
GUIA DE *DESIGN* DE TRANSPARÊNCIA

para Sistemas de Inteligência Artificial no Setor Público Federal

Versão 1.0

Escopo de aplicação: Órgãos e entidades da administração pública federal direta, autárquica e fundacional

2026

1.1. Objetivo.....	4
1.2. Escopo de aplicação.....	4
1.3. Público-alvo do Guia.....	6
1.4. Relação com outros instrumentos.....	7
2. POR QUE TRANSPARÊNCIA É FUNDAMENTAL.....	9
2.1. Confiança pública.....	9
2.2. Exercício de direitos.....	10
2.3. Responsabilidade institucional.....	11
2.4. Inclusão e diversidade de público.....	11
2.5. Transparência como decisão institucional.....	11
3. CENÁRIOS DE INTERAÇÃO COM SISTEMAS QUE UTILIZAM IA.....	13
3.1. Visão geral.....	13
3.2. Cenário A — Interação direta com as pessoas.....	14
3.3. Cenário B — Geração de conteúdo sintético.....	15
3.4. Cenário C — Deepfakes.....	16
3.5. Cenário D — Texto de interesse público gerado por IA.....	17
3.6. Cenário E — Decisões automatizadas que afetam interesses do titular.....	19
4. CATÁLOGO DE COMPONENTES VISUAIS DE TRANSPARÊNCIA.....	21
4.1. Introdução.....	21
4.2. Componente 1 — Badge persistente.....	21
4.3. Componente 2 — Ícone (tag IA).....	22
4.4. Componente 3 — Banner informativo.....	24
4.5. Componente 4 — Tooltip explicativo.....	26
4.6. Componente 5 — Modal de ciência.....	28
4.7. Componente 6 — Disclaimer textual.....	29
5. RECOMENDAÇÕES PARA SELEÇÃO DE ELEMENTOS DE TRANSPARÊNCIA..	
32	
5.1. Como usar este capítulo.....	32
5.2. Princípios de seleção.....	32
5.3. Cenário A — Interação direta com as pessoas.....	33
5.4. Cenário B — Geração de conteúdo sintético.....	33
5.5. Cenário C — Deepfakes.....	34

5.6. Cenário D — Texto de interesse público gerado por IA.....	35
5.7. Cenário E — Decisões automatizadas que afetam interesses do titular.....	35
5.8. Sobreposição de cenários.....	36
5.9. Matriz consolidada — cenário × risco × componentes.....	36

OBJETIVO E ESCOPO

1.1. Objetivo

Este Guia estabelece recomendações de *design* de interface para a comunicação transparente do uso de inteligência artificial em serviços públicos digitais da Administração Pública Federal. Seu propósito é orientar a escolha e a construção de interfaces que informem o usuário dos sistemas, em especial os cidadãos, de forma clara e acessível, sobre a presença, a natureza e as consequências do uso de inteligência artificial na prestação de serviços públicos e na tomada de decisões administrativas.

O Guia não substitui obrigações legais vigentes, mas busca traduzi-las em recomendações implementáveis nas interfaces dos sistemas de IA do governo. A premissa é que a transparência no setor público não se esgota na publicação de documentos técnicos ou de relatórios de governança: ela exige materialização nas interfaces com os usuários, no ponto em que a decisão ou a interação efetivamente ocorrem.

O Guia também não regula o uso de inteligência artificial no setor público. Não define quais usos são permitidos, restritos ou vedados, nem estabelece critérios de avaliação técnica de modelos. Sua função é estritamente delimitada: uma vez que o órgão tenha decidido adotar tecnologias baseadas em inteligência artificial em determinado serviço digital, o Guia orienta como indicar, nas interfaces, a utilização dessas tecnologias de uma forma clara, acessível e padronizada.

1.2. Escopo de aplicação

1.2.1. Ambientes, canais e interfaces

Este Guia aplica-se aos seguintes ambientes, canais e interfaces digitais de interação entre a administração pública federal e o cidadão:

a. Interfaces *web* de serviços digitais

Portais de serviço, sistemas de consulta pública, plataformas de acompanhamento de processos administrativos e demais interfaces *web* acessíveis aos cidadãos.

b. Interfaces *mobile* de serviços digitais

Aplicativos oficiais, em suas versões para dispositivos Android, iOS ou outro sistema operacional *mobile*, que incorporem funcionalidades que utilizam tecnologias de inteligência artificial.

c. Chatbots e assistentes virtuais

Sistemas conversacionais automatizados em quaisquer canais como, por exemplo, interfaces *web*, aplicativos móveis, *WhatsApp*, *Telegram* ou qualquer outro canal de mensageria adotado oficialmente.

d. Comunicações automatizadas

E-mails, notificações push, mensagens via plataformas oficiais (por exemplo, Gov.br, e-Processo, SEI e assemelhados) e quaisquer outras comunicações geradas ou substancialmente redigidas por sistemas de inteligência artificial e endereçadas ao cidadão.

e. Painéis e relatórios automatizados

Painéis de informação e relatórios gerados, total ou parcialmente, por inteligência artificial, tanto os disponibilizados diretamente ao público quanto os acessados por servidores para fins de análise ou tomada de decisão que afete direitos ou interesses do cidadão.

f. Sistemas internos com impacto externo

Interfaces de uso interno em que o servidor público interage com resultados de sistemas de inteligência artificial que fundamentam, subsidiam ou influenciam decisões sobre direitos, interesses, benefícios ou obrigações do cidadão.

1.2.2. Exclusões expressas

As seguintes matérias não são objeto deste Guia e devem ser endereçadas por instrumentos próprios de governança:

a. Transparência algorítmica e explicabilidade de modelos

A explicação de como um modelo de IA funciona, suas variáveis, pesos, lógica decisória e critérios de treinamento é matéria de transparência algorítmica, distinta da transparência de interface. Este Guia tem como objetivo, exclusivamente, apresentar as formas mais adequadas de como informar o usuário, em especial o cidadão, de que a IA está presente e atuando, não como o modelo chega a seus resultados.

b. Auditoria técnica de sistemas de IA

A avaliação de viés, acurácia, robustez, segurança e desempenho dos modelos de inteligência artificial é matéria de auditoria técnica, fora do escopo deste documento.

c. Regulação do uso de IA no setor público

A definição de quais usos de IA são permitidos, restritos ou vedados na Administração Pública Federal é matéria de política pública, cujos limites são definidos por lei e regulação infralegal. Este Guia não estabelece, sob qualquer perspectiva, definições sobre as formas de uso de IA pela administração pública.

d. Política de dados abertos e transparência ativa

As obrigações de transparência ativa e de disponibilização de dados abertos, previstas na Lei de Acesso à Informação (Lei nº 12.527/2011) e no Decreto nº 8.777/2016, seguem lógica e instrumentos próprios que não se confundem com a transparência de interface de que trata este Guia.

e. Sistemas de IA sem interface humana

Sistemas puramente internos, sem interface de interação com pessoas — como rotinas automatizadas de processamento de dados, automação de infraestrutura de TI e fluxos internos de tratamento de informação — estão fora do escopo. A exceção ocorre quando os resultados desses sistemas alimentam interfaces cobertas pelas categorias anteriores: nesse caso, a interface que apresenta o resultado ao cidadão ou ao servidor está sujeita ao Guia, ainda que o sistema que gerou o dado não esteja.

