

Consulta Pública**SUPGA/GACOM/GABSA Nº 0666/2023****Consulta pública eletrônica ao mercado no modelo Request For Information - RFI para subsidiar estudo para futura contratação de solução de Motor SQL para Data Hub**<https://www.serpro.gov.br/consultas-publicas/sede/0666-2023>**PROPOSTAS TÉCNICAS RECEBIDAS**

Itens	Data	Interessado	Produto
1	14/07/2023	IBM Brasil	Solução IBM WatsonX.data
...			

Planilha de Atendimento aos Requisitos

2.1.1. Requisitos Funcionais para conexão e processamento de dados utilizando linguagem SQL	Forma de Atendimento	Observações
	Nativo (N)/ Nativo por Parametrização (NP) / Customizado (C) / Não Atende (NA)	
2.1.1.1. Possuir conectividade à fonte de dados relacionais, no mínimo Teradata, Oracle Exadata, Oracle, Microsoft SQL Server, PostgreSQL, IBM DB2, MySQL e MariaDB.	C	
2.1.1.2. Permitir realizar importação e exportação, utilizando recursos próprios, para fonte de dados de arquivo não estruturados como: TXT, PDF, MS Word e MS Excel.	C	Suportamos formato Open por definição CSV e PNG, que podem ser convertidos facilmente para consumo no formato desejado.
2.1.1.3. Possuir conectividade às fontes de dados em Data Lake Storage, no mínimo, S3, Azure Data	N	

Lake Storage, Google Cloud Storage, Parquet, Avro e ORC.		
2.1.1.4. Permitir realizar processamento utilizando linguagem SQL a partir dos tipos de armazenamento, no mínimo, S3, Azure Data Lake Storage, Google Cloud Storage, Parquet, Avro e ORC.	N	
2.1.1.5. Possuir compatibilidade com o padrão SQL ANSI para construções de filtros (queries).	N	
2.1.1.6. Possuir conectividade e integração de forma nativa a protocolos de serviços de diretórios: LDAP e AD, de forma a utilizar autenticação e controle de permissões baseado em objetos destes serviços.	N	
2.1.1.7. Permitir realizar cruzamento de dados entre databases diferentes de uma mesma origem, bem como de origens e tecnologias diferentes.	N	
2.1.1.8. Possuir conectividade a ambientes baseados em Apache Hadoop utilizando kerberos.	C	
2.1.1.9. Permitir consultas externas aos dados (exposição dos dados) via protocolo JDBC e ODBC.	N	
2.1.1.10. Permitir conectividade com ferramentas analíticas e de integração para consumo de dados: MS PowerBI, Tableau, Qlik Sense, Visual insight, Talend e Pentaho.	N	
2.1.1.11. Possuir sistema de controle de permissão a objetos como banco de dados, tabelas, schemas e similares.	N	
2.1.1.12. Possuir sistema para construção facilitada de metadados a partir de arquivos em formato parquet.	N	
2.1.1.13. Permitir a utilização de base de metadados de objetos baseados em Hive Metastore.	N	
2.1.1.14. Permitir exportação de dados das coleções de metadados para diversos formatos, tais como: CSV, XLSX e Tableau data.	N	Suportamos formanto Open por definição: CSV. Não suportamos: XLSX e Tableau Data.
2.1.1.15. Permitir utilizar SSL/TLS (1-way) ou (2-way), criptografando o canal de comunicação com uso de chave pública, bem como certificado digital (segurança a nível do canal de comunicação).	N	

2.1.1.16. Possuir funcionalidade onde as informações referentes ao consumo de recursos computacionais de cada query devem estar disponíveis para ser coletada de forma fácil (ex: API, arquivo de logs, base de métricas).	NP	
2.1.1.17. Permitir recurso de gestão de metadados para consumo interno e externo, possibilitando, ainda, consulta com diferentes perfis de acesso.	N	
2.1.1.18. Permitir a execução das instruções SQL de forma paralela e distribuída entre nós de processamento que compõe o ambiente.	N	
2.1.1.19. Possuir mecanismos de escalabilidade horizontal e vertical, bem como alta disponibilidade.	N	
2.1.1.20. Possuir Interface gráfica (GUI – Graphical User Interface) WEB.	N	
2.1.1.20.1. A interface deve permitir o acesso por meio de navegadores web em suas versões mais recentes (Firefox, Edge e Chrome) para sistema operacional Linux e Windows.	N	
2.1.1.20.2. A interface deve conter, no mínimo, Organização hierárquica por projetos.	NA	Suportamos via árvore de objetos com base no contexto da solução (Presto) e perfil de acesso.
2.1.1.20.3. A interface deve permitir a reutilização de scripts.	N	
2.1.1.20.4. A interface deve permitir restringir o acesso a usuários (perfis) definidos pela área técnica.	N	
2.1.1.20.5. A Interface deve permitir o uso por perfis de usuário com menor qualificação técnica ou restrição de acesso a módulos administrativos, como analistas de negócio, permitindo que eles desenvolvam seus processos de forma self-service com independência da equipe de TI.	N	
2.1.1.20.6. A Interface deve possuir gestão centralizada para monitoramento, auditoria e troubleshooting.	N	Dispomos de diferentes interfaces para monitoramento, auditoria e troubleshooting: solução, plataforma, storage e etc.

