

Consulta Pública Eletrônica  
para Contratação de  
Solução de Governança de  
Dados e Catálogo de  
Dados.

DIOPE/SUPEC/ECTAN  
05/2020

**Sumário**

<b>1. Objeto</b>	<b>3</b>
2. Especificação do Objeto	3
3. Publicação	19
<b>4. Período</b>	<b>20</b>
<b>5. Responsáveis</b>	<b>20</b>

## **1. Objeto**

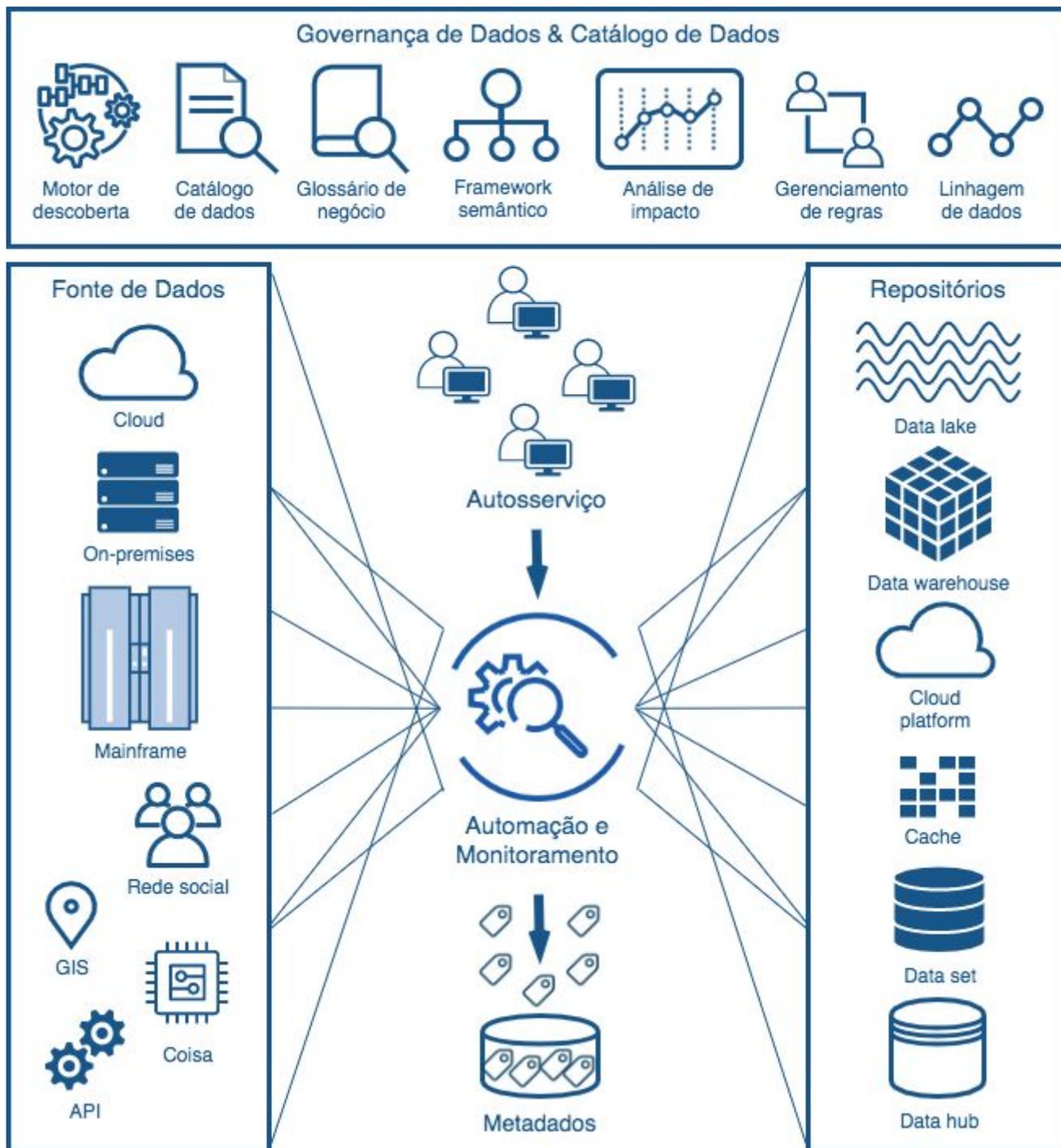
1.1. Contratação de solução de governança de dados e catálogo de dados, a ser inserida na arquitetura de dados do Serpro.

