

**PRESIDÊNCIA DA REPÚBLICA  
CASA CIVIL  
IMPRENSA NACIONAL**

**PLANO DE DADOS ABERTOS – PERÍODO  
2018/2020**

**APRESENTAÇÃO**

Instituída pelo Decreto nº 8.777, de 11 de maio de 2016, a Política de Dados Abertos do Poder Executivo Federal tem por objetivo promover a disponibilização na internet, em formato aberto, de informações públicas de órgãos e entidades da administração pública federal direta, autárquica e fundacional. De acordo com o normativo, o Plano de Dados Abertos (PDA) é o documento orientador para as ações de implementação da política em cada órgão ou entidade, obedecidos os padrões mínimos de qualidade, de forma a facilitar o entendimento e o reuso das informações.

O presente Plano de Dados Abertos (PDA) atualiza o modelo instituído em 2016, aprovado pela Portaria nº 337, de 14/11/2016, de abertura e divulgação, em formato aberto, dos dados públicos produzidos e custodiados pela Imprensa Nacional, órgão responsável pela publicidade, validação e preservação das informações oficiais, por meio da publicação do Diário Oficial da União (DOU).

Considerando o grau de relevância e interesse público dessas informações, e atendendo ao contido no Decreto, as unidades responsáveis pela sustentação do PDA-IN conduziram a atualização do plano direcionando esforços para a abertura dos dados legados das publicações do DOU, sendo parte composta de arquivos nato-digitais e o restante de arquivos digitalizados.

Este Plano abrange o período de 2018/2020, podendo sofrer eventuais adequações para melhor atendimento de suas finalidades, conforme os resultados alcançados e as manifestações registradas na Central de Atendimento e Ouvidoria. O PDA e suas versões atualizadas serão devidamente publicados no sítio eletrônico da Imprensa Nacional e catalogados no Portal Brasileiro de Dados Abertos.

**1. INTRODUÇÃO**

O Poder Executivo Federal instituiu a política de dados abertos, pelo Decreto nº 8.777, de 11 de maio de 2016, com o intuito de promover a publicação, em

formato aberto, de informações públicas armazenadas nas bases de dados de seus órgãos e entidades englobando administração pública federal direta, autárquica e fundacional, visando, dentre outros objetivos, aprimorar a cultura da transparência e franquear aos cidadãos o acesso dos dados produzidos e acumulados, sobre os quais não recaia vedação expressa de acesso. Posteriormente, o Decreto foi regulamentado pela Resolução nº 3, de 13 de outubro de 2017, do Comitê Gestor da Infraestrutura Nacional de Dados Abertos.

Conforme definição contida no citado Decreto, dados abertos são aqueles acessíveis ao público, representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na internet e disponibilizados sob licença aberta que permita sua livre utilização, consumo ou cruzamento, limitando-se a creditar a autoria ou a fonte.

As normas de publicação de matérias nos jornais oficiais são disciplinadas pela Portaria nº 283, de 2 de outubro de 2018, do diretor-geral da Imprensa Nacional, considerando a autonomia técnica do Órgão para edição e disponibilização do Diário Oficial da União, disposta no artigo 6º do Decreto nº 9.215, de 29 de novembro de 2017, com base na competência definida no art. 47 do Decreto nº 8.889, de 26 de outubro de 2016.

Neste contexto de decretos e normas, em que é fixada a competência exclusiva da Imprensa Nacional para publicação e disponibilização do Diário Oficial da União, cabe tão somente a este Órgão promover o acesso, sob a forma de dados abertos, das publicações oficiais armazenadas na sua base de dados.

