GUIA COMPLEMENTAR | 2026

Os Seus Dados Estão Prontos para a IA? 10 Verificações e Como Resolvê-las

Como cumprir cada uma das dez verificações da auto-avaliação do guia “Feeding AI the Right Information”, com passos práticos, opções de solução e provas claras de sucesso

Para responsáveis de TI, arquitectos e equipas de dados e de IA que querem melhorar a pesquisa e a recuperação de informação por trás de assistentes de IA, agentes e motores de pesquisa empresariais.

Preparado pela Sun Technosystems. Os nomes de ferramentas são apenas exemplos e não recomendações; confirme a sua adequação ao seu ambiente.

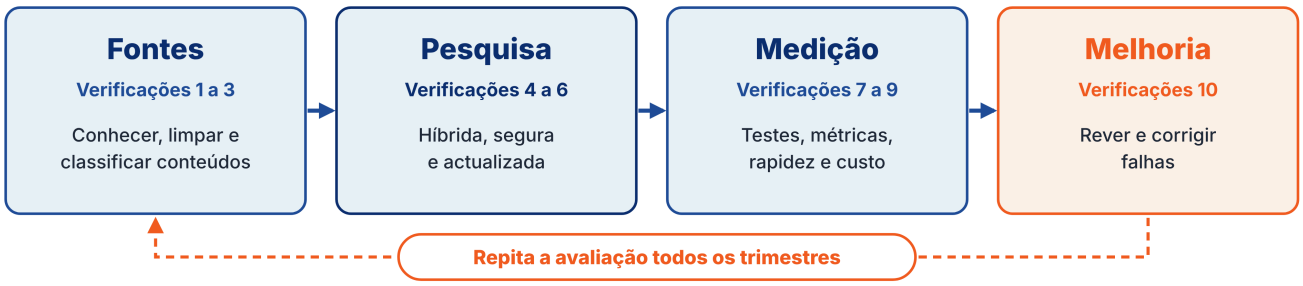
sales@suntechnosystems.net | +258 85 736 6286

Marcar Chamada de Descoberta

Como usar este guia

O nosso guia *Feeding AI the Right Information* (Alimentar a IA com a Informação Certa) termina com uma auto-avaliação de dez pontos. Muitos leitores fizeram-nos a pergunta óbvia que se segue: “Está bem, mas como é que lá chegamos?” Este guia complementar é a nossa resposta. Para cada verificação, encontrará porque é importante, os passos práticos, os tipos de solução disponíveis (com exemplos conhecidos) e as provas que lhe permitem assinalar a verificação com honestidade.

As dez verificações dividem-se em quatro grupos. Siga-as mais ou menos por ordem. Não adianta muito afinar a pesquisa ou medir a qualidade se os conteúdos que estão por baixo estiverem desactualizados ou mal classificados.



As dez verificações em quatro grupos, repetidas como um ciclo regular de melhoria.

Comece por avaliar-se

Antes de começar, dê a cada verificação uma pontuação honesta de 0 a 2. Depois, comece pelos zeros que mais pesam no seu caso de uso de IA mais importante.

Pontuação	Significado
0	Não existe, ou existe apenas na cabeça de algumas pessoas
1	Existe em parte, de forma inconsistente ou sem documentação
2	Existe, está documentado, tem responsável e é verificado regularmente

#	Verificação	Pontuação (0 a 2)
1	Sabemos de que fontes depende cada caso de uso de IA	
2	Os documentos desactualizados e duplicados são removidos ou claramente identificados	
3	Os documentos têm metadados como data, estado, responsável e nível de acesso	
4	Combinamos pesquisa por palavras-chave, por significado e estruturada quando necessário	
5	As permissões são aplicadas durante a pesquisa, e não depois	
6	As alterações importantes chegam ao índice de pesquisa com rapidez suficiente	
7	Temos um conjunto de teste com perguntas reais e respostas esperadas	
8	Medimos a pesquisa separadamente da resposta da IA	

Marcar Chamada de Descoberta

#	Verificação	Pontuação (0 a 2)
9	Acompanhamos a rapidez e o custo por pergunta, e não só a qualidade	
10	Revemos e corrigimos regularmente as falhas mais comuns	

VERIFICAÇÃO 1

Sabemos de que fontes depende cada caso de uso de IA

PORQUE É IMPORTANTE

Se não sabe onde um sistema de IA vai buscar a informação, não consegue avaliar se as respostas são fiáveis, se tem acesso a coisas que não devia ver, nem o que vai falhar quando uma fonte mudar.