1.3. Público-alvo do Guia

Este Guia é dirigido especificamente aos perfis institucionais que decidem sobre a adoção de inteligência artificial em serviços públicos digitais e que respondem por sua governança, conformidade e alinhamento com políticas públicas:

a. Gestores de TI e coordenadores de serviços digitais

Responsáveis por classificar as funcionalidades de inteligência artificial dos seus sistemas nos cenários descritos pelo Guia, dimensionar o esforço de adequação e incorporar os elementos de transparência ao ciclo de desenvolvimento e de manutenção dos serviços.

b. Encarregados de dados pessoais (DPOs) e áreas de conformidade

Responsáveis por avaliar se os mecanismos de transparência adotados em cada serviço são compatíveis com as obrigações de proteção de dados pessoais e com a política de governança de inteligência artificial do órgão, validando as escolhas feitas pelas equipes técnicas.

c. Gestores de políticas públicas e autoridades

Responsáveis por compreender o alcance institucional do Guia, seus limites e seu posicionamento em relação aos demais instrumentos de governança do órgão,

assegurando o respaldo necessário para a adoção das recomendações pelas equipes técnicas. Também apoiam a definição de cenários de uso e a avaliação de risco das funcionalidades de inteligência artificial.

1.4. Relação com outros instrumentos

O Guia não opera de forma isolada. Posiciona-se como instrumento operacional dentro de um ecossistema de governança composto por políticas, normas e processos complementares, com os quais mantém relações de subordinação, complementaridade ou conformidade.

a. Política de governança de IA

O Guia é um instrumento subordinado à política de governança de inteligência artificial adotada pelo órgão. As definições de uso permitido, os critérios de avaliação de risco e as instâncias deliberativas decorrem da política, não deste Guia. Havendo conflito entre uma recomendação do Guia e uma determinação da política de governança, prevalecem as determinações da política.

b. Autoavaliação de Impacto Ético da Inteligência Artificial no Setor Público (AIE)

A AIE é um instrumento que auxilia, entre outras dimensões, a identificação da necessidade de mecanismos de transparência para determinado sistema de IA. O Guia fornece os instrumentos concretos, tais como, componentes visuais, recomendações de seleção, especificações para implementar a transparência e deve ser utilizado como meio para atender às recomendações geradas pela AIE. A AIE também explicita o nível de risco da aplicação/funcionalidade, informação determinante para a escolha dos elementos de interface mais adequados para cada aplicação, funcionalidade ou interface.

c. Política de privacidade e proteção de dados pessoais

O Guia endereça um aspecto específico do princípio da transparência previsto no art. 6º, VI, da Lei Geral de Proteção de Dados Pessoais (LGPD): a comunicação ao cidadão sobre o uso de inteligência artificial no tratamento de dados pessoais. A política de privacidade do órgão trata a transparência em sentido amplo (informações sobre tratamento, finalidades, compartilhamento, direitos do titular); o Guia trata a transparência de interface em sentido estrito (como sinalizar visualmente a presença de inteligência artificial). Os dois instrumentos se complementam sem sobreposição.

d. Padrão Digital de Governo (*Design System* do Gov.br)

O Guia opera dentro das restrições impostas pelo Padrão Digital de Governo (*Design System* do Gov.br). Os componentes descritos no Capítulo 4 são padrões conceituais que

devem ser implementados utilizando os recursos do Padrão Digital de Governo e que devem atender aos requisitos do e-MAG. O Guia não substitui, modifica ou excepciona esses padrões, mas conforma-se a eles.

e. Cartilha de IA Generativa

Documento elaborado pela Secretaria de Governo Digital (SGD) e pelo Serviço Federal de Processamento de Dados (Serpro) com orientações para o uso de ferramentas de inteligência artificial generativa, mitigando riscos e maximizando benefícios. A cartilha possui informações relevantes, em especial sobre os riscos envolvidos na utilização dessas ferramentas, que podem ser úteis na aplicação dos critérios para definição dos elementos de interface mais adequados.

2. POR QUE TRANSPARÊNCIA É FUNDAMENTAL

Cada vez mais, a inteligência artificial fará parte do cotidiano dos serviços públicos. Sistemas baseados nessas tecnologias respondem dúvidas de cidadãos, redigem conteúdos informativos, analisam documentos, calculam valores devidos e, em alguns casos, recomendam ou tomam decisões que afetam diretamente a vida das pessoas. À medida que essas aplicações se multiplicam, uma pergunta se impõe: o usuário, em especial o cidadão, sabe que está interagindo com inteligência artificial ou com um conteúdo gerado por inteligência artificial?

Transparência é o compromisso de responder afirmativamente essa pergunta, de forma clara, acessível e oportuna. Não se trata de explicar como os algoritmos funcionam internamente, mas de garantir que o usuário perceba, sem esforço, que uma tecnologia de inteligência artificial está atuando naquele momento, apresentando o contexto e as potenciais consequências da utilização do serviço ou sistema.

Este Guia existe porque, sem a sinalização adequada nas interfaces, a presença da inteligência artificial permanece invisível. E quando permanece invisível, três consequências indesejáveis se produzem: a confiança é comprometida, a capacidade de exercer direitos é reduzida e a responsabilidade institucional se torna difusa.

2.1. Confiança pública

A relação entre governo e cidadão depende de confiança. Quando uma pessoa acessa um serviço público digital ou um sistema governamental, parte da premissa de que a informação que recebe é confiável, de que os processos são conduzidos de forma adequada e de que há responsabilidade institucional por aquilo que é comunicado ou decidido. Essa premissa também orienta o servidor público que utiliza um sistema interno para analisar requerimentos, elaborar pareceres ou subsidiar decisões: ele confia que os dados e as recomendações apresentados pelo sistema são resultado de um processo compreensível e rastreável.

A introdução de inteligência artificial altera pressupostos que ambos os públicos normalmente assumem. Para o cidadão, uma resposta em linguagem natural pode parecer escrita por uma pessoa quando, na verdade, foi produzida por um modelo de linguagem. Um resultado de análise pode parecer que foi avaliado individualmente por um servidor quando, na verdade, foi gerado automaticamente. Para o servidor, um parecer sugerido pelo sistema pode

parecer resultado de consulta a bases normativas consolidadas quando, na verdade, foi produzido por um modelo generativo cujas limitações não estão evidentes na interface. Em ambos os casos, a ausência de sinalização leva a expectativas equivocadas sobre a natureza daquilo que se recebe.

A transparência funciona como mecanismo de preservação da confiança nos dois contextos. Na interface voltada ao cidadão, permite que ele ajuste suas expectativas, exerça maior atenção crítica e mantenha sua confiança no serviço porque compreende o que está acontecendo. Na interface utilizada pelo servidor, permite que ele saiba quando o sistema está apresentando resultados produzidos por inteligência artificial, o que é essencial para que exerça adequadamente sua função de análise e validação, em vez de cancelar automaticamente um resultado cuja origem desconhece.

A confiança não é fragilizada pela presença de inteligência artificial. É fragilizada pela descoberta de que ela estava presente sem que ninguém tenha informado, seja o cidadão que recebe uma decisão, seja o servidor que a fundamentou sem saber que o insumo vinha de um sistema de inteligência artificial.

2.2. Exercício de direitos

Quando um sistema de inteligência artificial participa de uma decisão que afeta o cidadão, por exemplo, na análise de um requerimento de benefício, na classificação de urgência em um serviço de saúde ou na priorização de atendimento, a transparência deixa de ser apenas uma boa prática e passa a ser uma condição para o exercício de direitos.

Para que uma pessoa possa questionar, contestar ou solicitar a revisão de uma decisão, ela precisa primeiro saber que a decisão envolveu processamento por inteligência artificial. Se essa informação não aparece na interface, o cidadão sequer consegue formular o pedido. A transparência de interface é, portanto, o primeiro passo que viabiliza todos os demais: o pedido de explicação, a solicitação de revisão humana, a contestação do resultado.

Esse raciocínio se aplica também a situações menos evidentes. Um texto informativo sobre saúde pública gerado por inteligência artificial, por exemplo, pode influenciar decisões pessoais do cidadão, como vacinar-se ou procurar atendimento médico. Se o cidadão não sabe que o conteúdo foi produzido artificialmente, não consegue avaliar de forma adequada a confiabilidade e os limites daquela informação.

2.3. Responsabilidade institucional

Órgãos públicos são responsáveis pelos serviços que prestam, inclusive quando esses serviços utilizam inteligência artificial. A transparência de interface contribui para que essa responsabilidade se materialize de forma concreta, em vez de permanecer como um princípio abstrato.

