Tribunal de Contas da União
Secretaria-Geral de Controle Externo
Secretaria de Gestão de Processos

OFÍCIO 27378/2022-TCU/Seproc

Brasília-DF, 7/6/2022.

Ao(À) Senhor(a)
Diretor(a)-Geral do Tribunal Regional do Trabalho da 7ª Região/CE

Processo TC 006.662/2021-8 Tipo do processo: Relatório de Levantamento
Relator do processo: Ministro Aroldo Cedraz
Unidade responsável: Secretaria de Fiscalização de Tecnologia da Informação

Assunto: Notificação de acórdão.
Anexos: peças 628, 632, 633 e 634 do processo TC 006.662/2021-8.

Senhor(a) Diretor(a)-Geral,

1. Informo Vossa Senhoria do Acórdão 1139/2022-TCU-Plenário, de relatoria do Ministro Aroldo Cedraz, prolatado na sessão de 25/5/2022, por meio do qual o Tribunal de Contas da União apreciou o processo acima indicado.

2. Encaminho cópia do referido acórdão, cujo inteiro teor pode ser acessado no Portal TCU, endereço eletrônico www.tcu.gov.br/acordaos.

3. Por oportuno, solicito especial atenção às informações complementares que acompanham este ofício, bem assim para a necessidade de utilizar – para resposta a comunicações e envio de documentos – os serviços da plataforma Conecta-TCU ou do protocolo eletrônico, disponíveis no Portal TCU (www.tcu.gov.br), endereço em que também é possível acessar os autos do processo.

4. Esclarecimentos adicionais quanto ao processo indicado ou à presente comunicação podem ser obtidos junto à Secretaria de Gestão de Processos (Seproc) pelo e-mail cacidadao@tcu.gov.br, ou pelos telefones 0800-644-2300, opção 2, ou (61) 3527-5234 no horário das 13h às 17h.

Atenciosamente,

assinado eletronicamente
Maryzely Mariano
Chefe do Serviço de Comunicação Processual 1
(Subdelegação de competência: art. 2º, I, da Portaria-Seproc 1/2020)

Documento juntado por joao.lima - JOAO RIBEIRO LIMA JUNIOR

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 7156710.

INFORMAÇÕES COMPLEMENTARES

1) O acesso ao processo indicado nesta comunicação pode ser realizado por meio da plataforma de serviços digitais Conecta-TCU, disponível no Portal TCU (www.tcu.gov.br). A visualização de processos e documentos sigilosos depende de solicitação formal e posterior autorização do relator. Informações detalhadas sobre o uso da plataforma, inclusive para fins de cadastro e credenciamento, podem ser consultadas ao acionar o ícone "Conecta-TCU" do Portal TCU.

2) No caso de acórdãos apreciados por relação, na forma do art. 143 do Regimento Interno do TCU, não há relatório e voto. A fundamentação com que o Tribunal analisa as questões de fato e de direito encontra-se na instrução da unidade técnica juntada aos autos.

3) Nos termos do art. 18, §4º, da Resolução-TCU nº 170/2004, quando da apreciação de recurso interposto à deliberação do Tribunal, são expedidas comunicações sobre a deliberação adotada a todas as autoridades, responsáveis e interessados a quem foi dirigida comunicação quando da adoção da deliberação recorrida.

4) A juntada aos autos do instrumento de mandato, quando a parte for representada por procurador, é pressuposto essencial para a atuação do mandatário no processo, nos termos do art. 13, § 2º, da Resolução - TCU 36/1995.

5) Constitui dever das partes, de seus procuradores e de todos aqueles que de qualquer forma participem do processo indicar, no primeiro momento de falar nos autos, o endereço eletrônico e o endereço residencial ou profissional onde receberão as notificações, bem assim atualizar essas informações sempre que ocorrer qualquer modificação temporária ou definitiva, conforme previsto no art. 77, inciso V, do Código de Processo Civil, aplicável nos termos do art. 298 do Regimento Interno do TCU.

6) Nos termos dos arts. 31 a 35 da Lei nº 8.443/1992 e 285 a 288 do Regimento Interno do TCU, a parte poderá interpor recurso ao acórdão. A interposição de embargos de declaração é causa de mera suspensão e não de interrupção de prazo para os demais recursos, conforme disposto no art. 34, § 2º, da Lei nº 8.443/1992.

7) A apresentação de petição ou a interposição de recurso deve observar as seguintes orientações:

a) ser dirigida ao relator do processo;
b) indicar, com destaque, o número do processo e deste ofício;
c) utilizar dos serviços da plataforma digital Conecta-TCU ou do protocolo eletrônico disponíveis no Portal TCU;
d) a petição ou o recurso podem ser apresentados diretamente pelo destinatário do ofício ou por intermédio de procurador regularmente constituído nos autos, conforme disciplina o art. 145 do Regimento Interno do TCU;
e) caso haja procurador constituído nos autos, as comunicações processuais subsequentes serão dirigidas a esse representante, conforme disposto no art. 179, § 7º, do Regimento Interno do TCU. Se houver mais de um procurador, pode ser indicado o nome daquele a quem deverão ser encaminhadas as comunicações.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 71156710.

8) A informação classificada na origem com restrição de acesso deve ser acompanhada dos seguintes elementos, consoante a Lei de Acesso à Informação (Lei nº 12.527, de 18 de novembro de 2011), caso contrário será considerada de acesso público pelo Tribunal:

a) indicação objetiva da hipótese de restrição de acesso: informação imprescindível à segurança da sociedade ou do Estado; informação com sigilo atribuído por legislação específica; informação pessoal relativa à intimidade, vida privada, honra e imagem;

b) na hipótese de informação imprescindível à segurança da sociedade ou do Estado, indicar:

b.1) o grau de sigilo da classificação (reservado, secreto ou ultrassecreto);

b.2) o fundamento legal da classificação;

b.3) o prazo de restrição de acesso ou o evento que defina o termo final;

b.4) o assunto sobre o qual versa a informação.

c) na hipótese de informação com sigilo atribuído por legislação específica, indicar o fundamento legal da classificação;

d) na hipótese de informação pessoal relativa à intimidade, vida privada, honra e imagem, indicar o prazo de restrição de acesso e a pessoa a que se refere;

e) indicação do nome do responsável pela classificação.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 71156710.

GRUPO I – CLASSE V – Plenário
TC 006.662/2021-8
Natureza: Relatório de Levantamento
Órgãos/Entidades: Advocacia-Geral da União; Agência Brasileira de Desenvolvimento Industrial; Agência Brasileira de Inteligência; Agência Brasileira Gestora de Fundos Garantidores e Garantias S.a.; Agência Espacial Brasileira; Agência Nacional de Águas; Agência Nacional de Assistência Técnica e Extensão Rural; Agência Nacional de Aviação Civil; Agência Nacional de Energia Elétrica; Agência Nacional de Mineração; Agência Nacional de Saúde Suplementar; Agência Nacional de Telecomunicações; Agência Nacional de Transportes Aquaviários; Agência Nacional de Transportes Terrestres; Agência Nacional de Vigilância Sanitária; Agência Nacional do Cinema; Agência Nacional do Petróleo, Gás Natural e Biocombustíveis; Amazonas Geração e Transmissão de Energia S.a.; Amazônia Azul Tecnologias de Defesa S.a.; Associação das Pioneiras Sociais; Autoridade Portuaria de Santos S.a; Banco Central do Brasil; Banco da Amazônia S.a.; Banco do Brasil S.a.; Banco do Nordeste do Brasil S.a.; Caixa Econômica Federal; Casa da Moeda do Brasil; Centrais de Abastecimento de Minas Gerais S. A.; Centrais Eletricas Brasileiras Sa; Centrais Elétricas do Norte do Brasil S.a.; Centro de Tecnologia da Indústria Química e Têxtil do Senai/rj; Centro Federal de Educação Tecnológica Celso Suckow da Fonseca; Centro Federal de Educação Tecnológica de Minas Gerais; Centro Nacional de Tecnologia Eletrônica Avançada S.a; Colégio Pedro Ii; Comando da Aeronautica; Comando da Marinha; Comando do Exército; Comissão de Valores Mobiliários; Comissão Nacional de Energia Nuclear; Companhia Brasileira de Trens Urbanos; Companhia de Desenvolvimento dos Vales do São Francisco e do Parnaíba; Companhia de Geração e Transmissão de Energia Elétrica do Sul do Brasil; Companhia de Pesquisa de Recursos Minerais; Companhia Docas do Ceará; Companhia Docas do Espírito Santo; Companhia Docas do Pará; Companhia Docas do Rio de Janeiro; Companhia Docas do Rio Grande do Norte; Companhia Hidro Elétrica do São Francisco; Companhia Nacional de Abastecimento; Conselho Administrativo de Defesa Econômica; Conselho da Justiça Federal; Conselho Nacional de Desenvolvimento Científico e Tecnológico; Conselho Nacional de Justiça; Conselho Nacional do Ministério Público; Conselho Superior da Justiça do Trabalho; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Defensoria Pública da União; Departamento de Polícia Federal; Departamento de Polícia Rodoviária Federal; Departamento Nacional de Infraestrutura de Transportes; Departamento Nacional de Obras Contra As Secas; Eletrobrás Termonuclear S.a.; Empresa Brasileira de Correios e Telegrafos - ECT; Empresa Brasileira de Infraestrutura Aeroportuária; Empresa Brasileira de Pesquisa Agropecuária; Empresa Brasileira de Serviços Hospitalares; Empresa de Pesquisa Energética; Empresa de Tecnologia e Informações da Previdência -

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Dataprev; Empresa Gerencial de Projetos Navais; Financiadora de Estudos e Projetos; Fundação Alexandre de Gusmão; Fundação Biblioteca Nacional; Fundação Cultural Palmares; Fundação Escola Nacional de Administração Pública; Fundação Instituto Brasileiro de Geografia e Estatística; Fundação Instituto de Pesquisa Econômica Aplicada; Fundação Jorge Duprat Figueiredo, de Segurança e Medicina do Trabalho; Fundação Nacional de Artes; Fundação Nacional de Saúde; Fundação Nacional do Índio; Fundação Oswaldo Cruz; Fundação Universidade de Brasília; Fundação Universidade do Amazonas; Fundação Universidade Federal da Grande Dourados; Fundação Universidade Federal de Ciências da Saúde de Porto Alegre; Fundação Universidade Federal de Mato Grosso do Sul; Fundação Universidade Federal de Ouro Preto; Fundação Universidade Federal de Rondônia; Fundação Universidade Federal de São Carlos; Fundação Universidade Federal de São João Del Rei; Fundação Universidade Federal de Sergipe; Fundação Universidade Federal de Uberlândia; Fundação Universidade Federal de Viçosa; Fundação Universidade Federal do Abc; Fundação Universidade Federal do Acre; Fundação Universidade Federal do Amapá; Fundação Universidade Federal do Maranhão; Fundação Universidade Federal do Pampa; Fundação Universidade Federal do Piauí; Fundação Universidade Federal do Rio Grande; Fundação Universidade Federal do Tocantins; Fundação Universidade Federal do Vale do São Francisco; Fundo Nacional de Desenvolvimento da Educação; Furnas Centrais Elétricas S.a.; Hospital Nossa Senhora da Conceição S.a.; Indústrias Nucleares do Brasil S.a.; Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis; Instituto Chico Mendes de Conservação da Biodiversidade; Instituto do Patrimônio Histórico e Artístico Nacional - IPHAN; Instituto Federal de Educação, Ciência e Tecnologia Farroupilha; Instituto Federal de Educação, Ciência e Tecnologia Baiano; Instituto Federal de Educação, Ciência e Tecnologia Catarinense; Instituto Federal de Educação, Ciência e Tecnologia da Bahia; Instituto Federal de Educação, Ciência e Tecnologia da Paraíba; Instituto Federal de Educação, Ciência e Tecnologia de Alagoas; Instituto Federal de Educação, Ciência e Tecnologia de Brasília; Instituto Federal de Educação, Ciência e Tecnologia de Goiás; Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso; Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais; Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco; Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina; Instituto Federal de Educação, Ciência e Tecnologia de São Paulo; Instituto Federal de Educação, Ciência e Tecnologia do Acre; Instituto Federal de Educação, Ciência e Tecnologia do Amapá; Instituto Federal de Educação, Ciência e Tecnologia do Amazonas; Instituto Federal de Educação, Ciência e Tecnologia do Ceará; Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo; Instituto Federal de Educação, Ciência e Tecnologia do Maranhão; Instituto Federal de