Como lá chegar

- Liste cada caso de uso de IA e as perguntas a que deve responder.
- Para cada caso de uso, registe todas as fontes: bibliotecas de documentos, bases de dados, aplicações, sites e pastas partilhadas.
- Atribua a cada fonte um responsável, uma descrição, o nível de sensibilidade e a frequência com que muda.
- Registe como cada fonte está ligada (conector, exportação, API) e com que frequência é actualizada.
- Reveja a lista sempre que um caso de uso ou uma fonte mudar, e pelo menos todos os trimestres.

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Registo simples de fontes	Uma folha de cálculo ou página wiki partilhada por caso de uso	Suficiente para um ou dois casos de uso; pode começar hoje
Catálogo de dados (código aberto)	DataHub, OpenMetadata	Descoberta automática, linhagem e responsáveis; requer alguma configuração técnica
Catálogo de dados (comercial)	Microsoft Purview, Collibra, Alation	Funcionalidades de governação mais amplas; custos de licença
Configuração da plataforma de IA	Definições de fontes de dados e conectores na sua plataforma de pesquisa ou assistente de IA	Costuma listar as fontes ligadas, mas raramente os responsáveis ou a finalidade

PODE ASSINALAR QUANDO

Cada caso de uso de IA tem uma lista actualizada de fontes, cada uma com responsável, nível de sensibilidade e método de actualização, revista no último trimestre.

Marcar Chamada de Descoberta

VERIFICAÇÃO 2

Os documentos desactualizados e duplicados são removidos ou claramente identificados

PORQUE É IMPORTANTE

A IA não tem forma de saber qual de dois documentos semelhantes é o que está em vigor. Políticas antigas, rascunhos e cópias levam a respostas que parecem seguras e estão simplesmente erradas, e os duplicados afastam os melhores resultados.

Como lá chegar

- Acorde o que significa “em vigor”: estado aprovado, data de entrada em vigor e versão.
- Encontre duplicados exactos comparando as impressões digitais dos ficheiros, e quase-duplicados comparando a semelhança do texto.
- Peça aos responsáveis pelos conteúdos que confirmem, arquivem ou apaguem o material desactualizado, começando pelas fontes que a IA mais usa.
- Mova os documentos substituídos para um arquivo que a IA não pesquisa, ou identifique-os claramente como substituídos.
- Defina regras de retenção e revisão para que os conteúdos sejam revistos ou expirem automaticamente.

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Políticas de retenção e de ciclo de vida	Etiquetas de retenção no Microsoft 365, regras de retenção no Google Vault	Automatiza a revisão e eliminação; acorde as regras primeiro com a área jurídica
Detecção de duplicados	Ferramentas de hash de ficheiros; detecção de quase-duplicados com técnicas de semelhança de texto como MinHash	Pode ser automatizada por uma equipa de dados; reveja os resultados antes de apagar
Funcionalidades de gestão documental	Controlo de versões e fluxos de aprovação no seu sistema de gestão documental	Evita novos duplicados se forem usados de forma consistente
Ciclos de revisão de conteúdos	Lembretes de revisão aos responsáveis a cada 6 ou 12 meses	Baixo custo; depende da disciplina dos responsáveis

PODE ASSINALAR QUANDO

As fontes usadas pela IA não têm duplicados conhecidos, os documentos substituídos estão arquivados ou identificados, e cada documento tem uma data de revisão ou de expiração.

Marcar Chamada de Descoberta

VERIFICAÇÃO 3

Os documentos têm metadados úteis como data, estado, responsável e nível de acesso

PORQUE É IMPORTANTE

São os metadados que permitem à pesquisa filtrar e ordenar com sentido: só documentos aprovados, a área de negócio certa, os mais recentes primeiro e apenas o que este utilizador pode ver. Sem eles, a pesquisa só tem as palavras do texto para se orientar.