Quando a interface identifica claramente qual sistema de inteligência artificial foi utilizado, quando o conteúdo foi gerado, quais dados foram considerados e se houve revisão humana, o órgão demonstra que sabe o que acontece nos seus próprios sistemas e que se responsabiliza pelo resultado entregue ao cidadão. Essa documentação visível também facilita a auditoria, a rastreabilidade e a identificação de eventuais falhas.

A ausência de transparência, por outro lado, não elimina a responsabilidade. Se algo dá errado com um resultado produzido por inteligência artificial e a interface não contém nenhuma indicação de que a tecnologia foi utilizada, tanto o cidadão quanto o próprio órgão terão dificuldade para identificar a origem do problema e corrigi-lo.

2.4. Inclusão e diversidade de público

Os serviços públicos digitais são utilizados por um público extremamente diverso em idade, escolaridade, familiaridade com tecnologia e condições de acessibilidade. O que pode ser evidente para um usuário habituado a interagir com *chatbots* e assistentes virtuais pode não ser para outra pessoa que está acessando um serviço digital pela primeira vez.

A transparência padronizada, com componentes visuais reconhecíveis e textos em linguagem acessível, reduz a assimetria de informação entre diferentes perfis de usuários. Um ícone consistente, um *banner* com texto claro, um aviso em linguagem simples: esses elementos garantem que a informação sobre o uso de inteligência artificial chegue a todos, e não apenas a quem já saberia identificá-la por conta própria.

Essa preocupação é especialmente relevante no setor público brasileiro, onde uma parcela significativa dos cidadãos possui baixo letramento digital. Nesses casos, a transparência nas interfaces não é apenas recomendável, mas uma questão de equidade no acesso à informação.

2.5. Transparência como decisão institucional

A transparência sobre o uso de inteligência artificial não se concretiza apenas por meio de documentos institucionais, políticas de governança ou relatórios de auditoria. Esses instrumentos são necessários, mas não são suficientes. A transparência precisa acontecer no ponto em que o cidadão e o servidor público efetivamente interagem com os serviços e sistemas digitais: a interface.

É na tela do computador, do celular ou do tablet que o usuário deve tomar conhecimento de que está diante de inteligência artificial. É na interface que a transparência se materializa.

Por essa razão, este Guia trata a transparência como uma decisão institucional que envolve, simultaneamente, áreas de gestão, de governança, de *design* e de desenvolvimento. As recomendações apresentadas nos capítulos seguintes são instrumentos concretos para que esses atores, atuando de forma coordenada, possam traduzir o compromisso com a transparência em elementos visíveis, compreensíveis e acessíveis nas interfaces dos sistemas e serviços públicos digitais.

3. CENÁRIOS DE INTERAÇÃO COM SISTEMAS QUE UTILIZAM IA

3.1. Visão geral

Uma pessoa que utiliza um serviço público digital pode interagir com um sistema que utiliza inteligência artificial de diversas formas. Pode conversar com um *chatbot*, pode ler um texto informativo que foi total ou parcialmente escrito por máquina, utilizando um modelo de linguagem e pode, inclusive, receber o resultado de uma solicitação que foi analisada por um sistema de inteligência artificial sem interação humana prévia.

Como um dos pilares que sustentam o uso responsável e ético da inteligência artificial é a transparência, em todos esses casos o usuário precisa ser avisado de que está interagindo com um elemento de inteligência artificial ou de que o resultado a que teve acesso foi obtido com o uso dessas tecnologias. E, para cada situação, pode ser necessária a utilização de um tipo diferente de elemento de interface que esclareça, com grau razoável de exatidão, se e como essas tecnologias são utilizadas.

Para que as equipes responsáveis por projetar e desenvolver sistemas consigam identificar quais elementos visuais devem ser utilizados em cada circunstância, este capítulo descreve cinco cenários que representam as formas pelas quais as pessoas geralmente interagem com sistemas que utilizam inteligência artificial — ou são afetadas por seus resultados — nos serviços públicos digitais.

Para fins deste Guia, "usuário" designa qualquer pessoa que interaja com uma interface digital, seja o cidadão que acessa um serviço público, seja o servidor que opera um sistema interno.

Três regras orientam o uso dos cenários:

- i. **Os cenários podem se sobrepor.** Uma mesma interação pode se enquadrar em mais de um cenário ao mesmo tempo. Um *chatbot* que conversa com o usuário, gera textos informativos e avalia elegibilidade a um determinado benefício pode se enquadrar nos Cenários A, D e E simultaneamente.
- ii. **O cenário se faz pela funcionalidade, não pela tecnologia.** Um mesmo modelo de linguagem pode ser utilizado para implementar um *chatbot* de orientação em um serviço e para gerar relatórios em outro. O que determina o cenário não é

como o sistema funciona internamente, mas o que ele entrega (funcionalidades), as informações que apresenta ao usuário e como se dá a interação do usuário com o sistema.

- iii. **Na dúvida, escolha o cenário mais exigente.** Classificar um sistema em um cenário de menor exigência, quando deveria estar em cenário de maior exigência, resulta em transparência insuficiente. Por isso, é sempre recomendável que as decisões sobre os elementos de interface partam dos cenários mais exigentes.

3.2. Cenário A — Interação direta com as pessoas

3.2.1. Descrição

O usuário se comunica diretamente com um sistema de IA, seja por mensagem de texto, comandos de voz ou outra forma de diálogo, e recebe respostas em tempo real. A dinâmica se assemelha a uma conversa com um humano.

O que define este cenário é a **interlocução**: há troca direta entre pessoa e máquina. A preocupação é deixar claro que a pessoa está interagindo com um sistema que utiliza inteligência artificial para emular uma conversa.

Este cenário não inclui situações em que a IA apenas processa dados em segundo plano, sem dialogar com o usuário. Também não inclui casos em que a inteligência artificial gera conteúdo consumido pelo usuário sem troca interativa. Esses casos pertencem ao Cenário B ou ao Cenário D, conforme a natureza do conteúdo.

3.2.2. Exemplos no setor público

A interação direta entre uma pessoa e um sistema que utiliza inteligência artificial, no âmbito dos serviços públicos digitais, pode se dar de diversas formas como, por exemplo:

- *chatbot* que responde perguntas sobre documentos, prazos e procedimentos;
- assistente virtual em aplicativo oficial que ajuda o usuário a localizar serviços públicos;
- assistente (bot) em canal de mensageria institucional, *WhatsApp*, por exemplo, que informa o andamento de solicitações;
- assistente virtual em sistema interno que auxilia servidores a consultar normas ou procedimentos administrativos.

3.2.3. Nota sobre a exceção de obviedade

Em algumas situações, a natureza artificial do interlocutor pode ser considerada evidente, o que poderia levar à conclusão de que seria desnecessária a utilização de elementos de interface que explicitam a utilização de IA. Recomenda-se fortemente que essa exceção não seja utilizada em serviços públicos. Os usuários dos serviços digitais do governo são heterogêneos em idade, escolaridade e familiaridade com tecnologia. O que parece evidente para um usuário habituado a interfaces digitais pode não ser para outro. A exceção somente deve ser aplicada mediante deliberação expressa e documentada do comitê de governança de IA do órgão, quando essa estrutura de governança existir, com justificativa que considere o perfil do público-alvo do serviço.

3.3. Cenário B — Geração de conteúdo sintético

3.3.1. Descrição

O usuário consome conteúdo — texto, imagem, áudio ou vídeo — que foi total ou parcialmente gerado por um sistema de IA, sem que haja necessariamente diálogo entre o usuário e o sistema.

O que define este cenário é a autoria artificial: aquilo que o usuário lê, vê ou ouve foi produzido, no todo ou em parte significativa, por um sistema que utiliza tecnologia de inteligência artificial. A preocupação é que o usuário atribua ao conteúdo o mesmo grau de curadoria e responsabilidade que atribuiria a um material produzido integralmente por um ser humano.

