



## CONSULTA PRELIMINAR AO MERCADO DAG/DSI Nº 10

### Solução de Data Lake

#### Nota legal:

Esta apresentação é apenas uma versão preliminar do projeto pretendido, partilhada apenas para fins de informação geral, não podendo ser considerada versão final, nem vinculativa.

As informações contidas neste documento podem estar sujeitas a alterações, não comprometendo nem vinculando os Serviços Partilhados do Ministério da Saúde, EPE e/ou quaisquer outros serviços e/ou órgãos do Ministério da Saúde ou do Serviço Nacional de Saúde.

#### I. ENQUADRAMENTO

---

A SPMS tem por missão a prestação de serviços partilhados nas áreas de compras e logística, serviços financeiros, recursos humanos e sistemas e tecnologias de informação e comunicação às entidades com atividade específica na área da saúde, de forma a "*centralizar, otimizar e racionalizar*" a aquisição de bens e serviços no Serviço Nacional de Saúde.

Os Sistemas de Informação na Saúde permitem a cooperação, a partilha de conhecimentos e informação, bem como o desenvolvimento de atividades de prestação de serviços nas áreas dos sistemas e tecnologias de informação e comunicação. Desempenham um papel importante na reforma do sistema de saúde, tendo como principais objetivos a melhoria da acessibilidade, eficiência, qualidade e continuidade dos cuidados e o aumento da satisfação dos profissionais e cidadãos.

À SPMS cabe, ainda, a garantia da operacionalidade e segurança das infraestruturas tecnológicas e dos sistemas de informação do Ministério da Saúde, promovendo a definição e a utilização de normas, metodologias e requisitos que garantam a interoperabilidade e interconexão dos sistemas de informação da saúde entre si, e com os sistemas de informação transversais à Administração Pública, visando desenvolver e proteger a saúde dos cidadãos.

#### II. OBJETIVO

---

Pretende assim a SPMS, EPE vir a adquirir uma **Solução de Data Lake**, pelo que com vista à preparação do respetivo procedimento aquisitivo, e fazendo uso do disposto no artigo 35.º-A do Código dos



Contratos Públicos, a SPMS, EPE vem, nos termos da denominada "consulta preliminar ao mercado", solicitar informações sobre o objeto do contrato.

A solução de Data Lake para o SNS pretendida tem os seguintes objetivos macro:

1. Criar um repositório com elevada capacidade de armazenamento para salvaguardar e trabalhar todo o tipo de dados, estruturados e não estruturados. O repositório deverá ser distribuído entre diferentes centros de dados, deslocalizados;
2. Implementar estruturas de processamento e controlo da solução, integrando o repositório de armazenamento, permitindo processamento rápido e gestão de todas as ferramentas necessárias para montar o Data Lake;
3. Integração de diferentes streams de dados com tratamento em tempo real, com grande capacidade de processamento de dados em simultâneo, introduzindo inteligência de dados, análise preditiva e machine learning;
4. Exposição em tempo real de indicadores, dados e estruturas através de ferramentas de dashboarding, bus de integração ou serviços implementados à medida. Disponibilização de APIs de ingestão e consulta de informação em tempo real.

A definição e implementação do modelo de governação de dados para o SNS pretendida tem os seguintes objetivos macro:

1. Levantamento e análise da realidade e capacidade atual das soluções de Business Intelligence e das várias fontes de dados;
2. Apresentação de um modelo de governação de dados para o Data Lake e para o modelo de governação de dados atual, tendo consideração a necessidade de evolução e uniformização das várias soluções de Business Intelligence existentes;
3. Definição de um roadmap de implementação e de um plano de transição/gestão de mudança;
4. Conceitos de governação de dados e ferramentas a adotar para implementação da metodologia proposta.