## **2. Especificação do Objeto**

2.1. Solução de governança de dados e catálogo de dados, a ser inserida na arquitetura de dados do Serpro, de modo a sustentar o gerenciamento de dados, inventários e análise de impacto e análise de casos de uso alinhados com os objetivos empresariais; e estabelecer processos de controle, funções de curadoria e métricas de integridade dos dados para melhorar a qualidade dos dados.

2.2. O escopo da consulta pública eletrônica refere-se apenas às funcionalidades de aquisição de Tipos e instâncias de metadados, Catálogo de Dados, Motor de descoberta, Glossário de negócio, Framework semântico, Análise de impacto, Gerenciamento de regras, Linhagem de dados, Segurança e Mascaramento de Dados.

### **2.2.1. Arquitetura Conceitual de Dados**



2.2.1.1. Com esta arquitetura, o Serpro deseja modernizar seu gerenciamento de dados e padrões arquitetônicos que, quando combinados, possam suportar os casos de uso de dados e análises que exigem conexão com fontes de dados distribuídas, nos levando à necessidade de maior consistência semântica, entrega otimizada de análises e governança de informações para operacionalizar novos modelos analíticos e aproveitar a reutilização sustentável de insights com governança de dados e análises, inclusive como parte da curadoria e linhagem de dados, para fornecer confiança ao usuário final no processo e na tecnologia e aos titulares de dados.

2.3. Às empresas interessadas devem responder a consulta pública com as seguintes informações:

**2.3.1. Contatos:**

2.3.1.1. Nome completo do responsável pelas respostas desta Consulta Pública.

2.3.1.2. Cargo, telefones e endereço de e-mail.

**2.3.2. Identificação da Empresa:**

2.3.2.1. Nome completo e fantasia.

2.3.2.2. CNPJ.

2.3.2.3. Endereço completo.

2.3.2.4. Site WEB (www).

**2.3.3. Solução**

2.3.3.1. Nome da solução oferecida, objeto desta consulta pública.

2.3.3.2. Site WEB do fabricante da solução (www).

2.3.3.3. Descrição detalhada da solução e seus componentes (Documentos/datasheet, etc).

2.3.3.4. Forma de licenciamento da solução e seus componentes (Licença perpétua /subscrição anual, e outras ), conforme exemplo abaixo:

Part Number	Descrição da Solução	Licença de uso perpétuo / subscrição anual	Unidade / Métrica	Faixa / Quantidade	Estudo de referência do Valor Unitário (R\$)

2.3.3.5. Forma e condições de pagamento da solução e seus componentes (Licença perpétua /subscrição anual, etc ).

**2.3.4. Base de Clientes:**

2.3.4.1. Quantidade de clientes no Brasil.

2.3.4.2. Nomes dos entes públicos que já adquiriram a solução.

**2.3.5. Experiência e Suporte:**

2.3.5.1. Possui equipe de suporte técnico para atendimento fora do horário comercial e em dias não úteis.

2.3.5.2. O suporte é prestado pelo fabricante ou parceiro?

2.3.5.3. Quais os níveis de serviços ofertados para a solução (Tempo de atendimento, tempo de solução, etc).

2.3.5.4. Informar a forma de repasse de conhecimento, resumos das grades e carga horária.

**2.4. Os requisitos funcionais e não funcionais da solução estão descritos no documento “Anexo A - Requisitos Solução Governança e Catálogo de Dados”.**

2.4.1. Orientações para preenchimento do anexo A.

2.4.1.1. Os campos a serem preenchidos nesta planilha estão marcados em azul:

2.4.1.2. O escopo para o preenchimento das respostas é dividido em campos objetivos e descritivos.

2.4.1.3. Para os campos objetivos, utilize o seguinte padrão de respostas:

2.4.1.3.1. “0” = requisito NÃO atendido pela solução.

2.4.1.3.2. “1” = SIM, requisito totalmente atendido pela solução.