Este cenário **não** inclui:

- conteúdo produzido por pessoas, mesmo que publicado em plataforma que use IA em outros contextos;
- conteúdo que recebeu apenas revisão ortográfica ou gramatical automatizada, sem alteração de sentido ou estrutura;
- conteúdo cujo único envolvimento de IA foi a diagramação, formatação ou distribuição.

3.3.2. Exemplos no setor público

- Descrições de serviços públicos redigidas ou reescritas por IA;
- resumos automatizados de documentos oficiais como, por exemplo, portarias, editais, pareceres;
- traduções automatizadas de conteúdo institucional;

- imagens geradas por IA para campanhas de comunicação governamental;
- transcrições automatizadas de audiências públicas ou sessões de órgãos colegiados;
- relatórios gerenciais internos cujos textos analíticos são produzidos por IA para uso de servidores.

3.3.3. Fronteiras com outros cenários

O Cenário B é o mais amplo dos cenários relativos a conteúdo. Os Cenários C e D são mais específicos: o Cenário C trata de conteúdo audiovisual (imagens, vídeos ou áudios), criado ou manipulado por IA, para parecerem realistas (*deepfakes*) e o Cenário D trata de texto com finalidade informativa sobre assuntos de interesse público. Todos os conteúdos dos Cenários C ou D são também conteúdos sintéticos no sentido do Cenário B, mas possuem peculiaridades que os diferenciam e que trazem exigências adicionais em termos de transparência nas interfaces.

Na prática, o Cenário B funciona como categoria geral: aplica-se sempre que o conteúdo gerado por IA não atender aos critérios mais específicos dos Cenários C ou D.

3.4. Cenário C — *Deepfakes*

3.4.1. Descrição

O usuário é exposto a conteúdo em formato de imagem, áudio ou vídeo gerado ou manipulado por IA de modo a se assemelhar a pessoas, objetos, lugares ou eventos reais, com grau de fidelidade suficiente para ser confundido com registro autêntico por uma pessoa comum.

O que define este cenário é a verossimilhança enganosa. Não basta que o conteúdo seja sintético (o que bastaria para o Cenário B): é necessário que simule a realidade de forma convincente. A preocupação é que o usuário acredite estar diante de registro real quando o material foi gerado ou substancialmente manipulado por IA.

Este cenário é restrito a conteúdo audiovisual. Texto é tratado nos Cenários B ou D, independentemente do grau de verossimilhança.

Este cenário **não** inclui:

- conteúdo audiovisual claramente estilizado ou abstrato, sem capacidade de simular a realidade (ilustrações, infográficos, ícones gerados por IA);

- conteúdo audiovisual autêntico que recebeu apenas edição técnica automatizada (ajuste de cor, estabilização, compressão), sem alteração do conteúdo representado.

3.4.2. Exemplos no setor público

- Vídeo institucional com avatar digital de aparência fotorrealista representando porta-voz de órgão;
- imagens fotorrealistas geradas por IA para projeções de obras públicas ou simulações urbanísticas;
- vídeo de capacitação de servidores com instrutor virtual de aparência humana realista;
- reconstrução visual de eventos históricos com imagens fotorrealistas para material educativo.

3.4.3. Fronteiras com outros cenários

A fronteira entre os Cenários B e C está no grau de verossimilhança. Uma ilustração estilizada gerada por IA para campanha institucional, reconhecível como produção artificial, enquadra-se no Cenário B. Um vídeo com avatar digital que reproduz fielmente a aparência e a voz de uma pessoa, que age como um apresentador em um vídeo, enquadra-se no Cenário C.

Critério orientador: uma pessoa comum, sem aviso prévio, poderia razoavelmente confundir o conteúdo gerado ou manipulado por IA com um registro autêntico, ou seja, um registro real? Se sim, é o Cenário C.

3.4.4. Nota sobre o uso legítimo de *deepfakes* no setor público

O setor público pode utilizar técnicas de *deepfake* para finalidades legítimas, tais como acessibilidade, educação, treinamento, simulação. O enquadramento no Cenário C não implica vedação do uso, mas sim que a transparência deve ser proporcionalmente reforçada. Mesmo quando a finalidade é institucional, o usuário tem o direito de saber que o conteúdo foi gerado artificialmente.

A marcação deve ser aplicada inclusive em contextos de uso claramente institucional e educativo, por duas razões: respeito ao princípio da confiança legítima e prevenção contra reutilização do material fora de seu contexto original.

3.5. Cenário D — Texto de interesse público gerado por IA

3.5.1. Descrição

O usuário acessa texto gerado ou substancialmente assistido por IA, publicado com a finalidade de informar sobre assuntos de interesse coletivo.

Três características, quando observadas conjuntamente, definem este cenário: (i) o formato é textual (não audiovisual); (ii) a finalidade é informativa (o texto comunica fatos, dados, orientações ou análises) e (iii) o tema é de interesse público (saúde, educação, segurança, economia, direitos, políticas públicas e assemelhados).

A preocupação é que o usuário forme opinião ou tome decisão com base em informação de relevância coletiva sem saber que a origem é artificial, o que pode comprometer sua avaliação sobre a confiabilidade e a responsabilidade relacionada ao conteúdo.

Este cenário **não** inclui:

- texto gerado por IA para uso operacional interno, sem publicação dirigida a público externo ou interno amplo;
- texto de natureza promocional ou publicitária, que pode se enquadrar no Cenário B;
- conteúdo audiovisual sobre assuntos de interesse público, que se enquadra no Cenário B ou, se houver verossimilhança enganosa, no Cenário C.

3.5.2. Exemplos no setor público

- Boletins informativos automatizados como, por exemplo, boletins epidemiológicos, de indicadores econômicos, de segurança pública;
- informes e notas oficiais redigidos com assistência substancial de IA;
- respostas automatizadas a pedidos de acesso à informação;
- conteúdo de seções de perguntas frequentes (FAQ) gerado ou atualizado por IA em portais de serviço;
- resumos de políticas públicas ou programas governamentais produzidos por IA para publicação em portal institucional.

3.5.3. Fronteiras com outros cenários

Os Cenários B e D diferem em duas dimensões. Na modalidade: o Cenário D é restrito a texto, enquanto o Cenário B abrange qualquer formato. Na finalidade: o Cenário D exige finalidade informativa de interesse público, enquanto o Cenário B não impõe esse requisito.

Uma descrição operacional de serviço gerada por IA pode se enquadrar apenas no Cenário B. Um boletim epidemiológico automatizado enquadra-se no Cenário D.

O Cenário D existe como categoria distinta porque textos informativos de interesse público exigem transparência mais rigorosa do que conteúdo sintético genérico. Quem forma opinião ou toma decisão com base em informação pública tem interesse qualificado em saber se essa informação foi produzida por uma máquina.

3.6. Cenário E — Decisões automatizadas que afetam interesses do titular

3.6.1. Descrição

Um sistema de IA toma, recomenda, fundamenta ou influencia de forma preponderante uma decisão que afeta diretamente direitos, interesses, benefícios, obrigações ou o acesso de uma pessoa a serviços públicos.

O que define este cenário é o impacto decisório. Diferentemente dos demais, que tratam de diálogo, conteúdo ou informação, aqui a IA decide ou subsidia decisão sobre algo que importa concretamente para a vida de alguém: a concessão ou negativa de um benefício, a posição em fila de atendimento, o valor de uma obrigação, a elegibilidade para um programa.

A pessoa afetada pode ser o próprio usuário da interface (o cidadão que preenche formulário e recebe decisão automatizada) ou um terceiro que não interage diretamente com o sistema (o cidadão cujo requerimento é analisado por IA em sistema operado por servidor). Em ambos os casos, o cenário se aplica.