O Plano de Dados Abertos instituído pela IN, na forma concebida pelo Decreto nº 8.777, de 11 de maio de 2016, e pela Resolução nº 3, de 13 de outubro de 2017, é o documento orientador para as ações de implementação e promoção de abertura de dados, consistente na publicação e divulgação de dados, em formato aberto, sob sua responsabilidade, considerando em toda a sua estrutura, os princípios e diretrizes dos normativos nacionais vigentes a seguir:

- A Lei Complementar nº 101, de 4 de maio de 2000 (Lei de Responsabilidade Fiscal), que dispõe sobre a obrigatoriedade do Poder Público em adotar instrumentos de transparência na gestão fiscal que permitam o acesso público às informações orçamentárias, visando o controle social como pilar no processo de prestação de contas;
- A Lei nº 12.527, de 18 de novembro de 2011 (Lei de Acesso à Informação – LAI), bem como o Decreto nº 7.724, de 16 de maio de 2012, que a regulamentou, no qual o Poder Público passou a ter o dever de não apenas publicar as informações na internet, mas de disponibilizá-las em formatos abertos e processáveis por máquina;

- O Decreto s/nº de 15 de setembro de 2011, que instituiu o Plano de Ação Nacional do Governo Aberto, o qual estabelece o compromisso do governo de implantar a Infraestrutura Nacional de Dados Abertos (INDA);
- A Instrução Normativa nº 4 de 13 de abril de 2012, que cria a INDA, o Portal Brasileiro de Dados Abertos e estabelece conceitos referentes a dado, informação, dado público, formato aberto, licença aberta, dados abertos e metadados;
- Os parâmetros de arquitetura de interoperabilidade do governo eletrônico estabelecidos na e-PING (Portaria SLTI/MP Nº 92, de 24 de dezembro de 2014), e os vocabulários e ontologias de Governo Eletrônico e-VoG e e-MAG - Modelo de Acessibilidade de Governo Eletrônico (instituído pela Portaria SLTI/MP nº 03, de 07 de maio de 2007); e
- Os parâmetros atuais ou que venham a ser estabelecidos no âmbito de Planejamento Estratégico Institucional ou setorial, bem como os relacionados às áreas de tecnologia da informação (PETIC e PDTIC), sob orientação da Estratégia de Governança Digital – EGD, instituída pelo Decreto nº 8.638, de 15 de janeiro de 2016, e aprovada pela Portaria SLTI/MP nº 68, de 7 de março de 2016, do Ministério do Planejamento Orçamento e Gestão.

O presente PDA pretende ser um marco geral de implantação e não um detalhamento da internalização de práticas e ações. Trata-se, portanto, de um instrumento de planejamento e coordenação das ações de disponibilização de dados da Imprensa Nacional no biênio 2018-2020.

## **1.1. Cenário Institucional**

Este PDA foi revisado em observância ao Planejamento Estratégico Institucional (PEI) da Imprensa Nacional para o período de 2017-2019, em alinhamento direto com sua missão precípua de publicizar e realizar a gestão do conhecimento sobre as informações dos atos oficiais para a sociedade, tornando a oferta, em formato aberto, das publicações do DOU um imperativo na prestação de serviço orientado às novas demandas, impactando diretamente na qualidade do serviço público prestado pelo Órgão.

Sendo o DOU a fonte primária da informação oficial, a divulgação desse conteúdo em dados abertos assume caráter prioritário, visto que contribui efetivamente para a disseminação e desenvolvimento de instrumentos voltados ao aprimoramento da gestão pública e maior controle social, em sintonia com a visão da Instituição de ser referência como fonte exclusiva de informações oficiais com modernidade, confiabilidade e acessibilidade.

Ademais, a elaboração deste PDA e suas revisões posteriores observarão o alinhamento não apenas com o PEI, mas também com os princípios e diretrizes definidas no Planejamento Estratégico de Tecnologia da Informação e Comunicações (PETIC) e no Plano Diretor de Tecnologia da Informação e

Comunicações (PDTIC) no sentido de aperfeiçoar a gestão da informação pública, a partir do aprimoramento das tecnologias de disponibilização da informação.

O Comitê Estratégico de Tecnologia da Informação e Comunicações da Imprensa Nacional (CETIC), instituído pela Portaria nº 317, de 17 de outubro de 2016, para formulação, proposição e acompanhamento do PDTIC, absorverá as ações de abertura de dados do PDA. O acompanhamento do PDA será realizado em reuniões específicas do CETIC.