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Educação, Ciência e Tecnologia do Pará; Instituto Federal de Educação, Ciência e Tecnologia do Paraná; Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro; Instituto Federal de Educação, Ciência e Tecnologia do Sertão Pernambucano; Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais; Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais; Instituto Federal de Educação, Ciência e Tecnologia do Tocantins; Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro; Instituto Federal de Educação, Ciência e Tecnologia Goiano; Instituto Nacional da Propriedade Industrial; Instituto Nacional de Colonização e Reforma Agrária; Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira; Instituto Nacional de Metrologia, Qualidade e Tecnologia; Instituto Nacional de Tecnologia da Informação; Instituto Nacional do Seguro Social; Ministério da Ciência, Tecnologia e Inovações; Ministério da Economia; Ministério Público da União; Petrobras Transporte S.A.- MME; Petróleo Brasileiro S.a.; Procuradoria Geral da Fazenda Nacional; Secretaria de Controle Interno do Ministério das Relações Exteriores; Secretaria de Empreendedorismo e Inovação; Secretaria Especial da Receita Federal do Brasil; Secretaria - Executiva da Casa Civil da Presidência da República; Secretaria - Executiva do Ministério da Agricultura, Pecuária e Abastecimento; Secretaria -Executiva do Ministério da Cidadania; Secretaria - Executiva do Ministério da Ciência, Tecnologia e Inovações; Secretaria -Executiva do Ministério da Economia; Secretaria - Executiva do Ministério da Educação; Secretaria -Executiva do Ministério da Infraestrutura; Secretaria -Executiva do Ministério da Justiça e Segurança Pública; Secretaria -Executiva do Ministério da Mulher, da Família e dos Direitos Humanos; Secretaria -Executiva do Ministério da Saúde; Secretaria -Executiva do Ministério da Transparência, Fiscalização e Controladoria-Geral da União; Secretaria -Executiva do Ministério das Comunicações; Secretaria - Executiva do Ministério de Minas e Energia; Secretaria -Executiva do Ministério do Desenvolvimento Regional; Secretaria -Executiva do Ministério do Meio Ambiente; Secretaria -Executiva do Ministério do Turismo; Secretaria -Geral do Ministério da Defesa; Superintendência da Zona Franca de Manaus; Superintendência de Desenvolvimento da Amazônia; Superintendência de Desenvolvimento do Nordeste; Superintendência de Seguros Privados; Superintendência do Desenvolvimento do Centro -Oeste; Superintendência Nacional de Previdência Complementar; Superior Tribunal Militar; Telecomunicações Brasileiras S.a.; Tribunal de Justiça do Distrito Federal e dos Territórios; Tribunal Regional do Trabalho da 1ª Região/rj; Tribunal Regional do Trabalho da 10ª Região/df e TO; Tribunal Regional do Trabalho da 11ª Região/am e RR; Tribunal Regional do Trabalho da 12ª Região/sc; Tribunal Regional do Trabalho da 13ª Região/pb; Tribunal Regional do Trabalho da 14ª Região/ac e RO; Tribunal Regional do Trabalho da 15ª Região - Campinas/SP; Tribunal Regional do Trabalho da 16ª

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Região/ma; Tribunal Regional do Trabalho da 17ª Região/es; Tribunal Regional do Trabalho da 18ª Região/go; Tribunal Regional do Trabalho da 19ª Região/al; Tribunal Regional do Trabalho da 2ª Região/sp; Tribunal Regional do Trabalho da 20ª Região/se; Tribunal Regional do Trabalho da 21ª Região/rn; Tribunal Regional do Trabalho da 22ª Região/pi; Tribunal Regional do Trabalho da 23ª Região/mt; Tribunal Regional do Trabalho da 24ª Região/ms; Tribunal Regional do Trabalho da 3ª Região/mg; Tribunal Regional do Trabalho da 4ª Região/rs; Tribunal Regional do Trabalho da 5ª Região/ba; Tribunal Regional do Trabalho da 6ª Região/pe; Tribunal Regional do Trabalho da 7ª Região/ce; Tribunal Regional do Trabalho da 8ª Região/pa e AP; Tribunal Regional do Trabalho da 9ª Região; Tribunal Regional Eleitoral da Bahia; Tribunal Regional Eleitoral da Paraíba; Tribunal Regional Eleitoral de Alagoas; Tribunal Regional Eleitoral de Goiás; Tribunal Regional Eleitoral de Mato Grosso do Sul; Tribunal Regional Eleitoral de Minas Gerais; Tribunal Regional Eleitoral de Pernambuco; Tribunal Regional Eleitoral de Rondônia; Tribunal Regional Eleitoral de Roraima; Tribunal Regional Eleitoral de Santa Catarina; Tribunal Regional Eleitoral de São Paulo; Tribunal Regional Eleitoral do Acre; Tribunal Regional Eleitoral do Amapá; Tribunal Regional Eleitoral do Amazonas; Tribunal Regional Eleitoral do Ceará; Tribunal Regional Eleitoral do Distrito Federal; Tribunal Regional Eleitoral do Espírito Santo; Tribunal Regional Eleitoral do Maranhão; Tribunal Regional Eleitoral do Mato Grosso; Tribunal Regional Eleitoral do Pará; Tribunal Regional Eleitoral do Paraná; Tribunal Regional Eleitoral do Piauí; Tribunal Regional Eleitoral do Rio de Janeiro; Tribunal Regional Eleitoral do Rio Grande do Norte; Tribunal Regional Eleitoral do Rio Grande do Sul; Tribunal Regional Eleitoral do Sergipe; Tribunal Regional Eleitoral do Tocantins; Tribunal Regional Federal da 1ª Região; Tribunal Regional Federal da 2ª Região; Tribunal Regional Federal da 3ª Região; Tribunal Regional Federal da 4ª Região; Tribunal Regional Federal da 5ª Região; Tribunal Superior do Trabalho; Tribunal Superior Eleitoral; Universidade da Integração Internacional da Lusofonia Afro -Brasileira; Universidade Federal da Fronteira Sul; Universidade Federal da Integração Latino - Americana; Universidade Federal da Paraíba; Universidade Federal de Alagoas; Universidade Federal de Alfenas; Universidade Federal de Campina Grande; Universidade Federal de Goiás; Universidade Federal de Itajubá; Universidade Federal de Juiz de Fora; Universidade Federal de Lavras; Universidade Federal de Minas Gerais; Universidade Federal de Pelotas; Universidade Federal de Pernambuco; Universidade Federal de Roraima; Universidade Federal de Santa Catarina; Universidade Federal de Santa Maria; Universidade Federal de São Paulo; Universidade Federal do Cariri; Universidade Federal do Ceará; Universidade Federal do Espírito Santo; Universidade Federal do Estado do Rio de Janeiro; Universidade Federal do Oeste da Bahia; Universidade Federal do Oeste do Pará; Universidade Federal do Pará;

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Universidade Federal do Paraná; Universidade Federal do Recôncavo da Bahia; Universidade Federal do Rio de Janeiro; Universidade Federal do Rio Grande do Norte; Universidade Federal do Rio Grande do Sul; Universidade Federal do Sul da Bahia; Universidade Federal do Sul e Sudeste do Pará; Universidade Federal do Triângulo Mineiro; Universidade Federal dos Vales do Jequitinhonha e Mucuri; Universidade Federal Fluminense; Universidade Federal Rural da Amazônia; Universidade Federal Rural de Pernambuco; Universidade Federal Rural do Rio de Janeiro; Universidade Federal Rural do Semiárido; Universidade Tecnológica Federal do Paraná; Valec Engenharia Construções e Ferrovias S/a tagTitInteressados tagInteressados Representação legal: Álcio Antônio Lopes Guimarães (29.946/OAB-RS), Vitto Giancristoforo dos Santos (58566/OAB-RS) e outros, representando Hospital Nossa Senhora da Conceição S.a.; Leonor Chaves Maia de Sousa, Arnaldo de Morais Moreira Fernandes Vieira, Danielle Gonçalves e Silva, Ari Barbosa Ferreira e Ernesto Lima Cruz, representando Banco do Nordeste do Brasil S.A..

SUMÁRIO: LEVANTAMENTO DE AUDITORIA. INTELIGÊNCIA ARTIFICIAL. IDENTIFICAÇÃO DE OPORTUNIDADES ADOÇÃO DA TECNOLOGIA. EXEMPLOS DE USO DE IA NA ADMINISTRAÇÃO FEDERAL. AVALIAÇÃO DA ESTRATÉGIA BRASILEIRA DE INTELIGÊNCIA ARTIFICIAL. POSSÍVEIS IMPACTOS PARA O CONTROLE. LEVANTAR O SIGILO DOS AUTOS. ARQUIVAMENTO.

RELATÓRIO

Transcrevo o Relatório de Levantamento elaborado pela Secretaria de Fiscalização de TI (Sefti), à peça 328, e que contou com a anuência do corpo diretivo da unidade (peças 329 e 330):

1 Introdução

2 Deliberação que originou a fiscalização

1. O Levantamento em tela foi autorizado por meio de despacho do Ministro Relator Aroldo Cedraz, de 10/2/2021, decorrente de proposta de fiscalização apresentada pela Sefti, no âmbito do TC 036.797/2020-0 (Administrativo, peça 5).

3 Identificação do objeto

2. O objeto do presente levantamento são as tecnologias de Inteligência Artificial (IA) nas organizações da Administração Pública Federal (APF) em suas diversas formas de utilização.

4 Objetivo, Escopo e Questões de Auditoria

3. O objetivo deste trabalho é fazer um levantamento sobre o estágio atual da utilização de tecnologias de IA por diversas organizações que compõem a Administração Pública, sob vários aspectos.

4. O primeiro deles diz respeito a conhecer as principais tecnologias utilizadas no desenvolvimento de soluções de IA, e avaliar o estágio atual e as perspectivas de sua utilização nas organizações públicas, além de identificar os principais riscos relacionados ao tema.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

5. O segundo objetivo refere-se ao conhecimento do estágio atual de desenvolvimento e o conteúdo da proposta para uma Estratégia Brasileira de Inteligência Artificial, e identificar os principais riscos e oportunidades associados à sua implementação.

6. O escopo do trabalho abrangeu a utilização de IA nas organizações dos três poderes da República, tendo sido analisados somente aspectos em relação à IA enquanto tecnologias de aprendizado de máquina, conforme explicado mais detalhadamente adiante, sendo que não foram consideradas as soluções desenvolvidas por organizações da área privada.

7. Por fim, o levantamento buscou verificar de que forma o controle das soluções de IA desenvolvidas no âmbito da APF pode ser exercido, e como as tecnologias de IA podem auxiliar a atividade de controle.

8. Nesse sentido, elaboraram-se as seguintes questões de auditoria:

a. Quais as principais tecnologias de IA e seus benefícios para aplicação pela APF?

b. Qual o cenário atual de utilização da IA na APF?

c. Quais dificultadores e riscos devem ser considerados pelos gestores na implementação de soluções de IA na APF?

d. Qual o estágio atual da implementação da Estratégia Brasileira de IA e os riscos e oportunidades referentes à sua implementação?

e. Quais metodologias podem ser aplicadas pelas equipes de auditoria para fiscalizar soluções de IA na APF?

f. Como a IA pode ser utilizada para apoiar a atividade de controle?

5 Metodologia utilizada

9. Primeiramente, registra-se que os trabalhos foram realizados em conformidade com as Normas de Auditoria do Tribunal de Contas da União (NAT), definidas na Portaria-TCU 168/2011, e com o Roteiro de Levantamento, estabelecido na Portaria-Segecex 5/2021.

10. A presente auditoria foi realizada na modalidade de teletrabalho, na forma que foi direcionada pelas boas práticas desenhadas pelo TCU para auditorias virtuais utilizando as ferramentas já disponíveis (https://tcucloud.sharepoint.com/sites/praticasremotasauditoria).

11. Para refinar a matriz de planejamento que foi elaborada pela equipe de auditoria, foram realizadas entrevistas com especialistas, em que foram colhidas sugestões para as questões de auditoria, além de identificar organizações relevantes no uso de IA no país.

12. A fiscalização baseou-se em entrevistas com gestores para conhecer os principais casos de uso e funcionalidades da tecnologia no setor público. Além disso, verificou-se a existência de trabalhos realizados por vários organismos internacionais que abrangem diversos dos objetivos presentes neste levantamento, servindo assim de importante fonte para o trabalho. Também se utilizou a literatura referente ao tema, bem como fontes acadêmicas, conforme elencadas na seção "Referências".

13. Além disso, a equipe da fiscalização elaborou questionário eletrônico que foi encaminhado a 293 organizações abrangidas no presente escopo, com o objetivo de colher detalhes sobre como a IA é utilizada para resolver problemas de negócio e agregar valor nos serviços prestados aos cidadãos.

14. A equipe também entrevistou especialistas da área, tanto do meio acadêmico como da inciativa privada, com o objetivo de subsidiar a elaboração das informações constantes neste Levantamento.