Dois esclarecimentos são necessários para tornar mais clara a identificação deste cenário. O primeiro refere-se ao conceito de decisão automatizada. No contexto deste Guia, considera-se decisão automatizada não apenas aquela tomada inteiramente por máquina, mas também aquela em que o resultado do sistema de IA é o elemento preponderante da decisão, ainda que haja validação formal de um servidor. A mera chancela humana sem análise substantiva não afasta a aplicação deste cenário.

O segundo refere-se ao conceito de afetação de interesses. O cenário não se restringe a decisões com efeito jurídico formal (deferimento, indeferimento, sanção). Abrange também decisões com efeito prático: priorização em ordem de atendimento, classificação em faixa de risco, ordenamento em processo seletivo, cálculo de valores.

3.6.2. Exemplos no setor público

- Verificação automática de requisitos para inscrição em programa de transferência de renda;
- classificação algorítmica de urgência que determina a ordem de atendimento em serviços de saúde;
- triagem automatizada de requerimentos administrativos com recomendação de deferimento ou indeferimento;
- análise automatizada de documentos, com elaboração de recomendação feita por IA, em processos de fiscalização;
- cálculo automatizado de valores devidos ao cidadão ou por ele devidos ao poder público;
- ordenamento de candidatos em processos seletivos com base em critérios processados por IA.

3.6.3. Fronteiras com outros cenários

A fronteira entre os Cenários A e E é a que exige maior atenção, porque muitos sistemas combinam diálogo com função decisória. A distinção está na função: se o sistema conversa, orienta e direciona, trata-se do Cenário A. Agora, se há classificação, priorização, avaliação de elegibilidade ou recomendação de decisão que afeta a pessoa, trata-se do Cenário E, sem prejuízo de que a interação conversacional permaneça enquadrada no Cenário A.

Um assistente virtual em serviço de saúde que apenas informa horários e orienta sobre agendamento exerce função conversacional, portanto, enquadra-se exclusivamente no Cenário A. Se esse mesmo assistente solicita informações sobre sintomas e, com base nelas, classifica a urgência do caso para definir a prioridade de atendimento, essa segunda função é decisória, enquadrando-se, portanto, no Cenário E, independentemente de a coleta de dados ter ocorrido por meio de diálogo.

4. CATÁLOGO DE COMPONENTES VISUAIS DE TRANSPARÊNCIA

4.1. Introdução

Este capítulo apresenta o repertório de componentes visuais disponíveis para comunicar ao usuário a presença, o papel e as limitações de sistemas de inteligência artificial em interfaces de serviços públicos digitais.

Os componentes descritos a seguir são elementos conceituais e cada um deles deverá ser implementado, na prática, utilizando-se os padrões e requisitos determinados no Padrão Digital de Governo (*Design System* do Gov.br).

4.2. Componente 1 — *Badge persistente*

4.2.1. Descrição e finalidade

O *badge persistente* é um elemento visual compacto que permanece visível durante toda a sessão de interação do usuário com um sistema de IA conversacional. Posiciona-se na base ou no topo da área de conversa e acompanha a navegação, sem desaparecer ao rolar o conteúdo.

Sua finalidade é informar de forma contínua e não intrusiva que o interlocutor do usuário é um sistema de inteligência artificial, não uma pessoa. O *badge* não interrompe o fluxo da interação nem exige ação do usuário — apenas garante que a informação sobre a natureza artificial do interlocutor esteja sempre acessível.

4.2.2. Quando utilizar

Em interfaces conversacionais (Cenário A) em qualquer nível de risco, como *chatbots* de atendimento, assistentes virtuais e *bots* em canais de mensageria, neste último caso quando tecnicamente viável.

Pode ser utilizado, também, como complemento a outros componentes em cenários de risco moderado e alto, garantindo que a informação permaneça visível mesmo após o usuário interagir com modais ou *banners*.

4.2.3. Quando não utilizar

Não substitui o modal de ciência em cenários de risco alto. O *badge* informa continuamente, mas não exige confirmação ativa do usuário.

Não é adequado como único mecanismo de transparência para conteúdo estático (páginas, relatórios), em que não há sessão conversacional contínua.

4.2.4. Exemplo de aplicação

A imagem a seguir ilustra a aplicação do badge persistente em um assistente virtual em interface *mobile*. O elemento "Atendimento por IA", fixo na base da tela, mantém o usuário ciente, durante toda a interação, de que está conversando com um sistema de inteligência artificial.



4.3. Componente 2 — Ícone (*tag IA*)

4.3.1. Descrição e finalidade

O ícone de IA é um elemento gráfico compacto posicionado junto ao conteúdo gerado por inteligência artificial. Pode ser acompanhado de rótulo curto (ex.: "Gerado por IA").

Sua finalidade é sinalizar pontualmente que um trecho específico de conteúdo foi produzido ou substancialmente assistido por IA, sem demandar espaço significativo na interface. É o componente de menor intrusividade do catálogo.

O ícone possui duas variantes principais:

a. Variante "Gerado por IA"

Indica conteúdo integralmente produzido por inteligência artificial, sem revisão ou contribuição humana substancial.

b. Variante "Assistido por IA"

Indica conteúdo com participação humana e de inteligência artificial — por exemplo, texto produzido por inteligência artificial e revisado por servidor antes da publicação.

A distinção entre essas variantes é fundamental para a transparência: permite ao usuário compreender, de forma rápida, o grau de envolvimento humano no conteúdo apresentado.

O ícone de IA não representa uma sugestão de aplicação direta de elemento gráfico. Trata-se de um referencial conceitual: o elemento visual deve tornar claro, para qualquer pessoa, que aquela interface possui interação com inteligência artificial ou recebe conteúdo gerado por essa tecnologia. A associação entre o elemento visual e o contexto precisa ser direta.

4.3.2. Quando utilizar

Em conteúdo gerado ou assistido por inteligência artificial (Cenário B) em nível de risco baixo a moderado: descrições, resumos, traduções, imagens estilizadas. Pode ser usado também como marcação pontual de trechos específicos dentro de uma página que combina conteúdo humano e conteúdo gerado por inteligência artificial.

O ícone de IA pode ser inserido em outros componentes (*badges*, *banners*, modais, *disclaimers*), criando uma identidade visual persistente que se mantém em todas as circunstâncias em que o usuário interage ou recebe conteúdo gerado, total ou parcialmente, por inteligência artificial.

4.3.3. Quando não utilizar

O ícone é insuficiente como único mecanismo em cenários de risco moderado e de risco alto. Nessas situações, deve ser utilizado de forma complementar, junto a outros componentes do catálogo.

Em interfaces conversacionais, o ícone pode ser integrado ao *badge* persistente, pois este é o elemento mais indicado para esse cenário.

O ícone não deve ser utilizado sem rótulo textual em contextos em que o ícone isolado pode não ser compreendido pelo público-alvo do serviço.

4.3.4. Exemplo de aplicação

A imagem a seguir apresenta as duas variantes do ícone de IA — "Gerado por IA" (preenchido) e "Assistido por IA" (apenas com borda) — em diversos tamanhos e em uso integrado ao texto. A diferença visual entre as duas variantes permite ao usuário distinguir, à primeira vista, conteúdo integralmente artificial de conteúdo com colaboração humana.



4.4. Componente 3 — *Banner* informativo

4.4.1. Descrição e finalidade

O *banner* informativo é uma faixa destacada posicionada acima do conteúdo principal, com cor de fundo diferenciada e borda lateral de destaque. Contém texto explicativo sobre o uso de IA no conteúdo que se segue e pode incluir link para informações adicionais.

Sua finalidade é alertar o usuário, antes da leitura do conteúdo, de que aquele material foi gerado ou substancialmente assistido por inteligência artificial. É mais visível do que o ícone, mas não exige interação ativa do usuário, diferenciando-se do modal, que requer confirmação.