Como lá chegar

- Acorde um conjunto pequeno e obrigatório de campos: título, responsável, data, estado, tipo de documento, área de negócio e sensibilidade.
- Configure os seus sistemas documentais para exigirem estes campos quando os conteúdos são guardados.
- Use a classificação automática para preencher lacunas nos conteúdos existentes e peça aos responsáveis que confirmem os resultados.
- Aplique etiquetas de sensibilidade que acompanham o documento para onde quer que vá.
- Garanta que os metadados passam para o índice de pesquisa e não ficam esquecidos no sistema de origem.

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Definições do sistema documental	Tipos de conteúdo e metadados geridos no SharePoint; campos obrigatórios noutros sistemas documentais	O melhor sítio para registar metadados na origem
Etiquetas de sensibilidade	Microsoft Purview Information Protection e ferramentas semelhantes	Apoia o controlo de acessos e as obrigações de protecção de dados
Classificação automática	Serviços de classificação de texto com IA de fornecedores de cloud, ou modelos de código aberto	Rápida para grandes volumes; verifique a precisão numa amostra
Enriquecimento do índice	Processos de enriquecimento na plataforma de pesquisa que acrescentam etiquetas durante a indexação	Mantém os metadados consistentes entre fontes

PODE ASSINALAR QUANDO

Pelo menos os campos essenciais acordados estão preenchidos na grande maioria dos documentos usados pela IA e aparecem como filtros no índice de pesquisa.

Marcar Chamada de Descoberta

VERIFICAÇÃO 4

Combinamos pesquisa por palavras-chave, por significado e estruturada quando necessário

PORQUE É IMPORTANTE

A pesquisa por palavras-chave é boa a encontrar códigos e nomes exactos. A pesquisa por significado (vectorial) lida com perguntas escritas de forma diferente da fonte. As consultas estruturadas tratam datas, valores e estados. Cada uma falha coisas que as outras encontram, por isso precisa das três a trabalhar em conjunto.

Como lá chegar

- Reveja as suas perguntas de teste e identifique as que precisam de termos exactos, de significado ou de filtros estruturados.
- Indexe os conteúdos para pesquisa por palavras-chave e vectorial, com os metadados guardados para filtragem.
- Junte as listas de resultados, por exemplo com fusão por posição recíproca (reciprocal rank fusion), para que os documentos fortes em qualquer dos métodos subam para o topo.
- Acrescente um passo de reordenação que reorganiza os primeiros resultados com um modelo mais preciso.
- Compare os resultados antes e depois com o seu conjunto de teste (Verificação 7).

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Motores de pesquisa com suporte híbrido	Elasticsearch, OpenSearch, Apache Solr	Pesquisa por palavras-chave madura com funcionalidades vectoriais; amplamente suportados
Pesquisa gerida na cloud	Azure AI Search, Amazon OpenSearch Service, Google Vertex AI Search	Mais rápida de começar; considere a localização dos dados e os custos de utilização
Bases de dados com suporte vectorial	PostgreSQL com pgvector, a par da sua pesquisa de texto completo	Boa opção se os seus dados já estão em PostgreSQL
Bases de dados vectoriais	Qdrant, Weaviate, Milvus, Pinecone	Pesquisa vectorial forte; confirme as funcionalidades de palavras-chave e filtros
Modelos de reordenação	Modelos de reordenação de código aberto, ou serviços comerciais	Melhora a ordem dos primeiros resultados; acrescenta algum atraso e custo

PODE ASSINALAR QUANDO

A sua pesquisa executa consultas por palavras-chave, vectoriais e com filtros em conjunto, os resultados são combinados e reordenados, e os testes mostram melhoria face a um único método.

Marcar Chamada de Descoberta

VERIFICAÇÃO 5

As permissões são aplicadas durante a pesquisa, e não depois

PORQUE É IMPORTANTE

Se o índice de pesquisa ignorar as permissões, um assistente de IA pode ler e resumir tranquilamente conteúdos que o utilizador nunca deveria ver, mesmo quando o sistema de origem o teria impedido.

Como lá chegar

- Guarde no índice de pesquisa, junto com o conteúdo, a lista de acessos de cada documento (utilizadores e grupos).
- No momento da pesquisa, filtre os resultados pela identidade e pelos grupos do utilizador autenticado, antes de qualquer coisa chegar ao modelo de IA.
- Mantenha as permissões sincronizadas quando mudam no sistema de origem.
- Exclua totalmente as fontes muito sensíveis até os controlos estarem comprovados.
- Teste com contas que representem diferentes funções e inclua testes de permissões em cada nova versão.