6 Limitações

15. Como limitações do trabalho, citam-se o grande número de instituições abrangidas no levantamento e a dificuldade de obtenção de resposta ao questionário eletrônico por todos os envolvidos. Além disso, referente à avaliação da Estratégia Brasileira de Inteligência Artificial (EBIA), a recriação do Ministério das Comunicações em 2020, a qual motivou a remoção de vários dos servidores do MCTI que estavam encarregados da elaboração da referida estratégia para o novo ministério, trouxe dificuldades na obtenção de informações referentes ao tema.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

7 Benefícios Estimados da Fiscalização

16. Consideram-se os benefícios da presente fiscalização:

a. Fornecer subsídios para aprimorar o conhecimento dos gestores públicos federais em relação aos benefícios e critérios para a utilização de tecnologias de IA na disponibilização de serviços à sociedade e melhoria dos processos internos nas organizações públicas;

b. Estabelecer uma linha de base da situação atual de utilização de tecnologias de IA na Administração Pública Federal, de forma a subsidiar trabalhos futuros que busquem aferir a evolução do tema;

c. Dar subsídios para o aprimoramento da Estratégia Brasileira de Inteligência Artificial (EBIA), contribuindo para que alcance seus objetivos, de maneira legítima, econômica, eficaz e eficiente e, assim, possa atender à sociedade;

d. Aquisição de conhecimento pelo TCU nas metodologias e *frameworks* aplicáveis à fiscalização de soluções de IA.

8 Visão Geral do Objeto

17. Esta seção tem o intuito de conceituar o que é IA e apresentar suas principais técnicas, além de citar as principais áreas de aplicação e explicar as principais questões relativas ao desenvolvimento de soluções de IA.

18. Tendo em vista que existe farto material que discorre sobre o assunto em fontes públicas e na internet, não será feito um aprofundamento maior nas subseções a seguir, ficando o detalhamento sobre o tema a cargo dos apêndices deste relatório.

9 O que é Inteligência Artificial[1]

19. Ao longo do tempo, muitas definições foram dadas para o termo Inteligência Artificial, e a associação do mesmo a outros termos como aprendizado de máquina (*machine learning*) e aprendizado profundo (*deep learning*) tem gerado dificuldades para uma melhor compreensão sobre o tema.

20. O aspecto artificial da IA é bastante simples: refere-se a qualquer coisa não natural e, neste caso, feito pelo homem. Também pode ser representado por meio do uso de termos como máquinas, computadores ou sistemas. Já inteligência é um conceito muito mais amplamente contestado, explicando por que ainda não há consenso sobre como definir IA (Miaihle and Hodes, 2017).

21. Inteligência Artificial pode ser definida como o uso de tecnologia digital para criar sistemas capazes de realizar tarefas comumente pensadas para exigir inteligência.

22. Nesse contexto, cita-se a definição da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), que trata IA como um sistema baseado em máquina que pode, para um determinado conjunto de objetivos definidos pelo homem, fazer previsões, recomendações ou decisões que influenciam ambientes reais ou virtuais.

23. Dessa forma, pode-se falar que a IA atual envolve principalmente máquinas usando estatísticas para encontrar padrões em grandes quantidades de dados e realizar tarefas repetitivas com dados sem a necessidade de constante orientação humana. Assim, a IA não consiste numa solução tecnológica aplicável para todos os casos, uma vez que, em geral, não apresenta boa performance sem que haja uma quantidade de dados significativa, relevante e com boa qualidade.

10 IA Forte e IA Fraca

24. O termo IA Forte refere-se à ideia de que a inteligência humana geral - que abrange diferentes domínios e capacidades - pode ser igualada ou até superada por máquinas. A partir desse conceito, várias suposições são levantadas no sentido de que as crenças desse sistema possam não estar alinhadas

[1] O resumo apresentado nesta seção é baseada no documento "*Hello, World: Artificial intelligence and its use in the public sector*", publicado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE) (2019) e disponível em https://www.oecd.org/gov/innovative-government/working-paper-hello-world-artificial-intelligence-and-its-use-in-the-public-sector.htm?utm_content=buffer2b1a0&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

às humanas. Tal representação de IA na mídia de massa e na cultura popular capturaram a imaginação do público e podem contribuir para a desconfiança da tecnologia impulsionada por IA, embora haja opiniões mistas entre pesquisadores de IA quanto à probabilidade de tal cenário.

25. Já a IA fraca está menos preocupada com a criação de uma superinteligência unificada, mas aceita e utiliza a noção de que humanos e computadores têm diferentes pontos fortes e competências. A IA tira proveito do fato de que os computadores são excelentes no desempenho quando se trata de processamento rápido e consistente de grandes quantidades de dados, bem como na execução de tarefas com base em dados lógicos e regras explícitas, enquanto os humanos ainda são mais eficientes em lidar com situações ambíguas ou aquelas que requerem intuição, criatividade, emoção, julgamento e empatia. Atualmente, na prática, praticamente toda a IA utilizada é a IA fraca.

26. O presente levantamento tem seu escopo limitado às aplicações da IA Fraca, e mais especificamente no campo das tecnologias associadas à Aprendizagem de Máquina, apresentadas a seguir.

11 Aprendizagem de Máquina[2]

27. Nos primeiros dias da Inteligência Artificial, rapidamente foram resolvidos problemas que são intelectualmente difíceis para humanos, mas relativamente simples para computadores – problemas que podem ser descritos como uma lista de regras formais, matemáticas. O desafio real para a IA provou-se ser resolver as tarefas que são fáceis para as pessoas, porém difíceis para que elas as descrevam formalmente – problemas que resolvemos de forma intuitiva e automática, como reconhecer palavras pronunciadas ou rostos em imagens.

28. Muitos dentre os primeiros sucessos da Inteligência Artificial ocorreram em ambientes relativamente formais e não requeriam que computadores tivessem muito conhecimento sobre o ambiente. Um exemplo é o jogo de xadrez – ainda na década de 1990, máquinas já eram capazes de vencer humanos. Porém, derivar uma estratégia bem-sucedida para jogar xadrez requer tão somente que se descreva completamente uma grande lista de regras formais, o que pode ser codificado previamente por um programador. Nesse caso, por exemplo, o desafio não está em ser capaz de descrever o ambiente – composto basicamente por um conjunto limitado de peças que podem mover-se apenas de formas rigidamente circunscritas.

29. As tarefas diárias de uma pessoa requerem um conhecimento imenso sobre o mundo ao seu redor. Computadores precisam capturar esse mesmo conhecimento para se comportarem de forma inteligente. Um dos desafios na IA é como representar tal conhecimento formalmente para um computador.

30. Vários projetos de inteligência artificial buscaram codificar o conhecimento sobre o mundo em linguagens formais. Um computador pode raciocinar automaticamente em cima de linguagens formais utilizando regras lógicas de inferência, o que é conhecido como abordagem de base de conhecimento. Nenhum desses projetos conduziu a um maior sucesso.

31. As dificuldades enfrentadas por sistemas ao confiarem em conhecimento codificado sugerem que sistemas de IA precisam da habilidade de adquirirem seu próprio conhecimento, extraindo padrões a partir de dados brutos. Tal capacidade é conhecida como aprendizagem de máquina, ou *machine learning*. A introdução da aprendizagem de máquina permitiu que computadores tratassem problemas envolvendo conhecimento do mundo real e tomassem decisões que pareciam subjetivas. A figura abaixo ilustra a diferença entre a programação clássica e a aprendizagem de máquina.

[2] O resumo apresentado nesta seção é baseado em tradução livre do livro (Goodfellow, Bengio, & Courville, *Deep Learning*, 2016).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

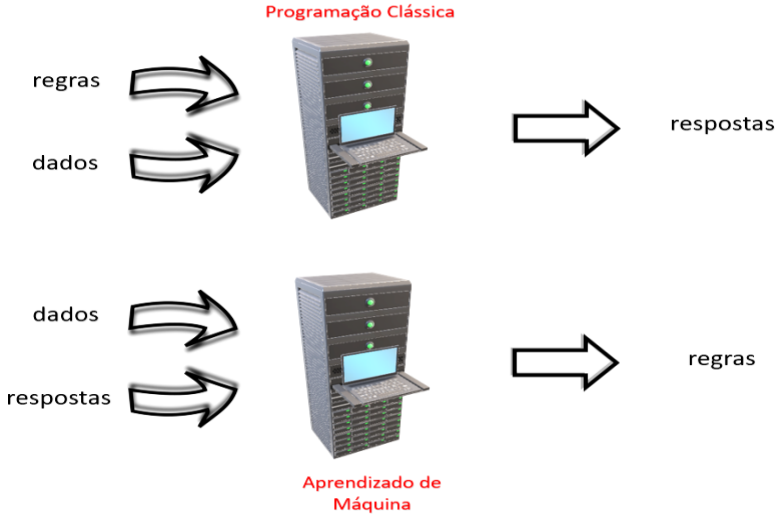


Figura 1 - Diferença entre programação clássica e aprendizado de máquina

32. Como visto acima, o conceito de aprendizado de máquina é complexo e engloba diversos aspectos técnicos. Para fins desse trabalho, será utilizada a definição formal da OCDE:

> O aprendizado de máquina é uma abordagem em que as máquinas aprendem a fazer previsões em novas situações baseadas em dados históricos. O aprendizado de máquina consiste em um conjunto de técnicas para permitir que as máquinas aprendam de maneira automatizada, sem instruções explícitas de um ser humano, por confiar em padrões e inferências. Aprendizado de máquina se aproxima frequentemente de ensinar as máquinas a chegarem a um resultado, mostrando-lhes muitos exemplos de resultados - chamados de "treinamento". Outra abordagem é os humanos definirem um conjunto de amplas regras e geralmente permitem que a máquina aprenda por conta própria por tentativa e erro.

12 Tipos de Aprendizagem de Máquina

Aprendizagem Supervisionada

33. Tendo em vista o paradigma de aprendizagem a partir de dados, considera-se aprendizagem supervisionada quando a máquina, para que seja capaz de fazer predições a partir de informações de entrada, precisa ter sido previamente treinada com exemplos nos quais, além das informações de entrada, as informações de saída esperadas também são especificadas, em geral, por um supervisor humano. Trata-se, por exemplo, do aprendizado a partir de dados históricos. Assim, quanto maior a quantidade de dados ou exemplos de treinamento, melhor o desempenho preditivo esperado do algoritmo.

34. Ao mesmo tempo, o algoritmo ou modelo resultante da etapa de treinamento deverá ter também capacidade de generalização, ou seja, deverá ser capaz de fornecer resultados corretos, próximos do esperado no mundo real, para dados que não tenham sido vistos anteriormente. Esses dados poderão, por exemplo, ser parecidos com os dados de treinamento, porém não serão idênticos. Aqui, o termo "parecidos" refere-se, tecnicamente, a pertencerem a distribuições de probabilidade semelhantes.

Aprendizagem não Supervisionada

35. Em situações em que há uma ampla quantidade de dados (ex.: informações de perfil de clientes de uma loja virtual), porém não há clareza sobre um valor-alvo a ser previsto, ou mesmo quando não há histórico de dados suficiente para que se possa definir o valor-alvo correto, ainda assim é possível extrair algum grau de informação útil a partir dos dados. É neste contexto que surge a aprendizagem não supervisionada, voltada para a identificação de padrões nos dados. Exemplos de padrões incluem agrupamentos (ex.: grupos de usuários semelhantes entre si, com algumas características em comum), detecção de anomalias (ex.: *outliers*) e outros tipos de padrão que podem ser úteis ao menos em uma etapa de análise exploratória dos dados.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

36. Em outras palavras, na aprendizagem de máquina não supervisionada não há a necessidade de uma supervisão humana para dizer à máquina o que especificamente procurar.

Aprendizagem por reforço

37. Entre todas as formas de aprendizado de máquina, o aprendizado por reforço é o que mais se aproxima da maneira com que os humanos e os outros animais aprendem. A abordagem neste tipo de aprendizado é de que não há uma resposta sobre como atingir determinado objetivo e, portanto, o sistema de IA terá que observar o estado do ambiente ao seu redor, agir sobre ele e colher a recompensa sobre sua ação para que possa ir aprendendo de forma interativa a descobrir qual a melhor maneira de maximizar a soma de suas recompensas no longo prazo e, com isso, atingir seu objetivo (Sutton & Barto, 2018).

38. A interação do sistema de IA com o ambiente faz com que aprenda a mapear estados observados em ações a serem executadas de acordo com sua probabilidade. Esse mapeamento é a política (*policy*) de comportamento do sistema de IA, uma regra de tomada de decisão.