## **1.2. Estratégia de Governança Digital**

Com o objetivo de gerar benefícios à sociedade mediante uso da tecnologia, estimular a participação social e assegurar o acesso do cidadão à informação pública, foi instituída a Política de Governança Digital para a administração pública, por meio do Decreto nº 8.638, de 15 de janeiro de 2016.

Em seu alinhamento com a política formulada para promover a melhoria dos serviços prestados e a transparência de informações, o PDA observará os seguintes princípios:

➤ **ESTÍMULO À DISPONIBILIZAÇÃO E O USO DE DADOS ABERTOS**

- a) Firmar compromissos de publicação de dados abertos, incluindo levantamento das bases de dados existentes e planejar a abertura de conjuntos de dados com metas e prazos definidos;
- b) Promover a construção coletiva e compartilhada e a sustentação de soluções que usem dados abertos para atender as necessidades ou resolver problemas específicos, estimular modelos inovadores de negócio que se baseiem no uso dos dados; e
- c) Reunir e aplicar as melhores práticas internacionais relacionadas à elaboração e implementação de políticas de dados abertos.

➤ **AMPLIAÇÃO DO USO DE TIC PARA PROMOVER A TRANSPARÊNCIA E DAR PUBLICIDADE À APLICAÇÃO DOS RECURSOS PÚBLICOS**

- a) Transformar em "transparência ativa" os pedidos de informação mais frequentes do Serviço de Informações ao Cidadão; e
- b) Dar publicidade e transparência da aplicação dos recursos públicos no maior número possível de dispositivos e plataformas digitais.

➤ **EXPANSÃO E INOVAÇÃO NA PRESTAÇÃO DE SERVIÇOS DIGITAIS**

- a) Estimular o desenvolvimento de novos serviços digitais com a participação da sociedade civil; e
- b) Expandir os serviços publicados no Portal Institucional.

➤ **SIMPlicidade de uso e universalidade do acesso aos serviços digitais**

- a) Construir interfaces para disponibilização dos dados governamentais; e

- b) Implantar ferramentas que promovam a comunicação proativa com o cidadão.

### **1.3. Objetivo Geral**

Promover a abertura de dados de interesse público na Imprensa Nacional, zelando pelos princípios da publicidade, transparência e eficiência, e visando o aumento da disseminação de dados e informações para a sociedade, bem como a melhoria da qualidade dos dados disponibilizados, de forma a dar maior suporte à tomada de decisão pelos gestores públicos e ao controle social.

### **1.4. Objetivos Específicos**

Em alinhamento com sua missão institucional, no âmbito da Imprensa Nacional, os objetivos específicos incluem:

- Abertura dos dados das publicações do DOU, em razão do grande interesse público e da priorização expressa no Decreto nº 8.777/2016;
- Identificar outras prioridades, a partir de demandas recebidas pelos canais de comunicação, e disponibilizá-los em formato aberto sempre que possível;
- Garantir uma política de dados abertos progressiva e continuada;
- Aprimorar os processos e ferramentas de transparência e acesso às informações públicas.

## **2. DEFINIÇÃO DOS DADOS A SEREM ABERTOS**

A abertura de dados deve, sobretudo, garantir os princípios da publicidade e da transparência da administração pública. Para o estabelecimento das metas de abertura de dados e etapas de priorização, deve-se levar em consideração os seguintes critérios:

1. O grau de relevância e interesse público dessas informações;
2. Volume de demandas encaminhadas via Central de Atendimento e e-SIC, bem como os setores e serviços mais procurados no portal institucional do órgão;
3. O alinhamento perante o Planejamento Estratégico Institucional (PEI) e o Plano Diretor de Tecnologia da Informação e da Comunicação (PDTIC); e
4. O nível de maturidade da organização das informações e dados existentes.