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Filtragem de segurança nas plataformas de pesquisa	Segurança ao nível do documento no Elasticsearch ou OpenSearch; filtros de segurança no Azure AI Search	Exige que as listas de acesso sejam indexadas e mantidas actualizadas
Pesquisa empresarial integrada	Pesquisa no Microsoft 365, Google Workspace e pacotes semelhantes	Normalmente respeita as permissões de origem automaticamente
Conectores que transportam permissões	Estruturas de conectores que importam as listas de acesso com o conteúdo	Confirme a rapidez com que as alterações de permissões são aplicadas
Integração de identidade	Microsoft Entra ID, LDAP ou outros fornecedores de identidade	Os grupos são a base de uma filtragem fiável

PODE ASSINALAR QUANDO

Os testes de permissões com várias funções passam de forma consistente, as alterações de permissões chegam ao índice dentro do prazo acordado e nenhum conteúdo sem filtro chega ao modelo de IA.

Marcar Chamada de Descoberta

VERIFICAÇÃO 6

As alterações importantes chegam ao índice de pesquisa com rapidez suficiente

PORQUE É IMPORTANTE

Preços, níveis de stock, políticas e estados de processos mudam. Se o índice ficar para trás, a IA dá respostas que estavam certas ontem e estão erradas hoje.

Como lá chegar

- Defina o atraso aceitável para cada fonte: minutos para stock ou preços, um dia para políticas.
- Substitua a reindexação completa por actualizações incrementais que só processam as alterações.
- Use notificações de alteração ou captura de alterações de dados (change data capture) para que as actualizações sejam enviadas assim que acontecem.
- Monitorize, com alertas, o atraso entre uma alteração na fonte e o seu aparecimento na pesquisa.
- Trate as eliminações tal como os acréscimos, para que os conteúdos removidos desapareçam dos resultados.

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Captura de alterações de dados	Debezium, muitas vezes com um fluxo de mensagens como o Apache Kafka	Actualizações quase em tempo real a partir de bases de dados; exige esforço de engenharia
Notificações de alteração	Webhooks e notificações de alteração de aplicações SaaS e pacotes de colaboração	Eficiente para documentos e registos em aplicações na cloud
Conectores incrementais	Sincronização incremental agendada na sua plataforma de pesquisa ou de IA	Simples de manter; confirme o intervalo mínimo suportado
Monitorização da actualidade	Painéis e alertas que medem o atraso do índice por fonte	Transforma a actualidade num nível de serviço mensurável

PODE ASSINALAR QUANDO

Cada fonte tem uma meta de actualização acordada, o atraso do índice é monitorizado e as alterações e eliminações recentes aparecem na pesquisa dentro dessa meta.

Marcar Chamada de Descoberta

VERIFICAÇÃO 7

Temos um conjunto de teste com perguntas reais e respostas esperadas

PORQUE É IMPORTANTE

Sem um conjunto de teste, cada alteração fica entregue a opiniões e a meia dúzia de exemplos marcantes. Com ele, pode medir a qualidade, comparar opções de forma justa e detectar problemas antes dos seus utilizadores.

Como lá chegar

- Recolha perguntas reais de utilizadores, pedidos de apoio e registos de pesquisa, e não perguntas inventadas.
- Para cada pergunta, registe o documento ou registo que a deve responder e uma resposta modelo curta.
- Inclua os casos difíceis: códigos exactos, perguntas em todas as línguas que os seus utilizadores usam, temas com documentos desactualizados e perguntas a que os utilizadores não devem ter resposta.
- Comece com cerca de 50 perguntas e vá aumentando; mantenha os dados pessoais fora do conjunto de teste.
- Controle as versões do conjunto de teste e acrescente-lhe todas as falhas importantes detectadas em produção.