#### **Arquitetura:**

##### ❖ **Gestão:**

1. Camada de Gestão de toda a solução, física e lógica;
2. Redundância física e geográfica da solução;
3. Escalável e capaz de orquestrar toda a solução;



4. Integração com diversos softwares e ferramentas do âmbito de salvaguarda, ingestão e tratamento de dados.

❖ **Processamento:**

1. Camada de Processamento de toda a solução, física e lógica;
2. Redundância física e geográfica da solução;
3. Mínimo (adaptável perante cada solução) 12Tb RAM;
4. Mínimo (adaptável perante cada solução) 768 cores;
5. Espaço em disco para cache ou aceleração de processamento.

❖ **Armazenamento:**

1. Camada de Armazenamento de toda a solução, física e lógica;
2. Redundância física e geográfica da solução;
3. Conetividade a 40Gb;
4. 3 Pb de espaço útil (sem compressão ou mecanismos de poupança de espaço).

### **Funcionalidades a suportar**

As funcionalidades identificadas devem ser parte integrante da solução, compatíveis com a arquitetura física descrita ou genéricas, devendo ser propostas tecnologias e arquitetura nas componentes:

1. Armazenamento de dados e indexação
2. Computação de pesquisa de informação
3. Integração de fontes de dados em tempo real
4. Reporting, dashboarding, análise e consulta de informação

### **Serviços cloud a Integrar**

A integrar com a solução de hardware e software montada, são pretendidos serviços *cloud* especializados nas seguintes áreas:

1. Serviços de machine learning
2. Serviços de analítica e dashboarding
3. Outros serviços relevantes identificados para a solução



### **Suporte da Solução**

O fornecimento do objeto do contrato contempla os direitos de utilização perpétua do hardware acima identificado, bem como de todo o software necessário para o funcionamento em pleno desse mesmo hardware. Incluindo atualizações e o fornecimento do respetivo suporte durante 5 anos na modalidade 24 horas, 7 dias por semana com tempo de resposta de 4 horas (24x7x4) nas seguintes vertentes:

1. Assistência na deteção de problemas;
2. Correção de anomalias;
3. Atendimento e resolução de dúvidas;
4. Acesso a todas as atualizações lançadas no mercado;
5. Substituição do hardware em caso de falha;
6. Abertura de casos de suporte diretamente com o fabricante.

### **Serviços de Instalação**

De modo a implementar a solução de *hardware e software* identificada, são necessários os seguintes:

1. Serviços de instalação de todas as componentes de hardware e software adquiridas, integrando todas de modo a providenciar a solução final preconizada;
2. Serviços de teste da solução, documentação e formação de utilização e tecnológica;
3. Serviços de levantamento da situação atual e de implementação do processo e atividade de mudança.

### **Governança de dados**

Com a criação de um *data lake* no SNS, existe a necessidade de ter bem desenhado um plano de governação de dados em massa, considerando o *as-is* da saúde em Portugal e definindo um *roadmap* para o *to-be*, tendo em consideração o contexto existente. Neste sentido, pretendemos que nos sejam apresentados os melhores conceitos e ferramentas de governação de dados associados ao *data lake* e para as restantes soluções de *Business Intelligence*, tendo em consideração as seguintes necessidades:

1. Listagens periódicas de dados (em diferentes formatos (xlsx, csv, xml, etc);
2. Acesso à informação em ferramentas de reporting, p.e, powerbi, considerando a necessidade uniformização do acesso transversal à informação.
3. Exploração de dados simplificada à medida, permitindo customização, mediante perfis de utilização.



### Informação pretendida

Assim, na presente consulta preliminar ao mercado, pretende-se identificar:

1. O preço base a considerar pela entidade adjudicante face aos equipamentos pretendidos;
2. O preço base a considerar pela entidade adjudicante para os serviços de instalação;
3. Prazo considerado necessário para a entrega dos equipamentos. Informar da necessidade de entregas faseadas;
4. Prazo considerado necessário para a instalação dos equipamentos;
5. Previsão de custos de manutenção a 10 anos;
6. Preço base por cada 1 PB útil adicional;
7. Preço base por cada nó de computação adicional;
8. Identificação dos serviços cloud a integrar (fornecedor, serviço específico) e respetivos custos associados.
9. High level Roadmap de implementação do data lake e restantes componentes, bem como o plano de transição referente à definição e operacionalização do modelo de governação de dados;
10. Custos associados a este plano de transição e implementação;
11. Custos das ferramentas propostas em cada cenário, bem como custos da sua manutenção.