No caso da Imprensa Nacional, a priorização da divulgação das publicações do DOU em dados aberto está prevista no Anexo ao Decreto nº 8.777/2016, em razão da relevância e interesse público dessas informações.

Inicialmente, serão disponibilizadas em formato aberto as novas publicações do DOU, com adequação imediata do formato e metadados mínimos exigidos. Posteriormente, as edições passadas, em ordem decrescente de data de publicação, até a oferta de todo o acervo do DOU em dados abertos.

Demais informações públicas poderão ser fornecidas em dados abertos a partir de análise futura, de acordo com os critérios acima estabelecidos.

### **3. ESTRATÉGIAS PARA ABERTURA**

Como conjunto de dados candidatos à abertura, foram definidos e priorizados o conteúdo das edições do Diário Oficial da União publicadas a partir da implementação do plano.

Considerando o conjunto de dados candidatos à abertura, este plano priorizou, inicialmente, a abertura dos dados das Seções 1 e 2, tendo sido publicados no Portal de Dados Abertos do Governo Federal desde dezembro de 2017. Posteriormente, no mês de março de 2018, foi adicionada ao conjunto a seção 3 do DOU. Para tanto, o órgão modernizou plataforma tecnológica de editoração eletrônica de matérias, permitindo hoje, a produção de dados do DOU em formato aberto.

Uma vez iniciado o processo de publicação em formato de dados abertos das edições do DOU, em dezembro de 2017, a Imprensa Nacional iniciou processo de conversão das publicações de anos anteriores para o formato de dados abertos, seguindo a ordem decrescente de datas até alcançar as publicações de 1988, ano da promulgação da Constituição Federal.

A previsão é de disponibilização, até dezembro de 2018, de todo o conteúdo das edições do DOU publicadas de 2002 a 2017. Para o período de 2001 a 1988, a Imprensa Nacional está avaliando tecnologias relacionadas com o reconhecimento ótico de caracteres (OCR) para promover a conversão dos atos oficiais publicados no DOU neste período. O plano de disponibilização em dados abertos do conteúdo de edições publicadas entre 2001 e 1988 será definido na próxima revisão deste PDA, e levará em conta as demandas recebidas pelos canais de atendimento ao público.

A preparação e a atualização dos dados do DOU a serem publicados em formato aberto é de responsabilidade da Coordenação-Geral de Publicação e Divulgação com o apoio técnico da Coordenação de Tecnologia da Informação.

A governança e propostas de alteração/revisão e atualização do PDA fica a cargo do Comitê Estratégico de Tecnologia da Informação e da Comunicação - CETIC da Imprensa Nacional.

A metodologia a ser utilizada para a produção dos dados do DOU em formato aberto segue os padrões definidos pela INDA, sendo adotado o XML (*eXtensible Markup Language*), na forma prevista no Consórcio World Wide Web (W3C), o formato adotado para os arquivos de dados abertos.

Os servidores envolvidos nos processos de publicação de dados abertos foram capacitados pela empresa responsável pela implantação da nova plataforma editorial a ser utilizada pela Imprensa Nacional. Para servidores envolvidos nos processos de catalogação dos metadados no Portal de Dados Abertos, a capacitação foi realizada pela equipe de Dados Abertos da Central de Serviços e Suporte do SISP (C3S). Além das formas de capacitação já realizadas, haverá capacitação dos servidores da Imprensa Nacional nos cursos de qualificação em governança de dados abertos providos pela Escola Nacional de Educação Pública - ENAP.

Os dados do DOU em formato aberto são publicados no Portal da Imprensa Nacional e catalogados no Portal Brasileiro de Dados Abertos pela Coordenação de Editoração e Divulgação Eletrônica dos Jornais Oficiais. Para publicação dos dados no Portal da Imprensa Nacional, observando-se o uso da URL fixa, é utilizado sistema automatizado desenvolvido e mantido pela Coordenação de Tecnologia da Informação. A catalogação dos atos oficiais do DOU em formato aberto ocorre regularmente até quinto dia útil do mês subsequente ao das edições.