2.4.1.3.1. “0,5” = requisito PARCIALMENTE atendido pela solução.

2.4.1.4. Se não atende, como poderia atender? Indique customizações necessárias e de baixa complexidade para ser atendido plenamente - a empresa se compromete a implementá-lo na totalidade.

**2.4.2. Requisitos Funcionais:**

Item	Requisitos Funcionais	Detalhamento dos Requisitos	Forma de Atendimento	
			Atende ? 0: Não 0,5: Parcial 1: Sim	Se não atende, como poderia atender?
1	Fonte de Dados	Conectividade à fonte de dados relacionais, no mínimo Oracle, Sql Server, PostGreSQL, DB2, My SQL e Maria DB.		
2		Conectividade à fonte de dados mainframe, no mínimo Adabas, DB2 e Vsam.		

3	Conectividade a fonte de dados não relacionais (NoSQL), no mínimo Elasticsearch, Hadoop, Hbase, Solr, MongoDB, Cassandra, Redis e Neo4J.		
4	Conectividade a fonte de dados em nuvem SaaS ERP, CRM (complementar Salesforce, SAP...)		
5	Conectividade a fonte de dados e appliance, no mínimo Exadata e Teradata.		
5	Conectividade à fonte de dados em vários formatos de arquivo (como XML, JSON, CSV, TXT, PDF Avro, ORC, RCFile e Parquet)		
6	Conectividade à fontes de dados disponíveis nas tecnologias e/ou protocolos JDBC, ODBC, HDFS (block size padrão 64MB), S3, NFS, SAN, LDAP, HTTP/HTTPS, FTP e SFTP.		
7	Conectividade a redes sociais, estando incluso no mínimo Facebook, Twitter, LinkedIn e Instagram;		
8	Conectividade à infraestrutura de mensageria, incluindo aquelas providas por middleware de integração, no mínimo Kafka, RabbitMQ, IBM MQ, TIBCO EMS, JMS e MQTT.		
9	Conectividade a ambientes baseados em Apache Hadoop, seja como origem ou destino dos dados, sendo compatível, no mínimo, com as seguintes tecnologias HDFS, Apache Hive, Apache Impala, HBase, Spark, Apache Kafka, Apache Kudu e Apache Solr.		
10	Conectividade a séries temporais e dados geoespaciais residentes em sistemas como GIS, ESRI e outros.		

11		Conectividade com soluções proprietárias, de provedores de nuvem, Azure SQL, Azure Cosmos DB, Amazon RDS, Amazon Redshift, Amazon S3, DynamoDB NoSQL, ElastiCache, Neptune graph, Snowflake, Glacier, services Google.		
12		Conectividade à dados de streaming via plataformas ESP (Event Stream Processing) tal como IoT, log, dados de sensores, etc.		
13	Catálogo de Dados	Existem mecanismos para criar e manter repositórios de metadados por domínios (multi-domínios, governança federada).		
14		Existem diversos tipos de metadados predefinidos, incluindo tipos com atributos primitivos, com atributos complexos, com referências a objetos ou herança de outros tipos.		
15		Suporta integração bidirecional entre metadados, inclusive a partir de conectores nativos e/ou APIs REST para procurar e extrair metadados das fontes e tipos de dados mais populares.		
16		Existem recursos para criar, manter e aplicar padrões de metadados, incluindo módulo de inteligência de negócios (BI).		
17		O repositório de metadados pode ser migrado automaticamente de uma plataforma para outra e funcionar como antes, sem intervenção manual.		
18		Definições de metadados predefinidas podem ser customizadas.		
19		Existem mecanismos para definir novos tipos para os metadados a		

		serem gerenciados.		
20		Importação/exportação de metadados em formato interoperável (XML, JSON, YAML ou CSV) ou APIs REST.		
21	Motor de descoberta	Existem mecanismos para criação de perfil de dados para descobrir o significado funcional dos dados em sua fonte de origem. Esse perfil tornar-se um objeto de repositório de metadados e pode ser utilizado na descoberta de metadados ou por outros mecanismos de terceiros, para identificar e classificar dados confidenciais em uma fonte de dados.		
22		Existem mecanismos automatizados para inventariar ativos de dados e selecionar, organizar e manter os metadados relativos a esses ativos com base em perfil de dados.		
23		Existem mecanismos automatizados para marcação automatizada de metadados com base em perfil de dados.		
24		Existem mecanismos automatizados para marcação de conteúdo ou detecção de DP (Dados Pessoais) com base em perfil de dados da ferramenta, criados segundo norma ABNT NBR ISO/IEC 27701:2019 Versão Corrigida:2020.		
25		Existem mecanismos para marcação automatizada e anotação de metadados, conforme política de governança de dados ou políticas regulatórias (LGPD).		
26		Existem conectores nativos e APIs REST para procurar e extrair metadados das fontes e tipos de dados mais populares (ver fonte de		