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Folha de cálculo ou tabela	Pergunta, fonte esperada, resposta esperada, categoria	O ponto de partida eficaz mais simples
Estruturas de avaliação de código aberto	Ragas, DeepEval, promptfoo, TruLens	Executam o conjunto de teste automaticamente e pontuam os resultados
Geração de perguntas sintéticas	Usar um modelo de IA para criar perguntas adicionais a partir dos seus documentos	Útil para cobertura; reveja sempre e misture com perguntas reais
Funcionalidades de avaliação das plataformas	Ferramentas de avaliação integradas nas plataformas de IA na cloud	Práticas se já usa essa plataforma

PODE ASSINALAR QUANDO

Existe um conjunto de teste documentado e com controlo de versões, com perguntas reais, fontes e respostas esperadas, executado antes de cada alteração importante.

VERIFICAÇÃO 8

Medimos a pesquisa separadamente da resposta da IA

PORQUE É IMPORTANTE

Quando uma resposta está errada, precisa de saber que parte falhou. A pesquisa não encontrou a informação certa, ou o modelo usou mal o que encontrou? As correcções são completamente diferentes.

Marcar Chamada de Descoberta

Como lá chegar

- Para cada pergunta de teste, verifique se a fonte esperada aparece nos primeiros resultados antes de a IA escrever qualquer coisa.
- Acompanhe as medidas de pesquisa habituais: abrangência (recall: encontrámos?), precisão (quanto do que devolvemos era útil?), posição recíproca média (MRR: a que altura aparece o primeiro resultado certo?) e nDCG (quão bem está ordenada a lista).
- Avalie em separado a resposta da IA quanto à correcção e à fidelidade às fontes encontradas.
- Registe que componente causou cada falha: fontes, pesquisa, ordenação ou modelo.
- Apresente os dois conjuntos de medidas lado a lado depois de cada alteração.

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Bibliotecas de métricas de pesquisa	pytrec_eval, ranx (bibliotecas Python)	Medidas padrão de recuperação de informação a partir de um ficheiro de resultados
Estruturas de avaliação de RAG	Ragas (precisão e abrangência de contexto); métricas de pesquisa do DeepEval	Medem a qualidade da pesquisa e da resposta numa só execução
Ferramentas das plataformas de pesquisa	Funcionalidades de avaliação de ordenação em motores como Elasticsearch e OpenSearch	Úteis para afinar a pesquisa por si só
Avaliação simples em folha de cálculo	Marcar “fonte certa nos 5 primeiros: sim ou não” por pergunta	Rápida e surpreendentemente informativa

PODE ASSINALAR QUANDO

Consegue reportar em separado a qualidade da pesquisa e a qualidade das respostas no seu conjunto de teste, e cada falha está atribuída a um componente.

VERIFICAÇÃO 9

Acompanhamos a rapidez e o custo por pergunta, e não só a qualidade

PORQUE É IMPORTANTE

Uma resposta certa que demora demasiado, ou que custa demasiado quando milhares de pessoas a usam, não é sustentável. Os agentes de IA que pesquisam muitas vezes por tarefa agravam ambos os problemas.

Como lá chegar

- Meça o tempo de resposta de cada passo: tratamento da pergunta, pesquisa, reordenação e geração pela IA.
- Acompanhe as respostas lentas nos períodos de maior movimento, e não apenas as médias.
- Calcule o custo por pergunta: uso do modelo de IA, serviço de pesquisa, infra-estrutura e reordenação.

Marcar Chamada de Descoberta

- Defina metas e alertas para ambos, por caso de uso.
- Reduza custos e atrasos com cache, modelos mais pequenos para passos simples, menos excertos mas melhores, e infra-estrutura bem dimensionada.

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Rastreio e observabilidade	OpenTelemetry, com painéis como o Grafana	Forma neutra de medir cada passo, sem depender de fornecedor
Ferramentas de observabilidade de LLM	Langfuse, Arize Phoenix	Rastreiam instruções, pesquisa, latência e custo de tokens por pedido
Gestão de custos na cloud	AWS Cost Explorer, Azure Cost Management, relatórios do Google Cloud Billing	Mostram os gastos em infra-estrutura e serviços; etiquete os recursos por caso de uso
Cache	Cache de respostas e cache semântica na aplicação ou no gateway de IA	Reduz custo e atraso em perguntas repetidas; gira com cuidado a actualidade e as permissões

PODE ASSINALAR QUANDO

Conhece os tempos de resposta típicos e mais lentos e o custo por pergunta de cada caso de uso, face a metas acordadas, com alertas quando se desviam.