39. Em geral, cada uma das três categorias de aprendizagem de máquina citadas tem sua aplicação voltada para a resolução de determinados tipos de problema. A figura abaixo ilustra as principais aplicações:

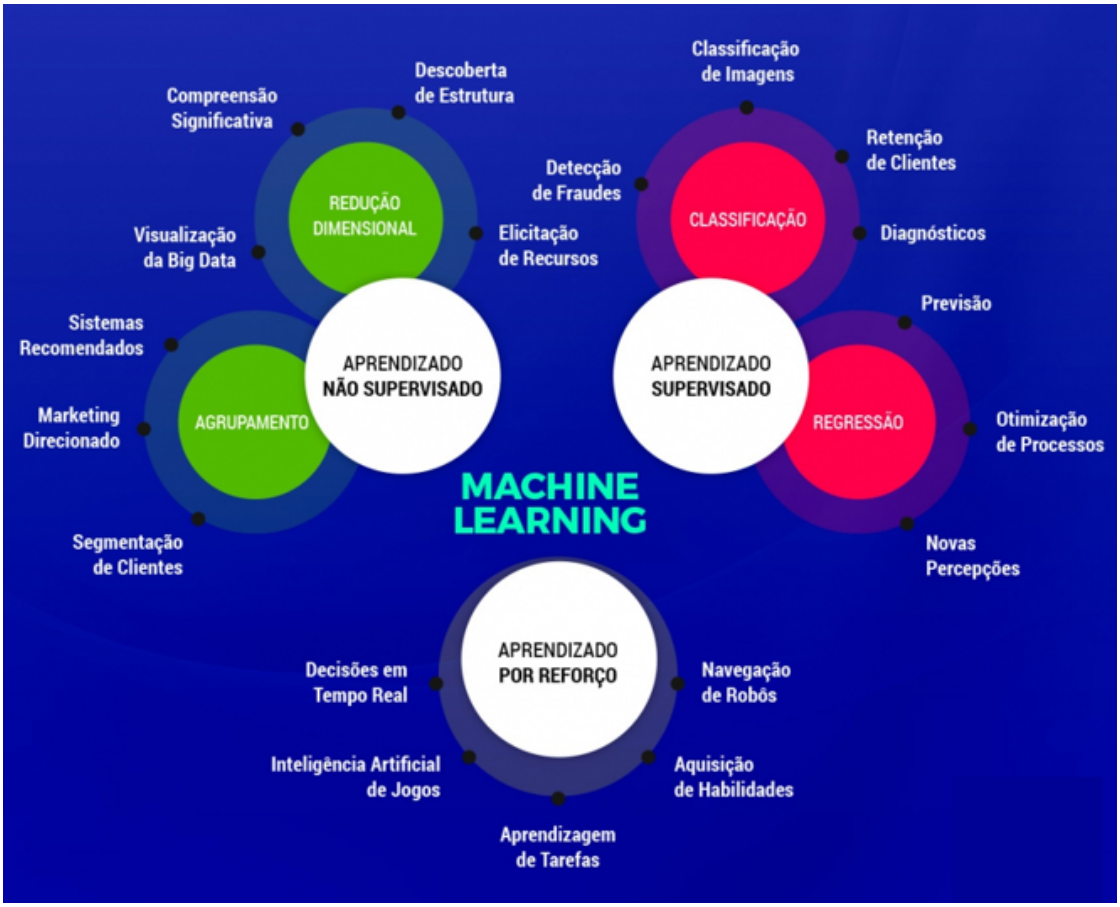


Figura 2 – Principais aplicações das tecnologias de aprendizagem de máquina[3]

16 Técnicas de aprendizagem de máquina

40. Não há uma tecnologia isolada que represente a IA. A exemplo, dentro de cada um dos tipos de aprendizagem de máquina identificados na secção anterior (supervisionada, não supervisionada e por reforço) existem diferenciados algoritmos e tecnologias que podem ser aplicadas, a depender dos objetivos a serem alcançados.

[3] Fonte: https://www.programaria.org/quais-sao-os-tipos-de-aplicacoes-de-inteligencia-artificial-mais-comuns/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

41. Nesse sentido, pode-se agrupar as técnicas de aprendizagem de máquina nos seguintes grupos:

Tabela 1 – Principais técnicas de Aprendizagem de Máquina[4]

Técnica	Descrição	Exemplos de aplicação
Classificação	Aprende as características de uma dada categoria, permitindo que o modelo de IA classifique dados desconhecidos em categorias pré-determinadas	• decidir se uma remessa de mercadorias passa por inspeção de fronteira; • decidir se um e-mail é *spam* ou não
Regressão	Prevê um valor de ponto de dados desconhecido	• prever o valor de mercado de uma casa a partir de informações como tamanho, localização ou idade, prevendo as concentrações de poluentes atmosféricos nas cidades
Agrupamento	Identifica grupos de dados similares em um arquivo	• agrupar clientes de varejo para encontrar subgrupos com hábitos de consumo específicos; • agrupar dados de medidores inteligentes para identificar grupos de aparelhos elétricos e gerar contas de eletricidade discriminadas
Redução de dimensionalidade	Limita os dados às variáveis mais relevantes para tornar os modelos mais precisos ou possibilitar a visualização dos dados	• usado por cientistas de dados ao avaliar e desenvolver outros tipos de algoritmos de aprendizado de máquina
Ranqueamento	Treina um modelo de IA para classificar novos dados com base em listas vistas anteriormente	• retornar páginas por ordem de relevância quando um usuário pesquisa em um *website*

42. Para implementar cada uma das técnicas de aprendizagem de máquina citadas na Tabela 1, existem inúmeros algoritmos já desenvolvidos e implementados. O Apêndice I apresenta alguns desses algoritmos. Apesar de sua variabilidade, um fato comum a todos os algoritmos de aprendizagem de máquina é que eles necessitam essencialmente de dados para produzir seus resultados. Nesse contexto, tem-se que os dados podem estar representados de forma estruturada ou mesmo na forma de imagens ou sons. Assim, dependendo do formato dos dados, um ou outro algoritmo será mais aplicável.

43. Uma classe de algoritmos que vem ganhando cada vez mais destaque e aplicação são as Redes Neurais Artificiais e o Aprendizado Profundo (*Deep Learning*). As redes neurais são classificadas como métodos supervisionados do aprendizado de máquina e são basicamente um conjunto de funções matemáticas que simulam o funcionamento do cérebro humano, através da simulação de seus neurônios e de suas ligações. O Apêndice II traz um maior detalhamento sobre essas tecnologias.

17 Principais domínios de aplicação da IA

44. A IA está se tornando muito presente no dia a dia dos cidadãos e cada vez mais a tecnologia é incorporada a programas de computador, sítios na internet e dispositivos conectados e vestíveis.

45. Embora a tecnologia atual de IA ainda tenha limitações significativas, ela está evoluindo rapidamente e vem trazendo significantes benefícios em certos domínios de aplicação. Entre eles, destacam-se os modelos preditivos, o Processamento da Linguagem Natural (PLN) e a Visão Computacional, conforme ilustrado na figura abaixo:

[4] Fonte: "*A guide to using artificial intelligence in the public sector" – Office for Artificial Intelligence – Goverment Digital Service – UK*. Disponível em: https://www.gov.uk/government/publications/a-guide-to-using-artificial-intelligence-in-the-public-sector

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

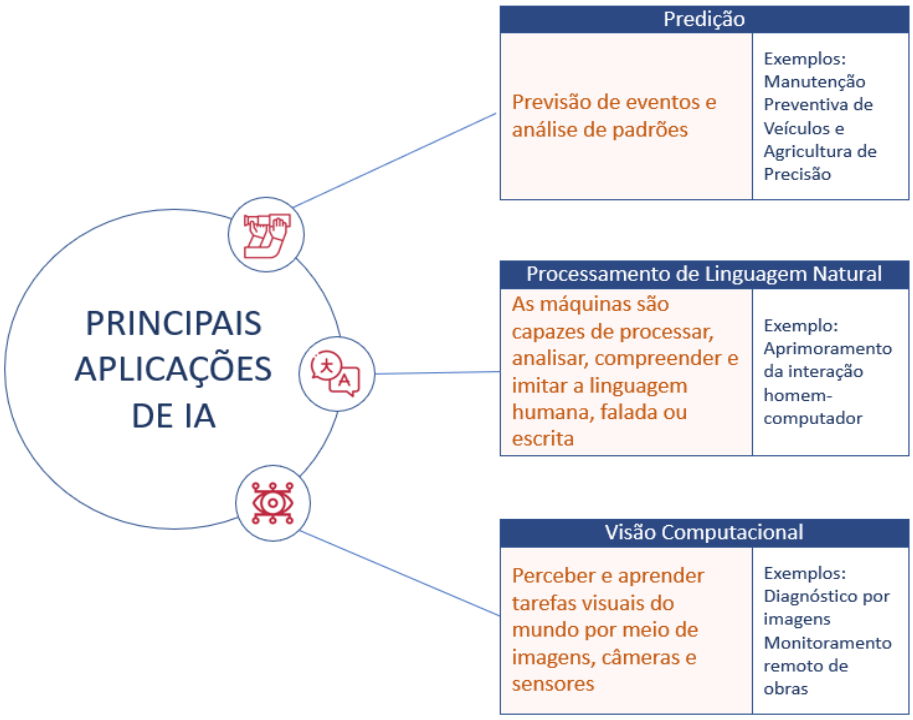


Figura 3 - Principais domínios de aplicação de IA

46. O Apêndice III apresenta um detalhamento sobre cada um desses domínios.

18 Ciclo de desenvolvimento de soluções de IA

47. O rápido crescimento das soluções que utilizam tecnologias de IA está tornando os sistemas cada vez maiores e complexos e os recentes avanços vêm fazendo com que estes sistemas sejam vistos como a base de uma futura geração de novos sistemas com inteligência. O fator mais importante é que se espera mais sucesso agora do que antes, pelo fato do grande interesse da comunidade científica. Com este advento, é importante que se tenham métodos e ferramentas que sirvam de auxílio nas etapas de desenvolvimento.

48. Dessa forma, já se desenvolveram várias propostas e modelos que endereçam o ciclo de vida de uma solução de IA. Dentre as metodologias existentes destacam-se as seguintes: KADS (*Knowledge Acquisition and Design Support*), MIKE (*Model-Based and Incremental Knowledge Engineering*), KOD (*Knowledge Oriented Design*), Vital, Desire, MKSM (*Methodology for Knowledge System Management*), e CYGMA (*Cycle de vie e Gestion des Métiers et des Applications*).

49. Porém, a maioria das soluções propostas para projetos que envolvem Aprendizagem de Máquina, apresenta, em resumo, as fases expostas na figura abaixo:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

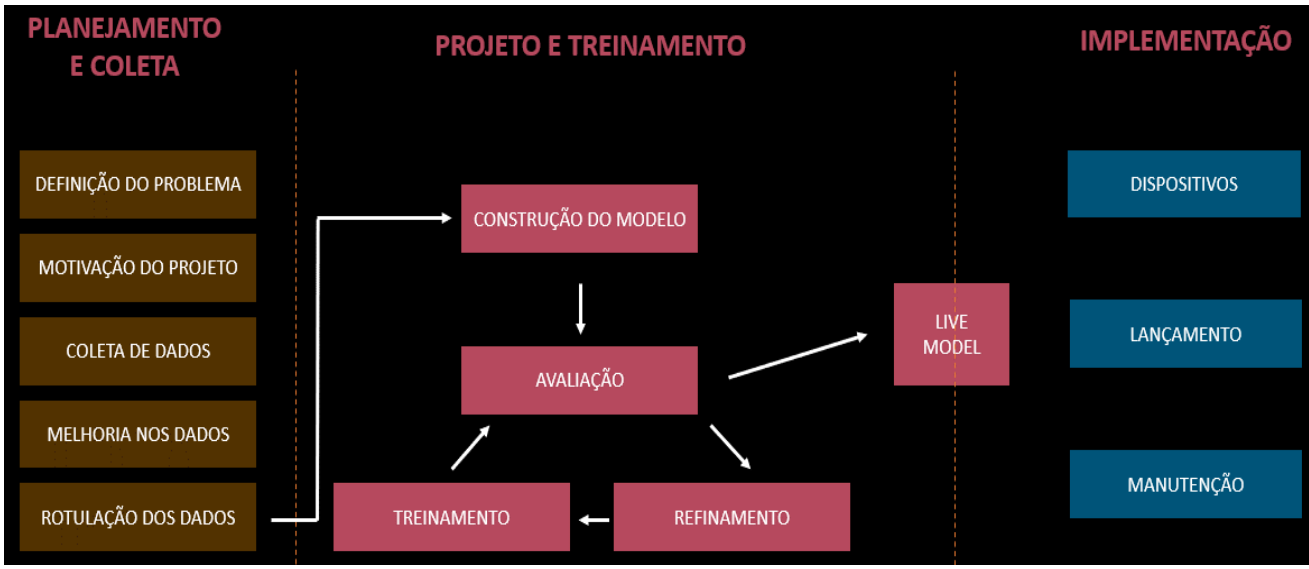


Figura 4 - Ciclo de desenvolvimento de um projeto de IA[5]

50. Em geral, todos os modelos de desenvolvimento abordam em sua fase inicial aspectos referentes à aquisição de dados a serem utilizados na elaboração dos modelos de aprendizagem de máquina. A construção do modelo é um processo no qual se busca desenvolver a solução com a maior acurácia e otimização do uso dos recursos computacionais. Por fim, após implementado o modelo num ambiente de produção, ele deve ser monitorado continuamente quanto ao seu desempenho. O Apêndice IV apresenta de forma mais detalhada as etapas do processo de desenvolvimento das soluções de IA.