2.1.1.20.7. A interface deve permitir a construção e execução de código SQL;	N	
2.1.1.20.8. A interface deve estar disponível no idioma português (PT-BR).	N	
2.1.1.20.9. A interface deve possuir ajuda on-line no idioma português (PT-BR).	NA	A documentação ainda não está disponível em PT-BR, somente EN-US
2.1.1.20.10. A interface deve permitir a visualização facilitada para exibir estatísticas de uso operacional, limites de capacidade de armazenamento e gráficos de performance.	N	Dispomos de diferentes interfaces para visualização de estatísticas e monitoração: solução, plataforma, storage e etc.
2.1.1.20.11. A interface deve permitir incluir descrição de tabela e descrição de colunas.	N	
2.1.1.21. Possuir interface de linha de comando (CLI – Command Line Interface).	N	
2.1.1.21.1. A interface deve permitir a execução de todas as operações por meio de linha de comandos nas plataformas Windows e Linux.	N	
2.1.1.22. Possuir sistema de auditoria para identificação e registro de operações nos dados, incluindo consultas SQL.	N	
2.1.1.22.1. O sistema de auditoria deve permitir a identificação do usuário, IP do usuário, resultado da consulta (success, cancel, error), data/hora de início e fim da operação, duração, quantidade de linhas retornadas.	N	
2.1.1.23. A solução deve estar aderente ao Regulamento Geral sobre a Proteção de Dados 2016/679 (General Data Protection Regulation – GDPR) da União Europeia e à Lei Geral de Proteção de Dados (LGPD), Lei nº 13.709/2018.	N	
2.1.2. Modelo de Implementação da Solução		
2.1.2.1. A solução deverá ser totalmente implantada no ambiente de nuvem privada do SERPRO, utilizando os recursos e a infraestrutura de TI disponíveis nessa nuvem.	N	Desde que os recursos de infraestrutura de TI disponíveis atendam aos requisitos da solução.

2.1.2.2. Todas as licenças de softwares que componham a solução e não estejam disponíveis na nuvem privada do SERPRO, inclusos sistemas operacionais e bancos de dados, deverão ser fornecidas, com suporte e atualizações durante toda a vigência do contrato.	N	
2.1.2.3. A arquitetura da solução deverá ser dimensionada e informada ao SERPRO de forma a atender aos requisitos expostos.	N	
2.1.2.4. A arquitetura do ambiente de produção da solução deverá ser dimensionada prevendo uma taxa de disponibilidade de 99,5%, operando 24 (vinte e quatro) horas por dia, 7 (sete) dias por semana.	N	
2.1.2.5. A arquitetura da solução deverá ser dimensionada prevendo a utilização de zonas de disponibilidade redundantes, garantindo que caso uma zona deixe de funcionar ou tenha degradação do tempo de resposta a segunda seja acionada para atendimento das requisições.		Precisamos de mais informação
2.1.2.6. Deverão ser previstos ambientes de homologação e produção.	N	O ambiente de homologação compartilha dos requisitos dos itens 2.1.2.4 e 2.1.2.5?
2.1.2.7. A solução deverá ser implantada sobre plataforma comum de mercado, sem qualquer dependência ou presença de hardwares específicos. Caso haja algum requisito obrigatório para o funcionamento da solução, como token de hardware ou instrução específica de processador, a possibilidade de adoção deverá ser avaliada previamente pelo SERPRO.	N	
2.1.2.8. A instalação e configuração da solução na nuvem privada do SERPRO deverá ser realizada de forma assistida pelos times do SERPRO e com base em procedimentos operacionais previamente fornecidos.	N	
2.1.2.9. A solução deve ser expansível.	N	
2.1.2.10. A instalação como qualquer outra operação pertinente à administração da mesma deverá estar documentada e entregue ao SERPRO.	N	
2.1.3. Requisitos de Compatibilidade da Nuvem Privada do SERPRO		

