</p> <p>Produto 4: 285 dias após a assinatura do contrato – R\$ 34.500,00 (Trinta e quatro mil e quinhentos reais)</p> <p>Valor total da consultoria: R\$ 115.000 (Cento e quinze mil reais)</p>
<p>7. Monitoramento e controle de andamento, inclusive exigências de relatórios, formato, periodicidade e prazo final.</p>	<p>A supervisão do trabalho será realizada por meio de análise do produto, reuniões periódicas com o consultor e acompanhamento do andamento do trabalho, de modo a possibilitar eventuais ajustes necessários.</p>
<p>8. Disposições de Supervisão:</p>	<p>O/a consultor/a desenvolverá as atividades e produtos sob a supervisão de um Oficial de Programa no UNFPA Brasil e de um assessor da presidência do IBGE.</p>
<p>9. Viagem prevista:</p>	<p>Não há viagem prevista</p>
<p>10. Capacitação, qualificações e competências necessárias, inclusive idiomas:</p>	<p>ETAPA – ANÁLISE CURRICULAR</p> <p>Requisitos obrigatórios:</p> <ul style="list-style-type: none"> ● Pós-graduação lato ou stricto sensu em Computação/Ciência de Dados, Processamento Digital de Imagens ou áreas afins; ● Experiência em escrita científica/report; ● 3 anos de experiência com pareamento/emparelhamento de dados; ● 2 anos de experiência em programação para modelos de aprendizado de máquina. <p>Requisitos desejáveis</p>

- Mestrado com ênfase em Computação, Ciência de Dados ou áreas afins (5 pontos);
- Doutorado com ênfase em Computação, Ciência de Dados ou áreas afins (10 pontos);
- Experiência de trabalho com algoritmos de aprendizado de máquina ou profundo (1 ponto por ano de experiência, máximo de 10 pontos);
- Experiência de programação em Python (2 pontos por ano de experiência, máximo de 10 pontos);
- Experiência de programação em R (2 pontos por ano de experiência, máximo de 10 pontos).

Habilidades Corporativas

- Integridade, ética e valores
 - o Agir de acordo com os valores do UNFPA e da ONU, regras administrativas, código de conduta e princípios éticos.
 - o Exercer julgamento crítico ao lidar com dados operacionais com foco no melhor cumprimento do mandato do UNFPA e garantir a confidencialidade das informações.
 - o Gerenciamento de conflitos / negociação e resolução de desacordos.
 - o Construção de apoio e perspicácia política. o Criatividade e inovação.
 - o Trabalho em equipe.
 - o Comunicação eficaz.
 - o Compartilhamento de conhecimento.
 - o Tomada de decisão justa e transparente.
- Orientação para cliente / parceiro
 - o Contribuir para a obtenção de resultados positivos para clientes e parceiros, antecipando necessidades e preocupações e respondendo a elas com eficiência.
- Sensibilidade à diversidade cultural
 - o Demonstrar um comportamento inclusivo com colegas e partes interessadas, desenvolvendo com sucesso relações interculturais.
 - o Ser adaptável e sensível às diferenças políticas, religiosas e culturais.
 - o Promover ativamente a equidade e a diversidade de gênero em todas as atividades.

	<ul style="list-style-type: none"> ● <u>Accountability</u> <ul style="list-style-type: none"> o Seja responsável por lidar com informações confidenciais em apoio ao trabalho do UNFPA no Brasil. o Garantir que as informações sejam coletadas, registradas e usadas corretamente para minimizar erros e fortalecer a implementação. o Tomada de decisão apropriada e transparente.
<p>11. Insumos / serviços a serem fornecidos pelo UNFPA ou parceiro de implantação (ex. serviços de suporte, sala de escritório, equipamentos), se aplicável:</p>	<p>O(a) consultor(a) deverá dispor dos meios e equipamentos necessários para a realização da consultoria.</p> <p>Para as atividades de pareamento, o IBGE deverá oferecer um computador capaz de suportar a manipulação das bases de dados</p>
<p>12. Outros dados pertinentes ou condições</p>	<p>O Edital e seus termos de referência estão disponíveis no site do IBGE (https://www.ibge.gov.br/acesso-informacao/institucional/trabalhe-conosco.html) e do UNFPA Brasil (https://brazil.unfpa.org/pt-br/vacancies). Para realizar a inscrição para este termo de referência, é necessário preencher o Formulário Google (https://forms.gle/x9BXqNnoEzpXzPeg7) e anexar nele o Formulário P11, conforme descrito no Edital.</p> <p>Não serão aceitas submissões feitas após o prazo final estipulado. Apenas as/os candidatas/os aprovadas/os serão contactadas/os.</p> <p>O UNFPA é uma agência comprometida com a diversidade em termos de gênero, raça, orientação sexual, nacionalidade e cultura. Pessoas de grupos minoritários, são encorajadas a participar do processo seletivo. Todas as candidaturas serão analisadas com total confidencialidade.</p>

ANEXO I

SIGILO DAS INFORMAÇÕES E ACESSO AOS DADOS EM PROCESSO DE PRODUÇÃO

O consultor contratado, no manuseio das informações obtidas por meio da atuação no projeto, obriga-se a observar e guardar, em toda a sua extensão, o sigilo e a confidencialidade dessas informações, observando os termos previstos no parágrafo único do art. 1º da Lei 5.534, de 14/11/68, regulamentada pelo Decreto 73.177, de 20/11/73, art. 1º, parágrafo 1º, e Decreto 74.084, de 20/05/74, art. 8º, que regulamenta o artigo 6º da Lei 5.878, de 11/05/73, que declara conhecer, bem como as disposições legais pertinentes à responsabilização penal do agente que infringir essas normas.