O *banner* possui, em geral, duas variantes:

a. Variante informativa

Aplicável quando o conteúdo gerado por inteligência artificial não representa risco a direitos e garantias dos cidadãos ou quando há garantia de que as informações são precisas ou foram previamente revisadas.

b. Variante de alerta

Adiciona um elemento de atenção e cor mais saliente. Deve ser utilizada quando o conteúdo gerado por inteligência artificial pode conter imprecisões ou limitações que o usuário deve considerar antes de tomar decisões.

4.4.2. Quando utilizar

O *banner* informativo é o elemento ideal para ser utilizado quando há conteúdo gerado por IA (Cenário B) em nível de risco baixo ou moderado. Também é o componente adequado para comunicações automatizadas dirigidas ao usuário (e-mails, notificações), bem como para os casos que envolvam texto de interesse público gerado por IA (Cenário D) e, também, em páginas cujo conteúdo principal foi gerado ou substancialmente editado por IA.

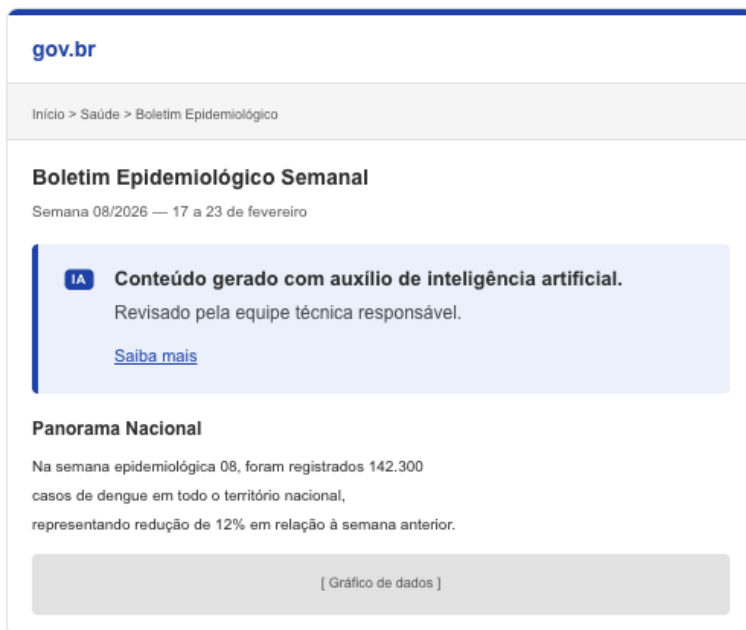
4.4.3. Quando não utilizar

O *banner* não pode ser o único elemento utilizado em cenários onde é necessária confirmação ativa (risco alto). Nestes casos, o *banner* deve ser complementado pelo modal de ciência. Também não é indicado para interfaces conversacionais, onde o *badge* persistente cumpre a função de forma mais integrada.

4.4.4. Exemplo de aplicação

A imagem a seguir mostra um *banner* informativo no topo de uma página de boletim epidemiológico, comunicando ao usuário que o conteúdo foi gerado com auxílio de inteligência

artificial e revisado pela equipe técnica responsável. O *banner* é claramente perceptível antes da leitura do conteúdo, sem bloquear o acesso à informação.



4.5. Componente 4 — *Tooltip* explicativo

4.5.1. Descrição e finalidade

O *tooltip* explicativo é uma caixa flutuante de informação que aparece quando o usuário interage com um elemento gatilho (passar o cursor, tocar em telas sensíveis ao toque ou navegar pelo teclado). Contém explicação breve sobre o uso de IA naquele ponto específico da interface.

Sua finalidade é fornecer informação adicional sob demanda, sem ocupar espaço permanente na interface. O *tooltip* complementa outros componentes — aparece, por exemplo, ao interagir com um ícone que identifica a existência de informações adicionais disponíveis.

4.5.2. Quando utilizar

Como camada adicional de informação em qualquer cenário, acionada por interação do usuário com ícones, badges ou outros indicadores. É especialmente indicado para

recomendações automatizadas e conteúdo personalizado em risco baixo ou moderado, bem como para explicação pontual sobre a origem de um dado específico em painéis e relatórios.

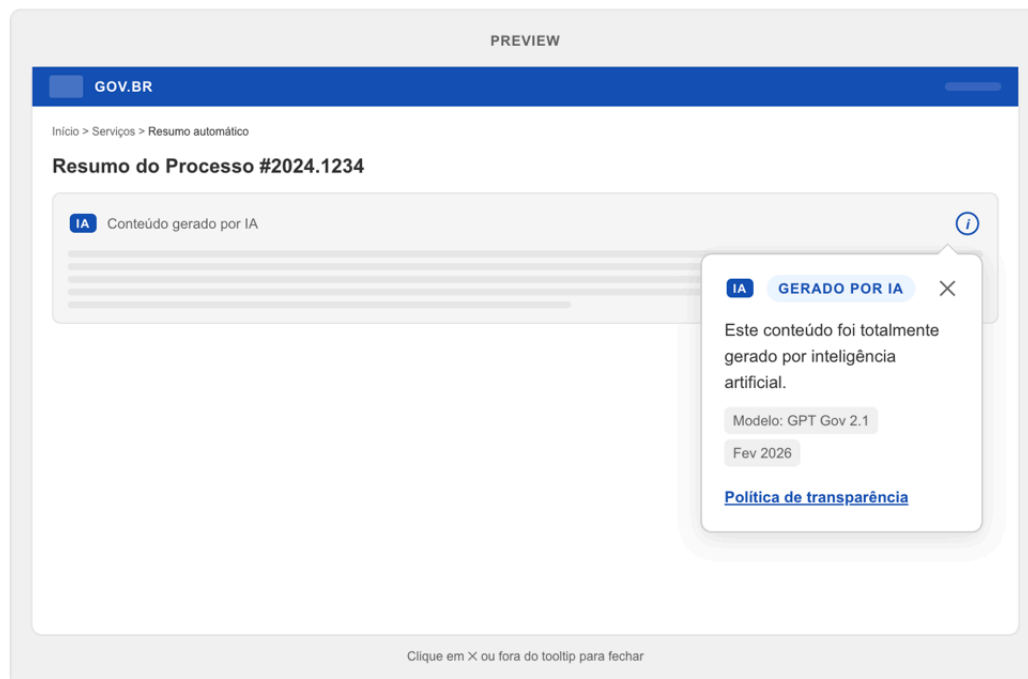
4.5.3. Quando não utilizar

Nunca deve ser o único mecanismo de transparência em cenários de risco alto ou em decisões automatizadas (Cenário E), uma vez que a informação depende de ação voluntária do usuário para ser exibida.

Não é adequado em aplicações móveis como mecanismo primário de informação, pois o gesto de passar o cursor não existe em telas de toque. Para esses casos, deve haver alternativas de toque ou outro componente visível.

4.5.4. Exemplo de aplicação

A imagem a seguir mostra um *tooltip* ativado a partir de um ícone de informação em uma página de resumo automático de processo. O *tooltip* apresenta detalhes adicionais sobre a origem do conteúdo (sistema utilizado, data) e oferece um link para a política de transparência do órgão.



4.6. Componente 5 — Modal de ciência

4.6.1. Descrição e finalidade

O modal de ciência é uma janela sobreposta à interface (*overlay*) que aparece antes de o usuário iniciar a interação com um sistema de IA de alto risco. Exige que o usuário leia as informações apresentadas e confirme ativamente que as compreendeu, em geral, por meio de marcação de uma caixa de seleção (*checkbox*) e do acionamento de um botão de confirmação antes de prosseguir.

Sua finalidade é garantir ciência informada: assegurar que o usuário saiba, antes de interagir com o sistema, que está diante de IA, quais são as implicações e quais são seus direitos, notadamente o direito à revisão humana. Quando o contexto exige, o modal pode gerar registro auditável da confirmação.