2.1.3.1. O ambiente de nuvem privada do SERPRO é composto por um conjunto de ferramentas do programa VMware Cloud Provider (ou VMware Cloud Provider Program – VCPP), de modo que a solução deverá adotar tais ferramentas como infraestrutura base, mantendo compatibilidade com as versões de produção dos seguintes componentes:	N	
2.1.3.1.1. VMware Cloud Director (vCD);		
2.1.3.1.2. Container Service Extension (CSE);		
2.1.3.1.3. VMware vSphere/ESXi.	N	
2.1.3.2. A solução deverá ser totalmente construída em torno do protocolo IPv4 e suportar o protocolo IPv6 para comunicação com o usuário final.	N	
2.1.3.3. A intercomunicação entre os componentes internos da solução deverá ocorrer de forma criptografada.		
2.1.3.4. A solução deverá ser capaz de interoperar em ambientes com NAT (Network Address Translation) e Firewalls.	N	
2.1.3.5. A solução deverá atender os seguintes requisitos de automação:		
2.1.3.5.1. A criação do ambiente deverá ser realizada de forma automatizada, em modelo de autosserviço dentro do ambiente de nuvem privada do SERPRO;	N	
2.1.3.5.2. A ferramenta de automação selecionada deverá ser compatível com os componentes da nuvem privada do SERPRO;	N	
2.1.3.5.3. No momento da construção da arquitetura da solução, deverá ser observada e utilizada a versão mais atual da ferramenta de automação;	N	
2.1.3.5.4. O código utilizado para automação deverá ser documentado e cedido como parte da solução;	N	
2.1.3.5.5. Tanto a documentação, quanto a construção do código, deverão seguir as melhores práticas de mercado;	N	
2.1.3.5.6. Fazer uso efetivo de um sistema de controle de versão.	C	
2.1.3.6. A arquitetura da solução deverá observar os seguintes requisitos:	N	
2.1.3.6. 1. Aproveitar os recursos de auto scaling disponíveis no VMware Cloud Director de nuvem	N	

para adequar a demanda ao consumo de recursos;		
2.1.3.6.2. Informar a quantidade média de recursos (CPU e RAM) que a solução assumirá quando ociosa;	N	
2.1.3.6.3. Apresentar uma matriz de projeção da quantidade de recursos (CPU e RAM) para atender ao crescimento da demanda e facilitar cálculos futuros de capacidade;	N	
2.1.3.7. A solução poderá ser baseada em máquina virtual (VM) ou Contêiner, sendo que:	N	
2.1.3.7.1. Se for VM, a arquitetura da solução deverá explorar o uso de VMs efêmeras e a persistência dos dados (da aplicação e logs) deverá ficar desvinculada do ciclo de vida da máquina virtual;	NA	São utilizados containers/pods sobre o Red Hat OpenShift
2.1.3.7.2. Se for contêiner, a solução deve possuir compatibilidade com pelo menos uma das tecnologias: VMware Tanzu ou Red Hat OpenShift. Somada a isto, deverá aderir às boas práticas para execução em ambientes baseados em contêineres, como:	N	
2.1.3.7.2.1. Uso de namespaces;	N	
2.1.3.7.2.2. Uso de auto scaling (elasticidade);	N	
2.1.3.7.2.3. Uso de limites de requisições e recursos;	N	
2.1.3.7.2.4. Uso balanceado dos nós disponíveis (regras de afinidade e anti-afinidade);	N	
2.1.3.7.2.5. Uso de Role-Based Access Control (RBAC);	N	
2.1.3.7.2.6. Uso de contêineres de tamanho reduzido;	N	
2.1.3.7.2.7. Uso de rótulos (labels) para organização dos objetos;	N	
2.1.3.7.2.8. Uso de annotation para extensão de funcionalidades;	N	
2.1.3.7.2.9. Uso efetivo de políticas de rede e segurança;	N	
2.1.3.7.2.10. Uso de contêineres inerentemente efêmeros;	N	
2.1.3.7.2.11. Uso de volumes externos aos contêineres para persistência de dados;	N	
2.1.3.7.2.12. Uso do paradigma de desenvolvimento para microsserviços.	N	

2.1.3.8. Caso utilize armazenamento de objetos, a solução deverá observar os seguintes requisitos:	N	
2.1.3.8.1. Ser compatível com a solução de armazenamento de objetos disponível na nuvem privada do SERPRO, sendo esta compatível com o protocolo S3-like;	N	
2.1.3.8.2. Utilizar a API de programação disponível com a plataforma de armazenamento de objetos;	N	
2.1.3.8.3. Os buckets e objetos criados deverão utilizar o nível mais restrito de acesso possível, reduzindo assim a exposição de informações;	N	
2.1.3.9. A solução deverá adotar apenas os sistemas operacionais e bancos de dados suportados pelo SERPRO para composição da sua arquitetura, a saber:	N	
2.1.3.9.1. Windows Server, Red Hat, CentOS e/ou Rocky/Alma Linux, nas suas versões mais atuais;	N	
2.1.3.9.2. Oracle, Microsoft SQL Server, MariaDB, MySQL, PostgreSQL e/ou Elastic, nas suas versões mais atuais.	N	
2.1.3.9.3. Caso seja necessária a adoção de qualquer solução diferente das apontadas acima, o SERPRO deverá ser consultado previamente sobre a viabilidade de manutenção posterior, sendo que em caso de soluções proprietárias, o fornecedor deverá incluir o suporte do fabricante.	N	