### III. FORMA DA CONSULTA

---

É imperativo que a consulta preliminar ao mercado seja conduzida com transparência e não haja tratamento desigual de operadores económicos, conforme dispõe o artigo 35.º-A do Código dos Contratos Públicos.

Assim, a consulta preliminar ao mercado será publicitada no portal de internet público da SPMS, EPE, em <https://spms.min-saude.pt/>, e num jornal de grande tiragem nacional, devendo os operadores económicos interessados em apresentar contributos no âmbito da Solução de *Data Lake*, remeter email para o endereço [consulta.preliminar@spms.min-saude.pt](mailto:consulta.preliminar@spms.min-saude.pt) até ao dia 11/01/2022, de modo a poderem estar presentes numa sessão plenária (data prevista 13/14 janeiro).

Em data a designar (prevista 27 e 28 de janeiro de 2022), a entidade adjudicante convidará cada um dos operadores económicos individualmente a efetuar apresentação de contributos para a solução que terá



a duração máxima de 1 hora. Nessa sessão a entidade adjudicante reserva um período adicional de 20 minutos para solicitar esclarecimentos.

#### **IV. MODELO DE APRESENTAÇÃO**

---

1. A apresentação deverá ser em modelo remoto;
2. Apresentação focada na solução pretendida, sendo que a apresentação da componente de infraestrutura não pode ultrapassar mais de 5 minutos;
3. Podem ser apresentadas propostas para apenas algumas vertentes desta consulta (por exemplo: apenas apresentar solução de hardware ou apenas apresentar solução de software);
4. Demonstração da solução ou casos práticos reais, quando aplicável;
5. Estimativa de custos (tal como referenciado no ponto II);
6. Metodologia de projeto;
7. A apresentação poderá incidir em apenas um dos temas levantados nesta consulta;
  1. Hardware – 5 minutos;
  2. Arquitetura e tecnologia – 25 minutos;
  3. Governação de dados – 25 minutos;
8. O tempo disponível para apresentação dependerá dos pontos que os agentes económicos indicarem que pretendem apresentar;
9. Em todas as vertentes, exceto no hardware, devem ser reservados cerca de 10 minutos para apresentação da metodologia de projeto e casos de uso reais já existentes.

#### **V. PRESTAÇÃO DE INFORMAÇÃO PELOS OPERADORES ECONÓMICOS**

---

Os operadores económicos interessados em apresentar contributos no âmbito da presente Consulta Preliminar devem remeter email para o endereço [consulta.preliminar@spms.min-saude.pt](mailto:consulta.preliminar@spms.min-saude.pt) até ao dia 11/01/2022, devendo os interessados indicar claramente no assunto do email a referência “Consulta Preliminar n.º 10 - Solução de Data Lake”.



## VI. INFORMAÇÃO PRETENDIDA

---

A informação a prestar pelos operadores económicos, considerada por eles como oportuna e relevante, é a seguinte:

- Detalhes do operador económico: Nome, endereço, site (se existir), contacto telefónico e email;
- Áreas de especialidade e atuação;
- Indicação do CAE
- Informação do equipamento, serviço ou do seu portefólio, com os detalhes que considerar relevante para o objeto da consulta preliminar.

## VII. PRAZO DA CONSULTA

---

Os operadores económicos interessados em apresentar contributos no âmbito da presente Consulta Preliminar deverão remeter email para o endereço [consulta.preliminar@spms.min-saude.pt](mailto:consulta.preliminar@spms.min-saude.pt) até ao dia 11/01/2022.