### **3.1. Premissas**

Além de estar alinhado aos princípios e diretrizes mencionados anteriormente neste PDA, o processo de abertura dos conjuntos de dados considera as seguintes premissas:

1. São disponibilizados os dados das publicações do DOU, em razão o grande interesse público, no formato definido pela INDA, sendo adotado o XML (*eXtensible Markup Language*), e informando as eventuais limitações de qualidade dos dados.
2. Sempre que possível, serão publicados dados e seus metadados conforme estabelecido no Plano de Ação da INDA, que institui que cada conjunto de dados deve conter, no mínimo:
  - a. Nome ou título do conjunto de dados;
  - b. Descrição sucinta;
  - c. Palavras-chave (etiquetas);
  - d. Assuntos relacionados do VCGE – Vocabulário Controlado do Governo Eletrônico;
  - e. Nome e e-mail do setor responsável pelos dados;
  - f. Periodicidade de atualização; e
  - g. Escopo temporal (anual, mensal, diário, bimestral).
3. Os dados abertos publicados são catalogados no Portal Brasileiro de Dados Abertos, ponto central de acesso aos dados do Governo Federal.

4. Manter os dados publicados atualizados e sincronizados com a origem, com a menor periodicidade e maior granularidade viáveis.
5. A atualização dos dados ocorre por meio de sincronização automática, estabelecendo-se um processo contínuo, com ganhos de eficiência em comparação a extrações pontuais.
6. Utilizar, como forma de disseminação, os ambientes do Portal Brasileiro de Dados Abertos do Governo Federal e, utilizar também a página institucional da Imprensa Nacional ([www.in.gov.br/dadosabertos](http://www.in.gov.br/dadosabertos)), mantendo URL fixa.

### **3.2. SITUAÇÃO ATUAL DOS DADOS ABERTOS**

As publicações das Seções 1 e 2 do DOU já vem sendo catalogados mensalmente no Portal de Dados Abertos desde dezembro de 2017, data que coincide com a descontinuidade da versão impressa do DOU. Em março de 2018, os dados da Seção 3 também passaram a integrar a rotina de catalogação. Desde então, os dados das novas publicações do DOU são regularmente disponibilizados em formato aberto.

Posteriormente, a Imprensa Nacional obteve a tecnologia necessária de extração estruturada de dados de arquivos PDF nativo-digitais. Dessa forma, foi possível realizar a conversão de PDF para XML estruturado do legado das publicações de 2002 a novembro de 2017, iniciando pela da Seção 1. Essa base de dados já se encontra catalogada no Portal de Dados Abertos desde julho deste ano.

O mesmo trabalho já está sendo realizado com o legado das seções 2 e 3, com previsão de disponibilização em dados abertos até dezembro de 2018.

### **4. CATALOGAÇÃO NO PORTAL BRASILEIRO DE DADOS ABERTOS**

O processo de catalogação no Portal de Brasileiro de Dados Abertos é feito diretamente pela Coordenação de Tecnologia de Informação – CORTI. Para isso, foi designada uma unidade responsável pela atividade, cujos servidores estão capacitados para realizar o cadastro, manutenção e revisão dos conjuntos de dados conforme o Manual de Catalogação.

O cadastro dos mantenedores dos metadados no portal também está a cargo da Coordenação de Tecnologia da Informação.

### **5. SUSTENTAÇÃO**

A condução do Plano de Dados Abertos conta com o envolvimento das áreas relacionadas à gestão da informação, que atuam de forma colaborativa, cujas responsabilidades estão assim distribuídas:

**Coordenação-Geral de Publicação e Divulgação (COGED):** preparar e atualizar os dados e detalhar o plano de ação com metas e prazos, com o apoio técnico da Coordenação de Tecnologia da Informação;

**Comitê Estratégico de Tecnologia da Informação e da Comunicação (CETIC):** elaborar e discutir matriz de responsabilidades, governança e propostas de alteração/revisão e atualização do PDA, segundo metodologia e padrões definidos pela IND;