19 O que considerar no desenvolvimento e uso de IA

51. Além de questões relativas à disponibilidade suficiente de dados, qualificação das equipes técnicas e aspectos operacionais de implementação, as soluções de IA demandam ainda aos gestores uma série de questões a serem consideradas na utilização dessa tecnologia e na busca do atingimento dos melhores resultados.

52. No documento "*A guide to using artificial intelligence in the public sector*", do Escritório para Inteligência Artificial do Serviço de Governo Digital do Reino Unido, são identificados aspectos relevantes que os gestores devem observar em relação ao desenvolvimento e à implementação de soluções de IA nas organizações públicas, destacando:

a) **qualidade de dados:** o sucesso do projeto de IA depende da qualidade dos dados utilizados no processo de aprendizagem de máquina;

b) **justiça** *(fairness)*: os modelos de IA devem ser treinados e testados utilizando conjuntos de dados relevantes, precisos e generalizáveis e os sistemas de IA devem ser implementados por usuários devidamente capacitados;

c) ***accountability:*** deve ser considerado quem é responsável por cada elemento da saída do modelo e como é atribuída a responsabilidade aos designers e implementadores dos sistemas de IA;

d) **privacidade:** a utilização dos dados deve estar em conformidade com a legislação e as políticas vigentes, a exemplo da Lei Geral de Proteção dos Dados (LGPD);

e) **interpretabilidade e transparência:** as partes interessadas devem poder entender como o modelo de IA chegou à sua decisão;

f) custos: deve ser considerado quanto custará para construir, executar e manter uma infraestrutura de IA, treinar e educar a equipe e se o trabalho para instalar AI pode superar qualquer economia potencial.

[5] Disponível em https://labelyourdata.com/articles/lifecycle-of-an-ai-project-stages-breakdown/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

53. Um outro exemplo de princípios referentes ao desenvolvimento e implementação de soluções que envolvem tecnologias de IA são elencados pelo *The Institute for Ethical AI and Machine Learning*[6]. Para cada um dos princípios, são estabelecidos compromissos a serem observados pelos responsáveis pela solução de IA, conforme abaixo:

Tabela 2 - Princípios para o desenvolvimento de soluções de IA (tradução livre)

	Princípio	Compromisso
1	**Participação humana**	avaliar o impacto de previsões incorretas e, quando razoável, projetar sistemas com processos de revisão humanos
2	**Avaliação de viés**	desenvolver continuamente processos que permitam compreender, documentar e monitorar preconceitos no desenvolvimento e na produção.
3	**Explicabilidade por justificativa**	desenvolver ferramentas e processos para melhorar continuamente a transparência e a explicabilidade dos sistemas de aprendizado de máquina onde for razoável.
4	**Operações reproduzíveis**	desenvolver a infraestrutura necessária para permitir um nível razoável de reprodutibilidade em todas as operações de sistemas de ML.
5	**Estratégia de deslocamento**	identificar e documentar informações relevantes para que os processos de mudança de negócios possam ser desenvolvidos para mitigar o impacto para os trabalhadores que estão sendo automatizados.
6	**Precisão prática**	desenvolver processos para garantir que as funções de precisão e métrica de custo estejam alinhadas aos aplicativos específicos do domínio.
7	**Confiança pela privacidade**	construir e comunicar processos que protegem e tratam os dados com as partes interessadas que podem interagir com o sistema direta e/ou indiretamente.
8	**Conscientização de risco de dados**	desenvolver e melhorar processos e infraestrutura razoáveis para garantir que os dados e a segurança do modelo sejam levados em consideração durante o desenvolvimento de sistemas de aprendizado de máquina.

54. Considerações sobre o uso de IA no Setor Público também são abordadas no documento "**Uso de Inteligência Artificial pelo Poder Público - Recomendações de Governança**"[7], elaborado pela organização Transparência Brasil no âmbito do Projeto Transparência Algorítmica, o qual traz, em resumo, as seguintes recomendações:

a) as bases de dados utilizadas para a construção dos modelos devem ser representativas e apropriadas para o contexto na qual serão utilizadas;

b) necessidade de supervisão humana como salvaguarda para a revisão de decisões automatizadas;

c) efetiva proteção dos dados pessoais do cidadão; e

d) transparência e explicabilidade dos sistemas.

[6] Disponível em https://ethical.institute/principles.html#commitment-1

[7] Disponível em https://www.transparencia.org.br/projetos/transparencia-algoritmica

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

55. Nesse contexto, o Apêndice V detalha os principais aspectos concernentes à implementação de soluções de IA que devem ser considerados pelas organizações.

20 Utilização de Inteligência Artificial no governo

56. O setor público é responsável por definir as prioridades, investimentos e regulamentações nacionais no que diz respeito à IA. Dessa forma, os governos devem buscar alavancar as tecnologias de IA para inovar e transformar o setor público, a fim de redefinir as maneiras pelas quais as políticas e serviços públicos são projetados e implementados. Essa inovação e transformação são críticas para os governos, pois eles enfrentam uma complexidade e demandas cada vez maiores de seus cidadãos e organizações públicas e privadas.

57. Assim, a IA pode ser integrada a todo o processo de formulação de políticas e desenho de serviços. À medida que a IA e o aprendizado de máquina evoluem, mais tarefas administrativas e orientadas a processos podem ser automatizadas, aumentando a eficiência do setor público e liberando os funcionários públicos para se concentrarem em um trabalho mais significativo. Com o auxílio da IA, os governos também podem entender melhor a realidade, tomar decisões dentro de suas organizações e antecipar as necessidades da população. Se bem executados, os processos automatizados podem ajudar o governo a tomar decisões mais justas e precisas do que antes.

58. Nesse sentido, o relatório "*How much time and money can AI save government?*" elaborado pela Deloitte University Press[8] aponta os principais benefícios da aplicação de IA:

a) redução de gastos em decorrência de maior eficiência operacional: de acordo com a Deloitte, a automação das tarefas dos funcionários do governo federal americano poderia economizar entre 96,7 milhões e 1,2 bilhão de horas anuais. O mesmo estudo da Deloitte também relata que a automação e a IA têm potencial para economizar entre US $ 3,3 bilhões e US $ 41,1 bilhões;

b) disponibilização de novos serviços e aprimoramento dos já existentes: aplicativos voltados para o cidadão, como ônibus autônomos e educação personalizada, melhoram a qualidade dos serviços que os governos oferecem aos cidadãos;

c) mais tomadas de decisão baseadas em dados: os governos estão coletando uma grande quantidade de dados todos os dias. No entanto, sem uma análise precisa, os dados não são adequados para *insights* acionáveis. Uma melhor tomada de decisão tem o potencial de melhorar os serviços e economizar custos.

59. A esse respeito, a publicação *Government AI Readiness Index* 2020, elaborada pela Oxford Insights e pelo Governo do Canadá[9], apresenta ranking de 172 países baseado em 33 indicadores que apontam a capacidade dos governos nacionais para implementar soluções de IA no fornecimento de serviços públicos aos cidadãos. Além disso, a publicação contém um *ranking* adicional de 32 países quanto à possibilidade de uso responsável de IA pelos governos. Tal trabalho é uma evolução metodológica de estudos já realizados e publicados respectivamente nos anos de 2017 e 2019[10]. A figura abaixo apresenta resumidamente as dimensões que foram consideradas na avaliação realizada em 2020:

---

[8] Disponível em https://www2.deloitte.com/content/dam/insights/us/articles/3834_How-much-time-and-money-can-AI-save-government/DUP_How-much-time-and-money-can-AI-save-government.pdf

[9] Disponível em https://www.oxfordinsights.com/government-ai-readiness-index-2020

[10] Disponíveis em https://www.oxfordinsights.com/government-ai-readiness-index (2017) e https://www.oxfordinsights.com/ai-readiness2019 (2019)

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

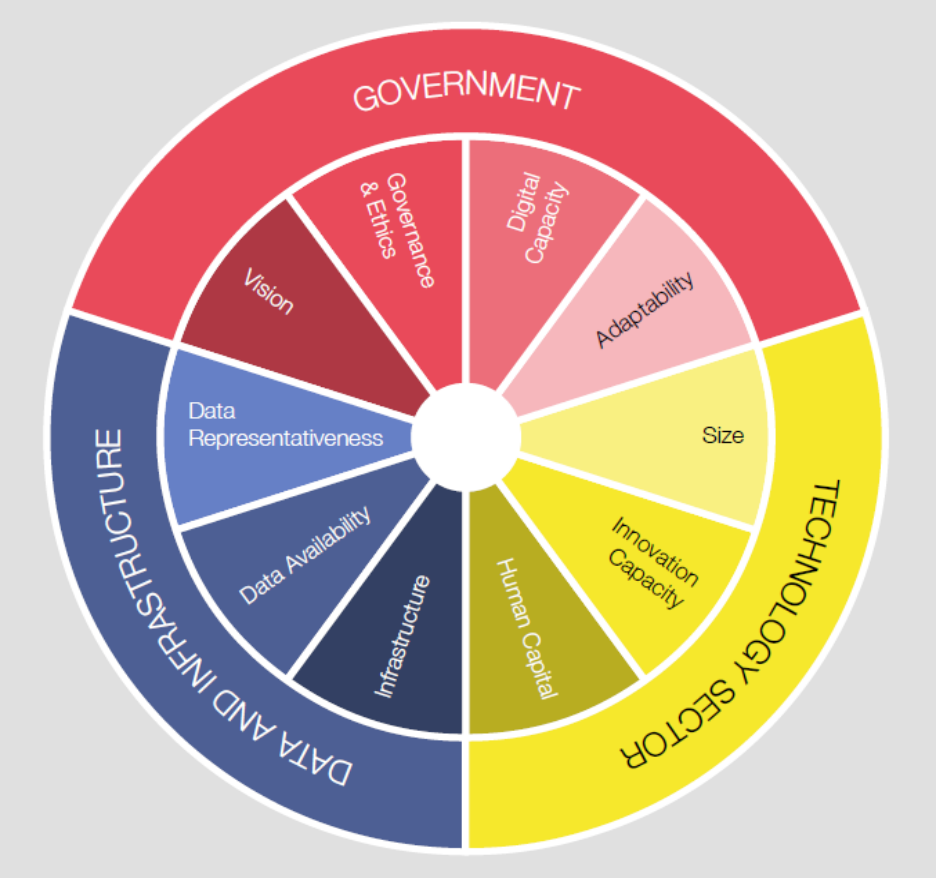


Figura 5 - Pilares e dimensões do *Government AI Readiness Index*

60. Conforme se observa, a avaliação considera três pilares (Governo; Setor de Tecnologia; e Dados e Infraestrutura), sendo que cada um dos pilares é dividido em dimensões de avaliação. O referido estudo apontou que os Estados Unidos tinham em 2020 o maior índice de capacidade de implementação de soluções de IA pelo governo, sendo que as próximas quatro posições no ranking são ocupadas por países da Europa Ocidental (Reino Unido, Finlândia, Alemanha e Suécia). As regiões de pontuação mais baixa, em média, são a África Sub-Saariana, a América Latina e o Caribe.