4.6.2. Quando utilizar

Em cenários de risco alto em qualquer categoria, especialmente interação direta (Cenário A) e decisões automatizadas (Cenário E). Também é o elemento indicado para o primeiro acesso a sistemas de IA que realizam classificação, triagem ou avaliação de elegibilidade. Sistemas cujas decisões afetam diretamente direitos, benefícios ou oportunidades de um cidadão, seja com interação direta, seja na interação de um servidor público com um sistema interno, devem utilizar sempre o modal de ciência.

4.6.3. Quando não utilizar

Não deve ser utilizado em cenários de risco baixo, pois gera fadiga de consentimento e cria barreira desnecessária de acesso ao serviço.

Também não deve ser exibido repetidamente ao mesmo usuário no mesmo serviço. Após a primeira confirmação, o sistema pode registrar a ciência e não exibir novamente o modal na mesma sessão (ou no mesmo dispositivo, a depender do contexto).

4.6.4. Exemplo de aplicação

A imagem a seguir mostra um modal de ciência aplicado a uma análise automatizada por inteligência artificial. O usuário recebe, em destaque, informações sobre o que acontece no processo, seus direitos (incluindo o direito à revisão humana, conforme a LGPD) e os dados utilizados. A confirmação ativa, por meio de caixa de seleção, é necessária para prosseguir, e gera registro auditável.

gov.br

IA

Análise Automatizada por IA

Antes de prosseguir, leia as informações abaixo.

O que acontece: Sua solicitação será analisada por um sistema de inteligência artificial com base em dados financeiros e cadastrais.

Seus direitos: Você pode solicitar revisão humana da decisão, conforme a LGPD (art. 20).

Dados utilizados: Histórico de crédito, renda declarada, dados cadastrais. [Política de uso de IA](#)

Li e compreendo que esta análise será realizada por sistema de IA.

△ Confirmação gera registro auditável (timestamp + ID)

4.7. Componente 6 — *Disclaimer* textual

4.7.1. Descrição e finalidade

O *disclaimer* textual é um bloco de texto posicionado ao final de uma página, relatório, resultado de consulta ou comunicação, com informações detalhadas sobre o uso de inteligência artificial na produção daquele conteúdo ou decisão.

Sua finalidade é fornecer informação completa, em linguagem acessível, sobre a natureza do envolvimento da IA, as limitações conhecidas do sistema, a existência (ou não) de revisão humana, e os canais disponíveis para contestação ou solicitação de revisão. É o componente com maior densidade informativa do catálogo.

O *disclaimer* pode assumir três modelos, conforme o nível de risco:

a. Modelo curto (risco baixo)

Apenas informa que o conteúdo foi gerado com auxílio de inteligência artificial e indica a equipe responsável.

b. Modelo padrão (risco moderado)

Acrescenta informações estruturadas: tipo de conteúdo, nome do sistema, data de geração, fonte dos dados e equipe ou coordenação responsável pela revisão.

c. Modelo detalhado (risco alto)

Inclui, além das informações anteriores, os critérios utilizados, as limitações conhecidas, o status de revisão humana, o responsável institucional e o canal disponível para contestação ou solicitação de revisão.

4.7.2. Quando utilizar

Deve ser utilizado quando houver decisões automatizadas (Cenário E) em nível moderado a alto, após a apresentação de resultados de análise, simulações e cálculos automatizados. Também deve ser utilizado sempre que um relatório ou documento, criado total ou parcialmente por IA, puder gerar consequências que limitem direitos e garantias do usuário ou de terceiros.

O *disclaimer* pode ser utilizado como componente complementar em qualquer cenário de risco alto, fornecendo o detalhamento que *badges*, *banners* e modais não comportam.

4.7.3. Quando não utilizar

O *disclaimer* textual não deve ser utilizado como mecanismo primário de aviso em interfaces conversacionais. O *badge* persistente é mais indicado. Também não deve substituir o modal de ciência quando confirmação ativa é necessária, já que o *disclaimer* informa, mas não exige confirmação.

4.7.4. Exemplo de aplicação

A imagem a seguir mostra um *disclaimer* detalhado aplicado ao resultado de uma análise de Benefício de Prestação Continuada (BPC). O bloco apresenta, de forma estruturada, o tipo de conteúdo, o sistema utilizado, a data, os critérios, as limitações conhecidas, o status de revisão humana, o responsável institucional e os canais para solicitação de revisão e contestação.

gov.br

Início / Meus benefícios / Resultado da análise – BPC

Resultado da análise – BPC

RESULTADO DA ANÁLISE

Benefício de Prestação Continuada (BPC)

Protocolo: 2026.0303.1847 Data: 03/03/2026

✓ Critérios iniciais atendidos – aguardando perícia

IA Informação de transparência

Conteúdo: Análise de requerimento de BPC

Sistema: PrevidênciaA v2.1

Data: 03 mar. 2026

Crítérios: Renda per capita, idade, CID, inscrição CadÚnico

Limitações: Não considera atividade especial, tempo rural ou acordos internacionais

Revisão: Não passou por revisão humana

Responsável: Coord. de Inteligência Artificial — INSS

[Solicitar revisão humana da decisão →](#)

[Acessar canal de contestação →](#)

[Acessar política de uso de inteligência artificial →](#)

5. RECOMENDAÇÕES PARA SELEÇÃO DE ELEMENTOS DE TRANSPARÊNCIA

5.1. Como usar este capítulo

Este capítulo orienta as equipes responsáveis pelos serviços digitais na escolha dos componentes visuais de transparência adequados a cada situação. A lógica de seleção segue três etapas:

- i. **Identificar o cenário de interação** (Capítulo 3). Uma mesma interface pode combinar mais de um cenário.
- ii. **Avaliar o nível de risco** da funcionalidade baseada em inteligência artificial (preferencialmente com a aplicação da Autoavaliação de Impacto Ético - AIE). Na dúvida, adotar o nível mais exigente.
- iii. **Consultar a combinação recomendada** na tabela do cenário correspondente (subseções 5.3 a 5.7) e aplicar os componentes indicados conforme as descrições do Capítulo 4.

Quando mais de um cenário se aplica simultaneamente, prevalece a combinação mais exigente entre os cenários identificados, conforme orientado na subseção 5.8.

5.2. Princípios de seleção

Três princípios orientam a composição dos elementos de transparência em cada interface:

Decisão baseada no nível de risco

A quantidade e a intrusão dos componentes devem ser proporcionais ao nível de risco. Interfaces relacionadas a funcionalidades de risco baixo utilizam menos componentes e de menor intrusão; interfaces para funcionalidades de risco alto exigem mais camadas de informação e mecanismos de ciência informada.

Complementaridade

Os componentes se complementam e não se substituem. Cada componente cumpre uma função distinta e, portanto, podem ser utilizados conjuntamente em uma mesma interface, sempre que o contexto exigir a aplicação de mais de um componente.

Não redundância

Componentes diferentes comunicam informações diferentes. O mesmo texto não deve ser repetido em múltiplos componentes. Se o *banner* informativo já informa que o conteúdo foi gerado por IA, o *tooltip* explicativo deve acrescentar informação complementar, e não repetir a mesma frase.

5.3. Cenário A — Interação direta com as pessoas

Neste cenário, o usuário dialoga com um sistema de IA em tempo real. A preocupação central é garantir que o usuário saiba, de forma contínua, que seu interlocutor é uma máquina.

Nível de risco	Componentes obrigatórios	Componentes recomendados
Baixo	Ícone de IA em cada mensagem do sistema + Badge Persistente	
Moderado	Ícone de IA em cada mensagem + Badge Persistente + Banner Informativo no início da sessão	Tooltip Explicativo com detalhes do sistema + Disclaimer Textual ao final da conversa.
Alto	Ícone de IA em cada mensagem + Badge Persistente + Banner Informativo (tipo alerta) ou Modal de Ciência no primeiro acesso, quando o conteúdo subsidia decisão do usuário que possa interferir negativamente em direitos e garantias dos cidadãos.	Tooltip Explicativo + Disclaimer Textual ao final, com metadados e canal de contestação.