**Coordenação de Tecnologia da Informação (CORTI):** sistematizar as rotinas de geração e disponibilização dos dados abertos; capacitar equipes para utilização dos recursos relacionados à publicação dos dados e catalogação dos metadados; definição de arquitetura de abertura para cada sistema; e efetuar catalogação no Portal de Dados Abertos, conforme as regras definidas, dos dados produzidos;

**Coordenação de Editoração e Divulgação Eletrônica dos Jornais Oficiais (COEJO):** efetuar a publicização dos dados abertos no Portal da Imprensa Nacional, por meio de sistema automatizado desenvolvido e mantido pela Coordenação de Tecnologia da Informação;

**Coordenação de Relacionamento Externo (COREX):** com base nas manifestações recebidas via Ouvidoria, Central de Atendimento e Serviço de Informações ao Cidadão, a COREX ficará responsável pela emissão de relatórios mensais contendo novas demandas e solicitações de abertura de dados a ser submetido ao Comitê Estratégico de Tecnologia da Informação e da Comunicação, que se reunirá para analisar a possibilidade de atendimento.

Havendo alteração na estrutura do órgão, o PDA será oportunamente atualizado, com as devidas adequações acerca das responsabilidades ora descritas.

## 6. MONITORAMENTO, CONTROLE E MELHORIA

Conforme estabelecido no art. 5º, §4º, do Decreto nº 8.777/2016, a autoridade designada nos termos do art. 40 da Lei nº 12.527, de 2011 (Lei de Acesso à Informação) é a responsável por assegurar a publicação e a atualização do Plano de Dados Abertos, dentre outras atribuições. Na Imprensa Nacional, esse encargo cabe atualmente ao titular da Chefia de Gabinete da Diretoria-Geral, o qual responde pela implementação e monitoramento do PDA, exercendo as seguintes atribuições:

- orientar as unidades sobre o cumprimento das normas referentes a dados abertos;
- assegurar o cumprimento das normas relativas à publicação de dados abertos, de forma eficiente e adequada;

- monitorar a implementação dos Planos de Dados Abertos; e
- apresentar relatórios periódicos sobre o cumprimento do PDA, com recomendações sobre as medidas indispensáveis à implementação e ao aperfeiçoamento da Política de Dados Abertos.

Além do monitoramento da autoridade designada, o PDA tem seu cumprimento acompanhado pelo CETIC com atualização de suas metas, prazos, indicadores e dados, a serem incorporados ao PDTIC. Cabe ao CETIC verificar o alinhamento do PDA com os instrumentos de planejamento aplicados ao Órgão, como o PEI e o PETIC, além da atualização do próprio PDTIC.

A gestão da qualidade dos dados está a cargo do **CETIC**, que avalia, em suas reuniões periódicas, melhorias com base nos indicadores e demandas recebidas pelos canais de comunicação da Imprensa Nacional.

## **6.1 Comunicação e Participação Social**

A institucionalização do Plano de Dados Abertos da Imprensa Nacional, bem como sua governança e revisões, serão divulgadas por meio do portal institucional na Internet e Intranet, de modo a alcançar, com a abordagem apropriada, tanto o público interno quanto a sociedade, aprimorando a disseminação da cultura da transparência no âmbito do Órgão e de toda a administração.

Da mesma forma, sempre que ocorrer a atualização ou inserção de novos dados, a informação será amplamente divulgada, inclusive por meio de ações específicas de comunicação coordenadas pela Assessoria de Comunicação (ASCOM).

Considerando o grande interesse sobre as publicações do DOU, a Central de Atendimento da Imprensa Nacional está preparada para receber manifestações de natureza técnica, bem como críticas e sugestões sobre o serviço, dando o devido encaminhamento à área responsável para resposta, solução ou avaliação, conforme o caso.