61. Ainda conforme o referido relatório, em 2020 o Brasil ocupava a 63ª posição no ranking mundial de capacidade do governo para implementar soluções de IA e o 6º na América Latina e Caribe. Já em relação ao uso responsável da IA, o Brasil ocupa a 30ª posição de 34 países avaliados. A figura abaixo detalha a pontuação atribuída ao Brasil nas dimensões avaliadas. Ressalta-se que o estudo é de 2020, quando o Brasil ainda não havia publicado sua estratégia para IA, o que ocorreu em 2021:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

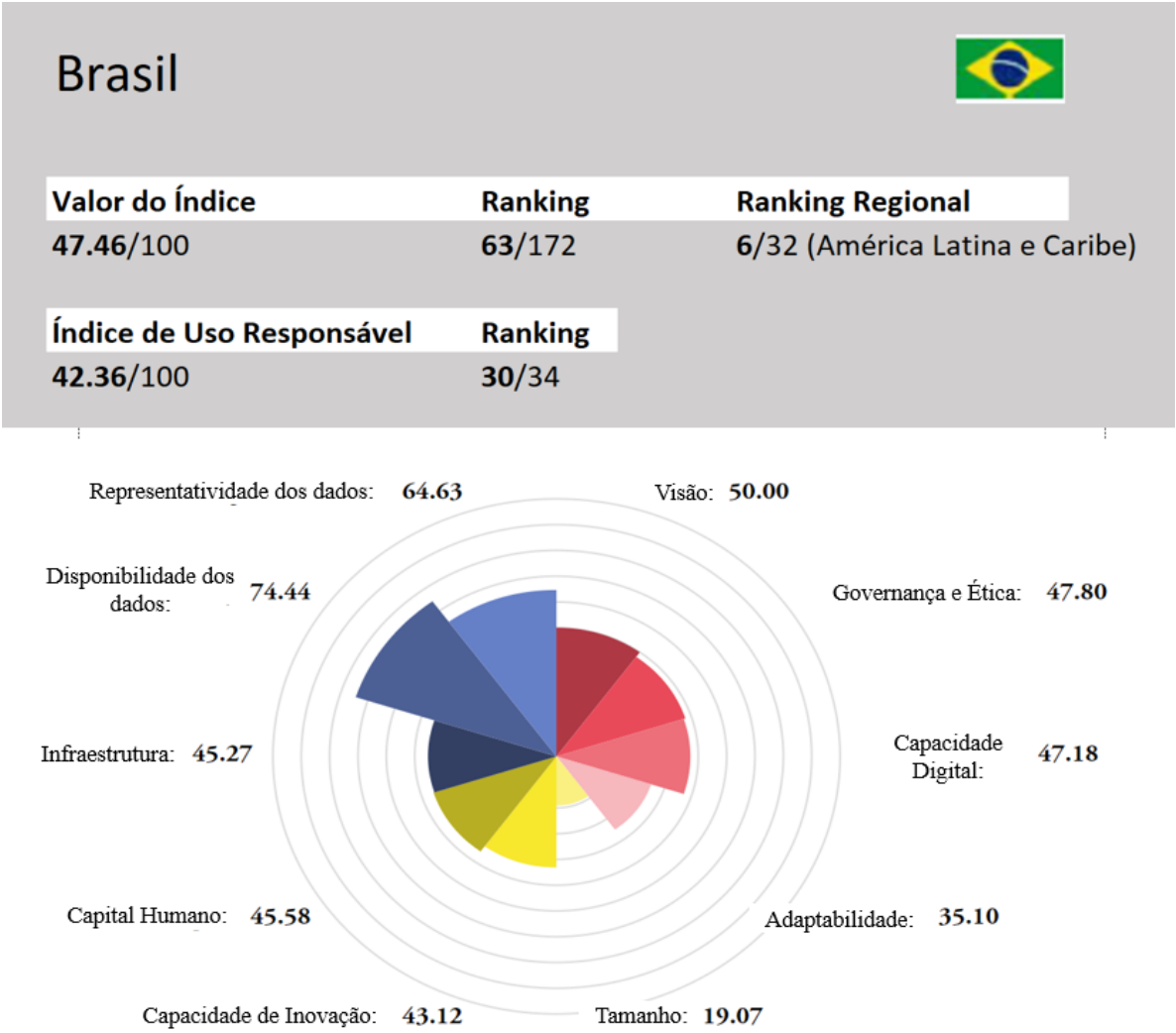


Figura 6 - Brasil – *Government IA Readiness Index*

62. A partir dos dados apresentados na figura anterior, verifica-se que os principais diferenciais para o incentivo à implementação de soluções de IA no Brasil são a disponibilidade e a representatividade de dados.

63. Nesse sentido durante o presente levantamento, buscou-se verificar os casos de uso de soluções de IA na APF, com o objetivo de se conhecer o estado atual de desenvolvimento, bem como de compartilhar no âmbito da APF as iniciativas de sucesso. Dentre as soluções implementadas, a equipe de planejamento destacou os seguintes exemplos, os quais são detalhados no Apêndice VIII:

Tabela 3 - Soluções de IA na APF

Organização	Solução	Objetivo	Principais Funcionalidades
Tribunal Superior Do Trabalho (TST)	Bem-Te-Vi	Gestão de acervo processual	• Triagem e encaminhamento de processos para analistas com base em modelo preditivo • Geração de modelos de instrução
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes)	Mimir	Desambiguação de dados	• Automatização do processo de preenchimento das plataformas de governo relacionadas à pesquisa científica • Apoio à integração entre

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Organização	Solução	Objetivo	Principais Funcionalidades
			sistemas e exposição de dados • Higienização, unificação e desambiguação de dados • Aprendizado ativo
Tribunal de Contas da União (TCU)	Zello	Chatbot[11] para levar informações do TCU para o cidadão em diversos canais e redes sociais	• Interação via portal do TCU • Interação via aplicativo móvel do TCU • Interação via redes sociais • Pergunta e resposta: dúvidas sobre funcionamento do TCU e como interagir e obter informações • Consultas processuais • Consultas a contas julgadas irregulares • Emissão de certidões
Agência Nacional de Energia Elétrica (Aneel)	Águia	Monitoramento de obras em construção de usinas solares fotovoltaicas utilizando sensoriamento remoto com imagens de satélite	• Identificação dos principais marcos de implantação dos empreendimentos de geração solar (início das obras e instalação dos painéis) • Apoiar a avaliação humana e decisão dos fiscais por meio de uso de ferramentas visuais (imagens de satélite e dashboards)
Tribunal Regional do Trabalho – 5ª Região	Gemini	Agrupamento de processos similares para instrução	• Agrupamentos de processos em pilhas que contém temas semelhantes • Apresentação de processos similares já julgados
Instituto Nacional de Metrologia, Qualidade e Tecnologia (Inmetro)	Identificação de irregularidades através de imagens	Monitoramento de inspeções de equipamentos para transporte de produtos perigosos	• Avaliação da completude de registros fotográficos • Detecção de fotos clonadas
Tribunal Regional Federal – 3ª Região	Sigma e Sinara	Produtividade e padronização de instruções judiciais com reuso de modelos mais adequados às características do processo	• Seleção do modelo mais adequado para novo processo • Preenchimento de dados extraídos das peças processuais anteriores
Tribunal de Contas da União (TCU)	Assistente Conjur	Sugestão de teses para resposta a questionamentos judiciais e geração de minuta de pareceres	• Seleção de teses a partir de petições iniciais ou mandados de segurança • Geração de minuta de pareceres

[11] *Chatbot* é um programa de computador que tenta simular um ser humano na conversação com as pessoas. O objetivo é responder às perguntas de tal forma que as pessoas tenham a impressão de estar conversando com outra pessoa e não com um programa de computador.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Organização	Solução	Objetivo	Principais Funcionalidades
Secretaria do Tesouro Nacional	PropLegis	Identificação automática de padrões em proposições legislativas	• Rastreio diário de proposições na Câmara e no Senado para classificar se são de interesse do Tesouro Nacional • Classificação das proposições legislativas de interesse por área de negócio
Câmara dos Deputados	Plataforma Ulisses	Conjunto de algoritmos de inteligência artificial construídos para sustentar funcionalidades diversas necessárias à plataforma digital legislativa	• Funcionalidades diversas distribuídas entre vários sistemas, descritas em mais detalhes no Apêndice VIII

21 Questionário eletrônico sobre utilização de IA na Administração Pública Federal

64. Com o objetivo de obter um panorama da situação atual da utilização de soluções de Inteligência Artificial pelos diversos órgãos da Administração Pública Federal, a equipe do levantamento elaborou questionário eletrônico que tratou de vários aspectos em relação ao tema, cujo inteiro teor encontra-se no sítio https://portal.tcu.gov.br/levantamento-Inteligência Artificial.htm

65. O referido questionário foi encaminhado para 293 organizações conforme relação constante no Apêndice IX, das quais 263 organizações encaminharam as suas informações no prazo definido. Além destas, também se realizaram entrevistas com outras organizações que buscaram levantar as mesmas informações constantes no questionário eletrônico.

66. O questionário elaborado teve como principais objetivos:

a) obter informações de referência para compor uma linha de base sobre a utilização de IA pelos órgãos da APF em 2021, de modo a subsidiar futuros trabalhos do TCU que busquem acompanhar a evolução do tema na APF;

b) conhecer as principais tecnologias de IA que estão sendo utilizadas na APF, bem como seus domínios de aplicação;

c) conhecer casos de uso e experiências de sucesso na utilização de soluções de IA, com o objetivo de dar publicidade aos mesmos e incentivar o intercâmbio de experiências entre as organizações públicas;

d) levantar os principais dificultadores que os gestores públicos enfrentam para a utilização de IA na APF.

67. Ressalta-se que, durante os trabalhos de planejamento do presente levantamento, verificou-se que já haviam sido realizadas recentemente por outras organizações pesquisas que também se referiram à utilização de tecnologias de IA em organizações públicas, as quais detalham-se a seguir.

68. A primeira pesquisa identificada foi realizada em setembro de 2020 pela Transparência Brasil, em parceria com a Controladoria-Geral da União, o Ministério da Ciência, Tecnologia e Inovação (MCTI) e o Centro de Estudos sobre Tecnologias Web (Ceweb.br). O questionário foi enviado a 319 órgãos do Executivo Federal. Além disso, foram realizados pedidos de acesso à informação para órgãos do Legislativo e do Judiciário.

69. Tal trabalho identificou 44 ferramentas de IA usadas por órgãos governamentais, as quais foram classificadas segundo o público-alvo (interno e externo) e segundo o processo de tomada de decisão implementado nas mesmas, resultando no quadro abaixo:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

	COM TOMADA DE DECISÃO	SEM TOMADA DE DECISÃO	TOTAL
Público interno	20 ferramentas	16 ferramentas	36
Externo	8 ferramentas	0 ferramentas	8
Total	28	16	44

Figura 7 - Classificação de ferramentas de inteligência artificial segundo uso para tomada de decisão e público-alvo[12]

70. Como produto, a Transparência Brasil elaborou o Relatório "Uso de Inteligência Artificial pelo Poder Público – Recomendações de Governança", o qual tem como foco principal a análise dos riscos e impactos a direitos do cidadão que as soluções de IA implementadas podem trazer. Nesse sentido, o relatório apresenta quatro considerações de governança que sintetizam os principais riscos encontrados nas soluções de IA já disponibilizadas[1]: i) bases de dados representativas e apropriadas para o contexto; ii) necessidade de supervisão humana como salvaguarda para a revisão de decisões automatizadas; iii) efetiva proteção dos dados pessoais do cidadão; e iv) transparência e explicabilidade dos sistemas.

71. Outro estudo relevante sobre a utilização de IA na APF foi publicado em 2021 pelo Centro de Inovação, Administração e Pesquisa do Judiciário da Fundação Getúlio Vargas[13], e teve como objetivo geral pesquisar o estado da arte e as iniciativas nacionais de aplicação de tecnologia com a utilização de IA nos tribunais, que busquem aperfeiçoar o sistema de justiça sob a perspectiva de seu funcionamento e de seu aparato.

72. A pesquisa da FGV evidenciou os projetos de IA já implementados e em desenvolvimento nos tribunais superiores, tribunais regionais federais, tribunais regionais do trabalho e tribunais de justiça, conforme figura abaixo:

TRIBUNAIS	EM DESENVOLVIMENTO	PROJETO-PILOTO	EM PRODUÇÃO
Tribunais Superiores (STF, STJ e TST)	4	-	5
Tribunais Regionais Federais	8	2	6
Tribunais Regionais do Trabalho	5	1	1
Tribunais de Justiça	12	4	15
Total por fase	29	7	27

Figura 8 - Fase de implementação dos projetos de IA no Poder Judiciário[14]

[12] Fonte: Transparencia.Brasil. Disponível em https://www.transparencia.org.br/downloads/publicacoes/Recomendacoes_Governanca_Uso_IA_PoderPublico.pdf

[13] Disponível em https://ciapj.fgv.br/sites/ciapj.fgv.br/files/estudos_e_pesquisas_ia_1afase.pdf

[14] Fonte: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/estudos_e_pesquisas_ia_1afase.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

73. O relatório detalha todos os projetos de IA identificados nos diversos tribunais, apresentando para cada um deles a motivação do desenvolvimento, o estágio atual da solução, as funcionalidades presentes e problemas que buscam solucionar, além de apresentar os resultados obtidos. Conforme informações constantes no sítio da FGV, o projeto encontra-se em sua 2ª Fase, na qual o objeto de pesquisa foi ampliado para abranger tecnologias de "inteligência computacional" (https://ciapj.fgv.br/projetos).

74. Dessa forma, considerando-se as informações já constantes nos trabalhos mencionados acima, o questionário eletrônico aplicado pelo presente Levantamento buscou aprofundar e tem como diferencial os seguintes aspectos:

a) informações requeridas para avaliação do nível de maturidade da organização em relação à utilização de IA;

b) condições existentes referentes aos colaboradores;

c) aspectos gerais sobre o desenvolvimento de soluções de IA;

d) Informações técnicas gerais e detalhadas sobre as soluções desenvolvidas (resposta opcional).