VERIFICAÇÃO 10

Revemos e corrigimos regularmente as falhas mais comuns

PORQUE É IMPORTANTE

A qualidade melhora com pequenas correcções contínuas baseadas em provas, e não com um grande projecto. Sem uma revisão regular, as mesmas falhas voltam sempre.

Como lá chegar

- Dê aos utilizadores uma forma simples de assinalar respostas fracas, com um comentário opcional.
- Reveja as respostas assinaladas e os resultados dos testes a cada uma ou duas semanas.
- Organize as falhas por causa: conteúdo em falta, conteúdo desactualizado, metadados, pesquisa, ordenação, permissões ou modelo.
- Corrija primeiro as causas mais frequentes e acrescente cada falha ao conjunto de teste (Verificação 7).
- Acompanhe o número de falhas por causa ao longo do tempo e reporte os progressos ao responsável do negócio.

Marcar Chamada de Descoberta

Soluções disponíveis

Tipo de solução	Exemplos	Bom saber
Feedback na aplicação	Polegar para cima ou para baixo, com comentários, na interface de IA	Pouco esforço para os utilizadores; sinal essencial
Filas de revisão e anotação	Funcionalidades de anotação no Langfuse, Arize Phoenix e ferramentas semelhantes	Organizam o trabalho de revisão e ligam-no aos rastreios
Registo de problemas	A sua ferramenta habitual de pedidos ou projectos, com um campo de categoria de falha	Torna as correcções visíveis e com responsável
Reunião de revisão regular	Uma sessão quinzenal de 30 minutos com os responsáveis pelos conteúdos e a equipa de IA	Transforma conclusões em acções com dono

PODE ASSINALAR QUANDO

Há uma revisão regular, as falhas são categorizadas e acompanhadas, as causas mais comuns estão a diminuir e as falhas corrigidas são acrescentadas ao conjunto de teste.

Pôr em prática

Período	Foco
Primeiros 30 dias	Pontue as dez verificações. Conclua a Verificação 1, comece a Verificação 7 com 50 perguntas reais e inicie a limpeza de duplicados e conteúdos desactualizados (Verificação 2).
Dias 31 a 60	Acrescente metadados (Verificação 3), confirme a filtragem por permissões (Verificação 5) e meça a pesquisa separadamente das respostas (Verificação 8).
Dias 61 a 90	Introduza pesquisa híbrida e reordenação (Verificação 4), metas de actualização (Verificação 6), acompanhamento de rapidez e custo (Verificação 9) e a rotina de revisão (Verificação 10). Depois, volte a pontuar todos os trimestres.

Como a Sun Technosystems pode ajudar

Serviço	O que fazemos
Workshop de auto-avaliação	Conduzimos a pontuação com a sua equipa e acordamos prioridades.
Limpeza de conteúdos e metadados	Registos de fontes, remoção de duplicados, regras de retenção e desenho de metadados.
Melhoria da pesquisa	Pesquisa híbrida, reordenação, filtragem por permissões e actualidade, na plataforma que escolher.
Avaliação e monitorização	Conjuntos de teste, métricas de pesquisa e de respostas, painéis de rapidez e custo.
Revisão contínua	Revisões regulares de falhas e ciclos de melhoria com a sua equipa.

Marque uma chamada de descoberta para pontuar a sua situação e acordar as primeiras acções. Nenhum resultado é garantido. Contacte sales@suntechnosystems.net ou +258 85 736 6286.

Marcar Chamada de Descoberta

Notas de publicação e aviso legal

Este guia foi escrito pela Sun Technosystems com base em boas práticas de gestão de informação, pesquisa e IA, sem reproduzir textos, estatísticas, casos de estudo ou gráficos de terceiros. As ferramentas são referidas apenas como exemplos, à data de Novembro de 2026, sem constituir recomendação; confirme funcionalidades, licenças, preços, localização dos dados e segurança antes de escolher. Todos os nomes são marcas dos respectivos proprietários, usados apenas para identificação. Informação de carácter geral, que não constitui aconselhamento jurídico, regulamentar ou de protecção de dados.

© 2026 Sun Technosystems. Todos os direitos reservados sobre o conteúdo original da Sun Technosystems. Os nomes e materiais de terceiros continuam a ser propriedade dos respectivos titulares.

Marcar Chamada de Descoberta