## 7. PLANO DE AÇÃO

### Quadro descritivo

Tema	Atividades	Produtos	Meta/Prazo	Requisitos	Unidade Responsável	Ponto Focal
Seção 3 – dados referentes aos meses de dezembro 2017 e janeiro de 2018	1 – Extração estruturada de dados de arquivos PDF; 2 – Conversão dos dados para o formato XML; 3 – Publicação no Portal de Dados Aberto do Governo Federal	1 – Catálogo de dados abertos publicados; 2 – Arquivo em formato aberto publicado no Portal da IN.	Dezembro de 2018	Customização da solução de extração de dados do DOU em formato PDF e validação do conteúdo	DIRGE COGED CORTI	Debora Beserra <a href="mailto:deborabeserra@in.gov.br">deborabeserra@in.gov.br</a> Alexandre Machado <a href="mailto:alexandremachado@in.gov.br">(alexandremachado@in.gov.br)</a> Symball Oliveira <a href="mailto:symballoliveira@in.gov.br">(symballoliveira@in.gov.br)</a>
Dados referentes às Seções 2 e 3 – <b>Legados</b> de 2002 a 2017	1 – Extração estruturada de dados de arquivos PDF; 2 – Conversão dos dados para o formato XML; 3 – Publicação no Portal de Dados Aberto do Governo Federal	1 – Catálogo de dados abertos publicados; 2 – Arquivo em formato aberto publicado no Portal da IN.	Dezembro de 2018	Customização da solução de extração de dados do DOU em formato PDF e validação do conteúdo	DIRGE COGED CORTI	Debora Beserra <a href="mailto:deborabeserra@in.gov.br">deborabeserra@in.gov.br</a> Alexandre Machado <a href="mailto:alexandremachado@in.gov.br">(alexandremachado@in.gov.br)</a> Symball Oliveira <a href="mailto:symballoliveira@in.gov.br">(symballoliveira@in.gov.br)</a>
Dados referentes às Seções 1, 2 e 3 – <b>Legados</b> (de 1988 a 2001)	1 – Desenvolver/obter solução para obter tecnologia baseada em reconhecimento ótico de caracteres para extração estruturada de dados de arquivos PDF; 2 – Extração estruturada de dados de arquivos PDF; 3 – Conversão dos dados para o	1 – Catálogo de dados abertos publicados; 2 – Arquivo em formato aberto publicação no Portal da IN.	Dezembro de 2019	Adquirir ou desenvolver solução tecnológica para extração estruturada de dados de arquivos PDF baseada em solução de reconhecimento	DIRGE COGED CORTI	Debora Beserra <a href="mailto:deborabeserra@in.gov.br">deborabeserra@in.gov.br</a> Alexandre Machado <a href="mailto:alexandremachado@in.gov.br">(alexandremachado@in.gov.br)</a> Symball Oliveira <a href="mailto:symballoliveira@in.gov.br">(symballoliveira@in.gov.br)</a>

	formato XML; 4 – Publicação no Portal de Dados Aberto do Governo Federal			ótico de caracteres		
Dados referentes às Seções 1, 2 e 3 – <b>Legados</b> (de 1808 até 1987)	1 - Publicação no Portal da Imprensa Nacional em formato PDF. 2 – Indexação do conteúdo e disponibilização para pesquisa.	1 – Arquivo em formato binário (PDF) com OCR para acervo digital.	A ser definido na próxima revisão deste PDA	Adquirir ou desenvolver solução tecnológica para extração estruturada de dados de arquivos PDF baseada em solução de reconhecimento ótico de caracteres	DIRGE COGED CORTI	Debora Beserra <a href="mailto:deborabeserra@in.gov.br">deborabeserra@in.gov.br</a> Alexandre Machado <a href="mailto:alexandremachado@in.gov.br">alexandremachado@in.gov.br</a> Symball Oliveira <a href="mailto:symballoliveira@in.gov.br">symballoliveira@in.gov.br</a>

## **8. GLOSSÁRIO**

**Arquitetura de interoperabilidade do governo eletrônico (e-PING)** define um conjunto mínimo de premissas, políticas e especificações técnicas que regulamentam a utilização da Tecnologia de Informação e Comunicação (TIC) na interoperabilidade de serviços de Governo Eletrônico, estabelecendo as condições de interação com os demais Poderes e esferas de governo e com a sociedade em geral.