75. Inicialmente, tem-se o seguinte quadro geral nas respostas das organizações respondentes quando perguntadas se já utilizam soluções que envolvem tecnologias de IA:

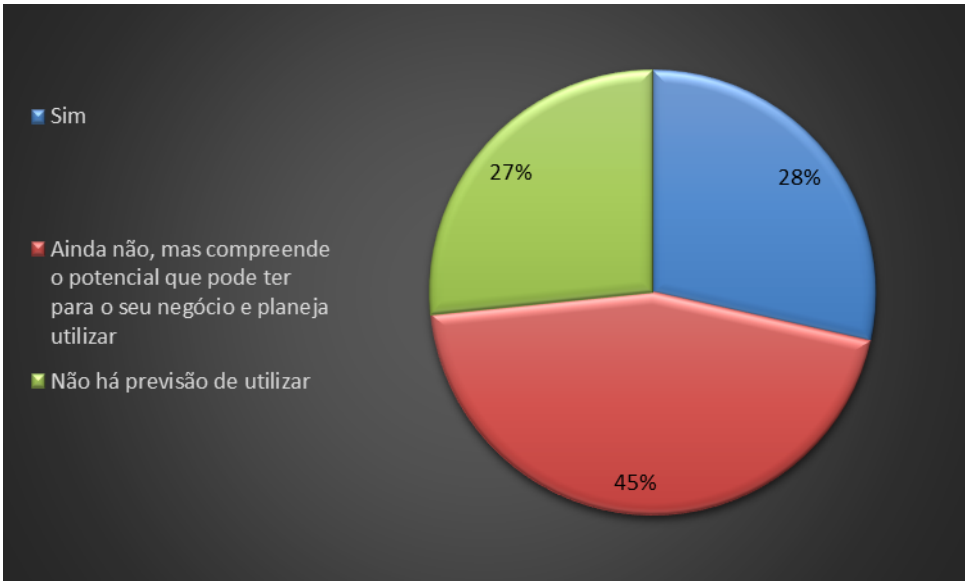


Figura 9 – Utilizam da IA pelas organizações pesquisadas

76. Partindo-se do conceito de que o presente levantamento objetiva definir uma linha de base inicial que venha a subsidiar futuros trabalhos do TCU que busquem avaliar a evolução da utilização de IA na APF, buscou-se no questionário eletrônico encaminhado às organizações públicas a inserção de questões que permitissem verificar o nível de maturidade em relação ao tema.

77. Nesse sentido, após pesquisa realizada pela equipe, optou-se por utilizar como referencial neste trabalho, com adaptações, o Modelo de Maturidade em Inteligência Artificial proposto pelo Gartner Group, publicado em 18/3/2020 (ID G00466009). O referido modelo propõe a estruturação da maturidade das organizações em relação ao uso de IA nos seguintes níveis:

Tabela 4 - Níveis de maturidade na utilização de IA - Gartner Group (ID G00466009)

Nível	Principais características
1 - Planejamento	As conversas internas sobre IA são ad hoc e especulativas. Apenas uma pequena minoria de colaboradores experimentou ou testou projetos. Nesta fase, os principais questionamentos são "O que é inteligência artificial?" e "Como a IA pode beneficiar minha linha de negócio?"

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Nível	Principais características
2 - Experimentação	Projetos iniciais de prova de conceito (POC) são elaborados e podem estar em fase piloto. Uma comunidade informal de prática pode ter se reunido várias vezes para compartilhamento de conhecimento e primeiros esforços de padronização.
3 - Estabilização	Os primeiros projetos de IA estão em produção. Existe um patrocinador da Alta Administração. Há orçamento exclusivo para projetos de IA. Um Centro de Especialização em IA foi criado, disponibilizando especialistas em IA, melhores práticas e tecnologia para projetos.
4 – Expansão	Todos os novos projetos digitais, incluindo revisões de processo para otimização, consideram a utilização de IA como forma de agregar valor. Existe *accountability* em torno de cada projeto. As unidades de negócio confiam nas técnicas de IA e estão prontas para usá-los. Os profissionais de DevOps / MLOps de TI são os principais responsáveis pela manutenção e atualização de soluções de IA.
5 – Transformação	A IA é rotineira e esperada como um elemento da execução de todos os processos de negócio. Todos os colaboradores no projeto e design de aplicativos são treinados e entendem os pontos fortes e riscos da IA. Os aplicativos que utilizam IA interagem de forma produtiva dentro da organização e em todo o ecossistema de negócios.

78. Além dos níveis de maturidade acima descritos, optou-se, a partir da análise das respostas ao questionário, pela criação do Nível 0, refletido pelas organizações que atualmente não utilizam e nem planejam utilizar tecnologias de IA nas soluções desenvolvidas. Dessa forma, verificou-se, no universo das 263 organizações respondentes ao questionário, a seguinte configuração em relação à maturidade para utilização de soluções de IA:

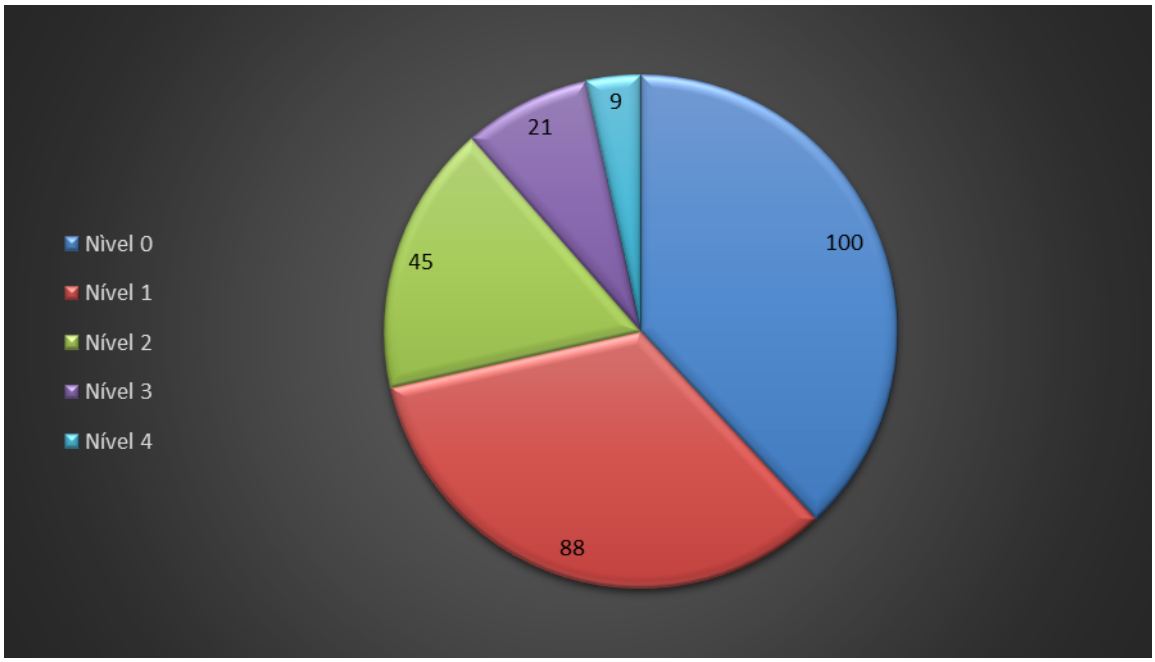


Figura 10 – Distribuição das organizações públicas federais por nível de maturidade quanto à utilização de IA

79. Conforme se observa no gráfico acima, 100 (38%) das 263 organizações públicas pesquisadas ainda não tem intenção de utilizar soluções de IA em seus processos. Dentre essas 100 organizações

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

(nível de maturidade 0), os principais dificultadores citados para a utilização da referida tecnologia são os seguintes:

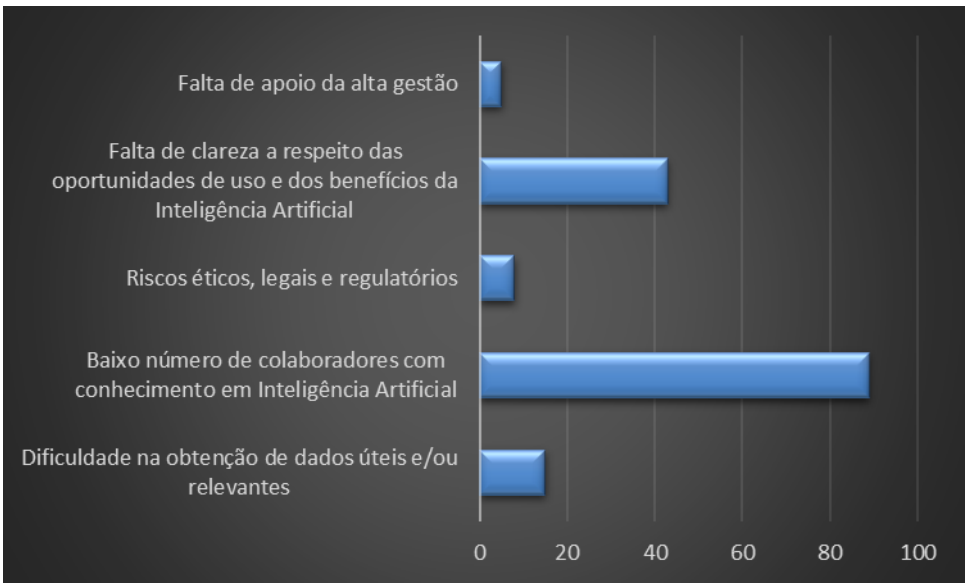


Figura 11 - Nível de Maturidade 0 - Principais Dificultadores apontados para implementação de Soluções de IA

80. Verifica-se que o ponto que mais impacta na intenção de utilização de IA nas organizações da APF é o baixo número de colaboradores com conhecimento na tecnologia. Com relação a esta questão, verificou-se no presente levantamento que, praticamente em todas as iniciativas identificadas de utilização de IA, o processo iniciou com um reduzido número de colaboradores que, em geral, trouxeram conhecimentos adquiridos externamente para dentro das organizações, antes mesmo de haver treinamentos institucionais relacionados à IA.

81. Corroborando esta afirmação, tem-se que dentre as 75 organizações que estão nos níveis de maturidade 2 a 4 e que desenvolvem internamente soluções de IA, 39 informaram dispor de no máximo cinco colaboradores que detém conhecimento em relação às tecnologias de IA.

82. Ainda com relação a este tema, tem-se que a escassez de colaboradores com conhecimento em IA pode ser suprida por meio da contratação de serviços terceirizados para o desenvolvimento de soluções ou a contratação de soluções de mercado. Neste contexto, tem-se, entre as 75 organizações com nível de maturidade 2 a 4, o seguinte cenário em relação a quem desenvolve as soluções de IA:

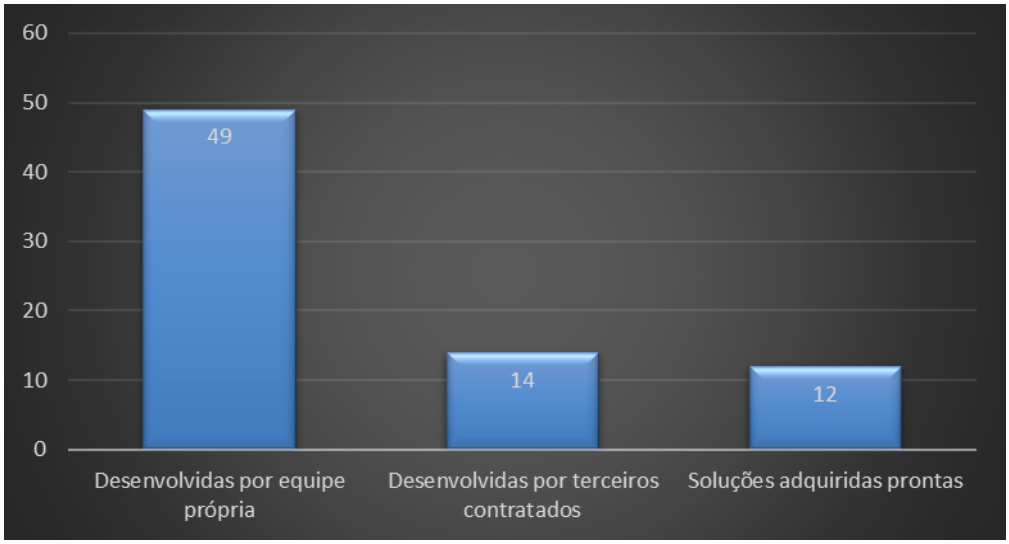


Figura 12 - Como são desenvolvidas as Soluções de IA - Níveis de Maturidade 2 a 4 (75 organizações)