		dados).		
27		Usa metadados de outras ferramentas para enriquecer o repositório.		
28		Existem mecanismos de integração e intreroperabilidade com repositórios do PowerCenter e do PowerDesigner.		
29		Existem mecanismos para identificar e "limpar" inconsistências nas informações preenchidas (como reconciliação de homônimos e sinônimos).		
30		Existem mecanismos em tempo real e em lote para sincronizar alterações em um repositório de metadados, cuja definição existe em vários repositórios (controle master/slave).		
31	Glossário de negócio	O glossário de negócio é personalizável		
32		O glossário de negócio contém termos comerciais existentes utilizam alguma especificação (por exemplo, OMG SBVR).		
33		O glossário de negócio pode ser categorizado.		
34		Glossários de vocabulários podem ser pesquisados.		
35		O glossário fornece uma trilha de auditoria.		
36		O glossário de negócio permite definir palavras-chave e identificar sinônimos de palavras-chave.		
37	Framework semântico	Existem algoritmos de ML para permitir que os usuários realizem pesquisa semântica sobre metadados para encontrar os conjuntos de dados mais relevantes e, em seguida, procurar e filtrar esses conjuntos de dados derivados,		

		conforme necessário.		
38		Existem algoritmos de ML para automatizar tarefas envolvidas na descoberta de metadados, ingestão, tradução, enriquecimento e a criação de relacionamentos semânticos entre os metadados.		
39		Existem mecanismos para a aplicação de padrões de ontologia e taxonomia de metadados (por exemplo, para domínios, classes, tipos de relacionamento, categoriais, nomes de dados e abreviações).		
40		Existem diferentes versões de grafos de conhecimento (Knowledge Graph) para visibilidade de integração de informações entre entidades (conceitos, objetos, coisas ou pessoas) voltado a análises de negócio.		
41		Existem mecanismos para captar metadados sobre as funções e os padrões de uso dos usuários para gerar grafos de conhecimento de modo a se ter informações contextualizadas para os usuários finais.		
42		Existem mecanismos de grafos de conhecimento baseados em gerenciamento de regras de negócios.		
43		Existem mecanismos de abstração para construção de modelos de dados semânticos (Taxonomia, Ontologia e Graph-Tool).		
44		Suporte a tecnologias e linguagens como RDF (Resource Description Framework), protocolo SPARQL, RDF Query Language (SPARQL), Apache TinkerPop e Cypher.		

45	Análise de impacto	É possível exibir e executar análises de impacto nos relacionamentos entre objetos armazenados no repositório (ativos de dados, perfil de dados, glossário de negócio, regras de negócio, taxonomia e ontologias) .		
46		É possível exibir e executar análises de impacto nos relacionamentos entre os termos de negócio no glossário de negócio.		
47		É possível realizar análises de impacto em mudança de qualreure objeto de metadados.		
48		É possível exibir e executar análises de impacto das dependências de informações ou o impacto de uma alteração em uma fonte de dados.		
49	Gerenciamento de regras	Suporta definição e avaliação de regras de negócios e relatórios de conformidade em caso de exceção a essas regras de negócios		
50		Suporta métricas de validação dos modelos de dados, métricas de armazenamento de dados, métricas de conscientização de segurança, métricas de desempenho, métricas de desempenho e resposta, métricas de gerenciamento de problemas, métricas de governança e conformidade, métricas de implementação de segurança, métricas de incidentes de segurança, métricas de operações, métricas de proteção de dados, métricas de serviço, e métricas de valor de negócio, incluindo KPI sobre estas métricas.		
51		Suporta scorecards de metadados, qualidade dos dados, valor, eficácia e sustentabilidade, permitindo		