No Cenário A, o *Badge Persistente* e o Ícone de IA são os componentes estruturantes: garantem que a informação sobre a natureza artificial do interlocutor esteja presente durante toda a interação, sem depender de ação do usuário.

5.4. Cenário B — Geração de conteúdo sintético

Neste cenário, o usuário consome conteúdo gerado ou substancialmente assistido por IA. A preocupação é evitar que o usuário atribua ao conteúdo o mesmo grau de curadoria humana de um material integralmente produzido por pessoas.

Nível de risco	Componentes obrigatórios	Componentes recomendados
Baixo	Ícone de IA junto ao conteúdo + Disclaimer Textual (modelo curto)	Tooltip Explicativo junto ao ícone de informação.
Moderado	Ícone de IA + Disclaimer Textual (modelo padrão)	Banner Informativo no topo do conteúdo Tooltip Explicativo acionável a partir do ícone de informação.
Alto	Ícone de IA + Banner Informativo (tipo alerta) + Disclaimer Textual (modelo detalhado) + Modal de Ciência no primeiro acesso, quando o conteúdo subsidia decisão do usuário que possa interferir negativamente em direitos e garantias dos cidadãos.	Tooltip Explicativo acionável a partir do ícone de informação.

No Cenário B, o Ícone de IA e o **Disclaimer Textual** são os componentes estruturantes: garantem que a natureza sintética do conteúdo seja identificada para o usuário e adicionam detalhes relevantes a respeito do conteúdo. O Modal de Ciência também é um elemento obrigatório nas circunstâncias onde o conteúdo subsidia decisões do usuário que possam afetar, negativamente, a esfera de direitos e garantias dos cidadãos.

5.5. Cenário C — Deepfakes

Neste cenário, o usuário é exposto a conteúdo audiovisual com verossimilhança enganosa. A preocupação é evitar que o usuário confunda material gerado por IA com registro autêntico.

Nível de risco	Componentes obrigatórios	Componentes recomendados
Moderado	Ícone de IA sobreposto ao conteúdo + Disclaimer Textual (modelo curto) sobreposto ao conteúdo.	Tooltip Explicativo , acionável a partir do ícone de informação, com descrição da técnica utilizada. Banner Informativo apresentado fora do conteúdo, preferencialmente em posição imediatamente posterior ao conteúdo.
Alto	Ícone de IA sobreposto + Disclaimer Textual (modelo curto) sobreposto ao conteúdo + Modal de Ciência , previamente ao acesso, quando o conteúdo subsidia decisão do usuário que possa interferir negativamente em direitos e garantias dos cidadãos.	Tooltip Explicativo , acionável a partir do ícone de informação, com descrição da técnica utilizada. Banner Informativo (tipo perigo) apresentado fora do conteúdo, preferencialmente em posição imediatamente posterior ao conteúdo. Disclaimer Textual (modelo longo) com sistema, técnica, limitações e link para política de uso.

No Cenário C, não existe variante de risco baixo. Conteúdo audiovisual com verossimilhança extrema exige, no mínimo, a combinação do nível moderado.

5.6. Cenário D — Texto de interesse público gerado por IA

Neste cenário, o usuário acessa texto gerado por IA com finalidade informativa sobre assuntos de interesse coletivo. O objetivo, portanto, é garantir que o usuário forme opinião ou tome decisão tendo ciência da origem artificial da informação.

Nível de risco	Componentes obrigatórios	Componentes recomendados
Moderado	Ícone de IA + Banner Informativo no topo.	Disclaimer Textual (modelo padrão com sistema, data, fonte e revisão) Tooltip Explicativo , acionável a partir do ícone de informação, com detalhes do processo de geração e revisão.
Alto	Ícone de IA + Disclaimer Textual (modelo detalhado com limitações e canal de contato para interação humana) ou Modal de Ciência , com registro auditável, previamente ao acesso, quando o conteúdo subsidia decisão do usuário que possa interferir negativamente em direitos e garantias dos cidadãos.	

No Cenário D, não existe variante de risco baixo. Texto de interesse público gerado por IA exige, no mínimo, a combinação do nível moderado. O *Disclaimer Textual* com metadados estruturados é essencial para a rastreabilidade do conteúdo.

5.7. Cenário E — Decisões automatizadas que afetam interesses do titular

Neste cenário, um sistema de IA toma, recomenda ou influencia de forma preponderante uma decisão com impacto direto sobre direitos ou interesses de uma pessoa. É o cenário de maior exigência de transparência.

Nível de risco	Componentes obrigatórios	Componentes recomendados
Alto	Ícone de IA + Modal de Ciência , com registro auditável, previamente ao acesso + Disclaimer Textual (modelo detalhado com sistema, critérios, limitações, status de revisão, responsável e canal para interação humana). Em determinadas circunstâncias,	Banner Informativo (tipo perigo), para informações complementares no curso do uso da aplicação.

as informações do *Disclaimer* Textual poderão ser incorporadas ao Modal de Ciência.

No Cenário E, o risco é sempre alto. O Modal de Ciência com checkbox obrigatório é imprescindível. O *Disclaimer* Textual no modelo detalhado deve incluir todos os metadados de rastreabilidade, incluindo canal de contestação e responsável pela decisão.

5.8. Sobreposição de cenários

Uma mesma interface pode combinar mais de um cenário. Um *chatbot* que conversa com o usuário (Cenário A), disponibiliza conteúdo de interesse público (Cenário D) e avalia elegibilidade a benefício (Cenário E) enquadra-se simultaneamente nos três cenários.

Quando há sobreposição, aplicam-se três regras:

- i. **Prevalece a combinação mais exigente.** Se o Cenário E exige Modal de Ciência e o Cenário A exige apenas *Badge* Persistente, ambos devem ser aplicados.
- ii. **Os componentes se somam, não se substituem.** A presença do Modal de Ciência (exigido pelo Cenário E) não dispensa o *Badge* Persistente (exigido pelo Cenário A).
- iii. **O conteúdo textual não se repete.** Cada componente deve acrescentar informação, não reproduzir o que já foi comunicado em outro.

O quadro abaixo ilustra, com os elementos obrigatórios, a combinação resultante para o exemplo citado (*chatbot* com funções dos Cenários A, D e E):

Componente	Origem	Função nesta composição
Ícone de IA	A + D + E	Identifica visualmente a origem IA em cada mensagem e no conteúdo gerado
<i>Badge</i> Persistente	A	Mantém a informação visível durante toda a sessão conversacional
Modal de Ciência	E	Exige ciência informada com checkbox e registro auditável antes da análise
<i>Disclaimer</i> Textual (detalhado)	D + E	Reúne metadados: sistema, critérios, limitações, revisão, contestação

5.9. Matriz consolidada — cenário × risco × componentes

O quadro a seguir sintetiza as recomendações de todos os cenários. As células indicam se o componente é obrigatório (O), recomendado (R) ou não aplicável (—).

Cenário	Risco	Ícone de IA	Badge Persistente	Banner Informativo	Tooltip Explicativo	Disclaimer Textual	Modal de Ciência
A	Baixo	O	O	—	—	—	—
A	Moderado	O	O	O	R	R	—
A	Alto	O	O	O	R	R	O*
B	Baixo	O	—	—	R	O	—
B	Moderado	O	—	R	R	O	—
B	Alto	O	—	O	R	O	O*
C	Moderado	O	—	R	R	O	—
C	Alto	O	—	R	R	O	O*
D	Moderado	O	—	O	R	R	—
D	Alto	O	—	—	—	O	O*
E	Alto	O	—	R	—	O	O

Legenda: O = Obrigatório O* = Obrigatório com condição específica R = Recomendado — = Não aplicável

Esta matriz é orientativa. Em caso de sobreposição de cenários, prevalece a combinação mais exigente, conforme a subseção 5.8. Na dúvida entre dois níveis de risco, adotar o nível mais exigente.