**Comitê Estratégico de Tecnologia da Informação e Comunicações da Imprensa Nacional (CETIC)** é um órgão colegiado de assessoria estratégica, ao qual compete deliberar sobre assuntos de Tecnologia da Informação e Comunicações (TIC) no âmbito da Imprensa Nacional, submetido à autoridade do Diretor-Geral tendo, especificamente como missão validar, supervisionar e monitorar as ações previstas no Plano Diretor de Tecnologia da Informação e Comunicações (PDTIC).

**Estratégia de Governança Digital (EGD)** define os objetivos estratégicos, metas e indicadores da Política de Governança Digital, estabelecida pelo Decreto nº 8.638, de 15 de janeiro de 2016. A publicação é composta de três eixos, dez objetivos e 51 iniciativas estratégicas.

**Infraestrutura Nacional de Dados Abertos (INDA)** é um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos. A INDA é a política do governo brasileiro para dados abertos.

**Padrão de Metadados do Governo Eletrônico (e-PMG)** objetiva assegurar às pessoas que pesquisam as informações do governo brasileiro na Web o acesso rápido e eficiente a descrições dos recursos. Os elementos e-PMG têm o propósito de facilitar a localização dos recursos que as pessoas necessitam, mesmo que elas não possuam conhecimento detalhado da localização ou da entidade responsável pela informação.

**Planejamento Estratégico de Tecnologia da Informação e Comunicações (PETIC)** tem por objetivo assegurar que as metas e objetivos da TI estejam fortemente vinculados às metas e objetivos do negócio/estratégia da Instituição e, portanto, alinhado com seu Planejamento Estratégico.

**Planejamento Estratégico Institucional (PEI)** constitui-se de um plano de trabalho para o alcance dos objetivos e metas que o Instituto se propõe.

**Plano de Dados Abertos (PDA)** é o documento orientador para as ações de implementação da política em cada órgão ou entidade, obedecidos os padrões mínimos de qualidade, de forma a facilitar o entendimento e o reuso das informações.

**Plano Diretor de Tecnologia da Informação (PDTI)** é o instrumento de diagnóstico, planejamento e gestão dos recursos e processos de Tecnologia da Informação e Comunicações que visa atender às necessidades tecnológicas, de informação e de comunicações de um órgão ou entidade para um determinado período.

**Política de Dados Abertos** tem por objetivo promover a disponibilização na internet, em formato aberto, de informações públicas de órgão e entidades da administração pública federal direta, autárquica e fundacional.

**SIOPLegis** é um sistema eletrônico que reúne a legislação relacionada direta ou indiretamente com o Orçamento da União. Disponível em: <https://www.siop.planejamento.gov.br/sioplegis/sof>.

**Vocabulário Controlado do Governo (VCGE)** é um esquema para ser utilizado no elemento assunto. Categoria (subject.category) do Padrão de Metadados do Governo Eletrônico (e-PMG). Termos retirados do VCGE tornarão mais direta, para os gerentes de sítios e portais governamentais, a apresentação dos serviços disponibilizados em uma estrutura de diretório baseada nos indexadores do VCGE.

**XML (eXtensible Markup Language)** é uma linguagem de marcação recomendada pela W3C para a criação de documentos com dados organizados hierarquicamente, tais como textos, banco de dados ou desenhos vetoriais. A linguagem XML é classificada como extensível porque permite definir os elementos de marcação.

**W3C (World Wide Web Consortium)** é a principal organização de padronização da World Wide Web. Consiste em um consórcio internacional com quase 400 membros, agrupa empresas, órgãos governamentais e organizações independentes com a finalidade de estabelecer padrões para a criação e a interpretação de conteúdos para a Web.