		inclusive customização e/ou criação de novos scorecards.		
52		Suporta scorecards relacionadas a políticas regulatórias e políticas internas de estratégia de dados.		
53		Suporta scorecards que permita identificar compartilhamento de dados, volume e utilização de compartilhamento de dados e proliferação de dados confidenciais.		
54		Combina algoritmos de ML a regras de negócio para gerar análises de impacto voltadas a tomada de decisão.		
55		Combina algoritmos de ML a regras heurísticas para produzir automaticamente regras de negócio otimizadas para gerar análises de impacto voltadas a tomada de decisão.		
54		Suporta implementatação de regras de negócio, governança de dados e/ou conformidade.		
55		Suporta variedade de recursos, desde a avaliação da política de informações até a operação e gerenciamento diários do MDM, com a criação de KPI, dashboards e relatórios de negócios.		
56		Suporta funções de governança de informações, como colaboração e criação de políticas de governança, gerenciamento de mudanças de políticas e análise de impacto, e reagir a mudanças feitas em uma camada de governança de informações interna ou externa.		
55	Linhagem de dados	Suporta diferentes níveis de granularidade (granulação grossa e fina) para responder perguntas "onde o dado foi originado?", "como		

		o dado foi transformado?", "quem tem acesso ao dado?", "justificativa do uso do dado?" e "quais consultas foram executadas no dado?".		
56		Suporta diferentes abordagens de captura (eager e lazy) para gerar e armazenar linhagem em tempo de execução ou sob demanda.		
57		Identifica riscos para a qualidade dos dados ao longo da linhagem de um determinado elemento de dados.		
58		Identifica problemas conhecidos relacionados a elemento de dados em qualquer ponto de sua linhagem.		
59		Existem mecanismos para a propagação de marcações via linhagem - garante automaticamente que as marcações sigam os dados à medida que passam por vários processos		
60		Interface do usuário intuitiva para visualizar a linhagem de dados à medida que ela se move através de vários processos, incluindo representação gráfica para que a linhagem de dados documente o ciclo de vida dos dados.		
61		APIs REST para acessar e atualizar linhagem.		
62	Relatórios	É possível gerar, exibir e imprimir representações gráficas completas de todos os modelos armazenados em repositório.		
63		É possível gerar, exibir e imprimir relatórios de comparação de versões detalhando as alterações feitas entre duas versões de um artefato.		

64		É possível gerar, exibir e imprimir relatórios de comparação de versões detalhando as alterações feitas entre duas coleções de artefatos (como as alterações refletidas em todos os artefatos em uma versão de produção versus a versão anterior que estava em produção).		
65		É possível gerar, exibir e imprimir uma representação gráfica de relacionamentos entre artefatos armazenados.		
66		Há recursos de relatórios personalizados e personalizáveis para relatórios, incluindo gráficos, operadores booleanos, palavras-chave parciais e sequências de caracteres.		
67		Os relatórios padrão e ad hoc podem ser iniciados e visualizados em todas as interfaces.		
68		A tecnologia suporta a saída de relatórios para ferramentas de produtividade do escritório, no mínimo .docx, .odt, .rtf, .xlsx, .ods.		
69	Integrações	APIs REST para consulta programática de metadados em repositório, isto é, uma integração mais fácil.		
70		APIs REST para pesquisa por critérios complexos.		
71		SQL-like para pesquisar objetos - DSL (Domain Specific Language)		
72		Integração com barramento de dados (ElasticSearch, Tibco), integração com ferramentas de monitoramento (Prometheus, Suite HPE, Suite Broadcom), integração com Analytics/BI, como PowerBI, Qlik e Tableau (entre vários outros), para que possam ser invocados		

		diretamente nessas ferramentas pelos usuários.		
73		Fornecimento de bibliotecas de desenvolvimento de software (SDK) e permita que os usuários criem conectores adicionais para fontes de dados exclusivas.		

### 2.4.3. Requisitos Não Funcionais

Item	Requisitos Funcionais	Detalhamento dos Requisitos	Forma de Atendimento	
			Atende ? 0: Não 0,5: Parcial 1: Sim	Se não atende, como poderia atender?
1	Interface com o usuário	Interfaces de usuário compatíveis com distribuições Ubuntu, Windows e MacOS em estações de trabalhos, navegadores Google Chrome, Firefox, Apple Safari e Microsoft Edge. A interface com o usuário em navegadores não poderá fazer uso de Applets Flash e Java.		
2		Interface do usuário intuitiva para pesquisar entidades por tipo, classificação, valor de atributo ou texto livre.		
3		Interface de usuário pode ser personalizada.		
4		Há barra de ferramentas que pode ser personalizada por grupo de usuário e por usuário.		
5		A interface do usuário permite a adição de menus e itens de menu personalizados.		

6		Os menus podem ser personalizados para diferentes grupos de usuários e por usuário.		
7		Além dos menus, a ferramenta fornece recursos de gerenciamento de processos para orientar os usuários nas tarefas que eles precisam executar.		
8		A solução fornece interface e ajuda a recursos sensível ao idioma português brasileiro.		
9		A solução fornece traduções de regras armazenadas em repositórios da solução no idioma português brasileiro.		
10		É possível criar "visualizações" agregadas (organização hierárquica) por escopo definido com base na função do usuário (role) e/ou domínio de dados (por exemplo, apenas os artefatos, consultas, análises etc. do domínio contábil).		
11	Gerenciamento de Configuração	Possui recursos de controle de versão.		
12		Suporta discussões encadeadas em torno de documentos/conteúdo.		
13		Suporta várias versões simultâneas de um artefato e seus relacionamentos.		
14		Tem a capacidade de "congelar" uma configuração para evitar atualizações indesejáveis enquanto permite que alterações em outros artefatos continuem.		
15		O repositório pode sincronizar alterações entre versões de configurações movendo coleções de artefatos (componentes) entre versões de configuração?		

16		No caso de fallback, o repositório inteiro deve ser restaurado para uma configuração congelada ou o fallback seletivo pode ser feito fazendo o backup das alterações apenas nos artefatos selecionados?		
17	Segurança	Integração com Red Hat Directory Server 10, para autenticação de passagem (pass-through authentication) e autenticação baseada em perfis (role-based authentication).		
18		Log e auditoria, permitindo identificar, responsável por inclusão, alteração ou exclusão de objetos de dados mantidos pela ferramenta.		
19		Segurança granular para acesso a metadados, permitindo controles sobre o acesso a instâncias e operações de entidades, como adicionar / atualizar / remover marcações, regras de negócio, ou quaisquer objetos mantidos pela ferramenta.		
19		Segurança granular baseada em perfis, permitindo associar funcionalidades da ferramenta a determinados perfis.		
20	Suporte a mascaramento dinâmico de dados	Em consultas realizadas via interfaces da solução, limitar a exposição de dados confidenciais através do mascaramento dinâmico dos dados para usuários sem privilégios.		
21		Permite integração com outras ferramentas de mercado para autorização / mascaramento de dados no acesso a dados com base nas marcações associadas aos objetos.		

22		Permite integração com outras ferramentas de mercado que possua as características de Criptografia, Mascaramento e Tokenização de forma paralela.		
23		O mascaramento em tempo real é aplicado quando um usuário acessa dados com base em privilégios de acesso.		

### 3. Publicação

3.1. Consulta pública com fulcro no Art. 31, da Lei nº 9.784/1999, objetivando esclarecimentos sobre produtos, processos, soluções e tecnologias junto ao mercado.

### 4. Período

4.1. A consulta pública eletrônica ficará publicada pelo período de 8 (dias) úteis, podendo ser prorrogado a critério do SERPRO.

### 5. Responsáveis

5.1. A Consulta Pública Eletrônica será acompanhada pelos empregados:

5.1.1. Charles Morais Magalhães, matrícula 21065411, lotado na DIOPE/SUPEC/ECTAN, Telefone: (61) 2021-7259, e-mail: [charles.magalhaes@serpro.gov.br](mailto:charles.magalhaes@serpro.gov.br).

5.1.2. Flavia Costa Bomfim, matrícula 01093649, lotado na SUPAI/AIGSB/AIGUP, e-mail: [flavia.bomfim@serpro.gov.br](mailto:flavia.bomfim@serpro.gov.br).