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

83. Verificou-se também, em relação à capacitação de colaboradores da organização, que das 163 organizações que estão no nível de maturidade 1 a 4, 32 informaram que já dispõe de plano de capacitação institucional para desenvolvimento e sustentação de soluções de IA. Nessas mesmas organizações, a pesquisa também identificou, independentemente da existência de um plano institucional de capacitação de IA, as seguintes quantidades de treinamentos já realizados:

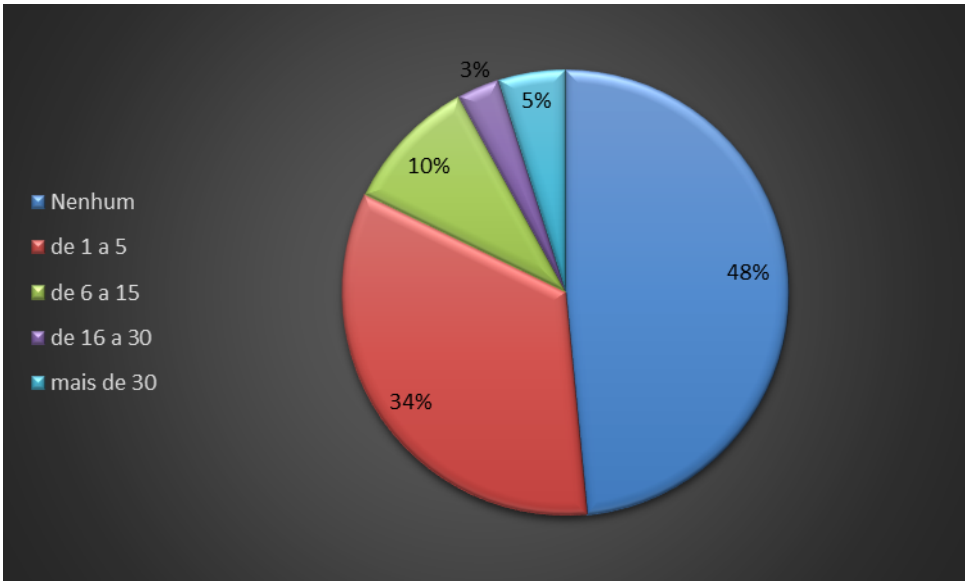


Figura 13 - Quantidade de colaboradores que receberam capacitação em desenvolvimento e sustentação de soluções de IA

84. Ainda em relação ao nível de maturidade 0, outro fator relevante que, segundo a pesquisa realizada, impede a adoção de tecnologias de IA nas organizações públicas é o desconhecimento em relação às oportunidades e benefícios que podem ser alcançados com o uso das tecnologias.

85. Se considerarmos as 163 organizações com níveis de maturidade de 1 a 4, tem-se os seguintes dificultadores apontados para implementação de soluções de IA:

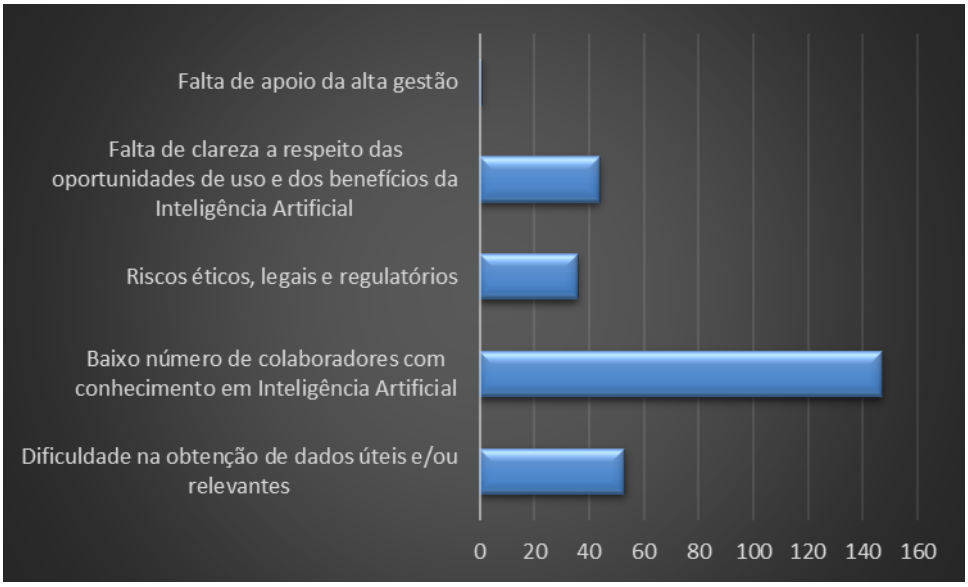


Figura 14 - Níveis de Maturidade 1 a 4 - Principais Dificultadores apontados para implementação de Soluções de IA

86. Verifica-se, assim, que mesmo entre as organizações que já planejam ou desenvolveram soluções de IA (níveis de maturidade 1 a 4), o baixo número de colaboradores também é o principal fator que impacta na implementação de tecnologias de IA nessas organizações. Apesar da relevância dessa

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

questão, tem-se que apenas 32 das 163 organizações informaram que desenvolveram plano de capacitação específico para os colaboradores envolvidos no desenvolvimento ou sustentação das soluções de Inteligência Artificial. Além disso, apenas 48 (29 %) das 163 organizações com nível de maturidade 1 a 4 consideram incluir nas próximas seleções públicas conhecimentos em tecnologias de inteligência artificial como pré-requisito para contratação.

87. A pesquisa também identificou as seguintes áreas finalísticas nas quais estão sendo aplicadas as soluções de IA identificadas no presente trabalho:

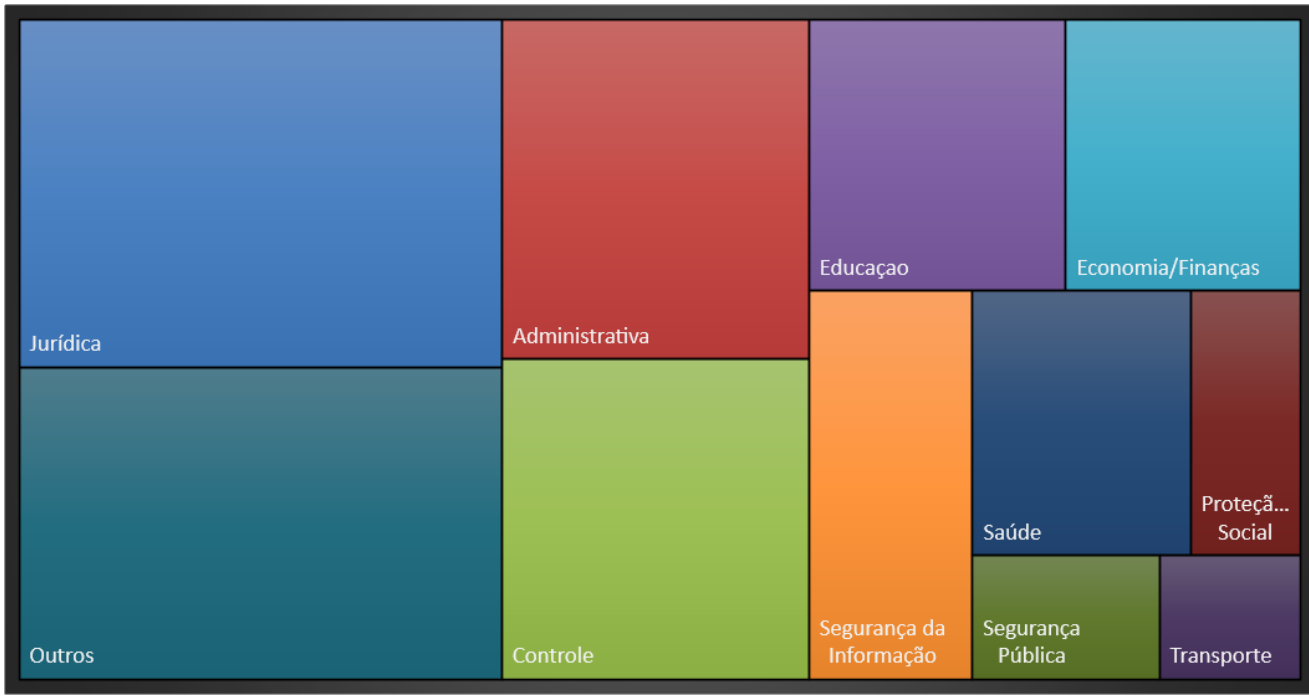


Figura 15 - Principais áreas de aplicação de soluções de IA na APF

88. Destaca-se na figura acima a predominância de utilização das soluções de IA pelos órgãos do Poder Judiciário, conforme já abordado pela pesquisa realizada pela Fundação Getúlio Vargas e detalhado nos parágrafos 71 a 73 supra. Além disso, verifica-se que há também maior utilização de tecnologias de IA nos processos administrativos internos das organizações públicas e em áreas de controle.

89. Ou seja, o questionário evidenciou que a maioria das soluções de IA implementadas até aqui não visaram o fornecimento de serviços diretos aos cidadãos, mas sim, no aprimoramento dos processos internos. Entende-se que tal quadro vai ao encontro do atual cenário pelo que passa a APF, de redução de pessoal e a consequente e necessária otimização de processos. Contudo, dado o aspecto disruptivo e inovador que IA traz, verifica-se que há espaço para crescimento do seu uso em soluções que atingem diretamente a sociedade.

90. A pesquisa também identificou que as soluções de IA em desenvolvimento, implementação ou em produção, estão distribuídas, conforme seus domínios de aplicação, conforme segue:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

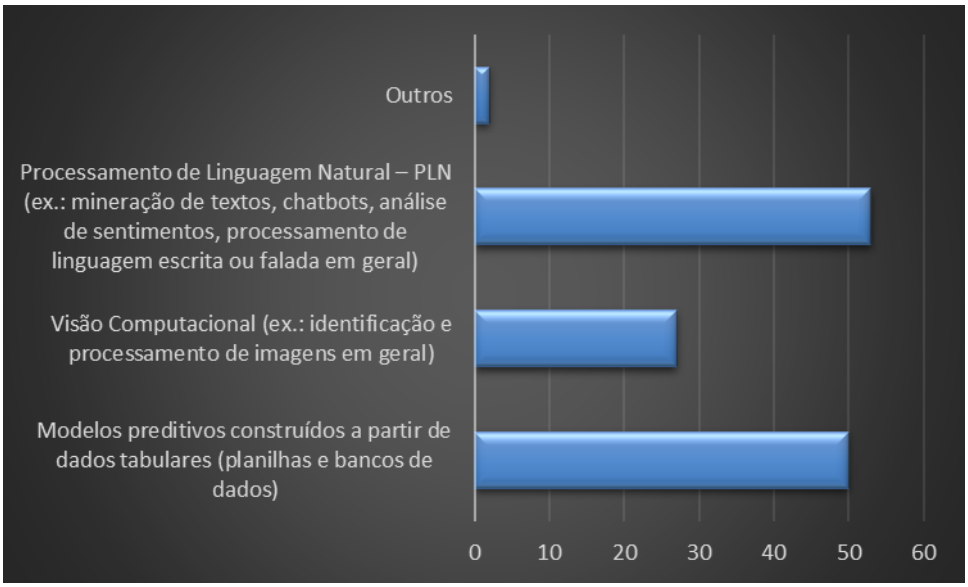


Figura 16 – Quantidade de soluções de IA na APF por domínio de aplicação

91. O domínio de processamento de linguagem natural (PLN) é amplamente utilizado nas soluções de IA que são desenvolvidas pelas organizações do Poder Judiciário, refletindo a necessidade de se tratar a grande quantidade de processos e documentos que compõem o dia a dia daquelas instituições. Além disso, destaca-se também a utilização de PLN no desenvolvimento de *chatbots* para interação com os usuários externos e o provimento de serviços. Verificou-se ainda a utilização de modelos preditivos, amplamente utilizados nas áreas de controle e fiscalização.

92. Já em relação às técnicas analíticas que são empregadas nas soluções de IA, evidenciou-se uma predominância da utilização de algoritmos tradicionais no desenvolvimento das soluções (considerando as soluções em desenvolvimento, em prova de conceito e em produção):

Técnica de IA	Qtd. Soluções implementadas
Algoritmos tradicionais de aprendizado supervisionado (Ex.: regressão linear, regressão logística, árvores de decisão, *support vector machines*, Naïve Bayes, GBM, etc.)	59
Redes neurais profundas (*Deep Learning*)	31
PLN avançado (Ex.: RNN, modelos de linguagem, *Transformers*, etc.)	32
Aprendizado não supervisionado (Ex.: agrupamentos, K-Means, DBSCAN, etc.)	32
Aprendizado por reforço (*Reinforcement Learning*)	15
Algoritmos evolucionários (Ex.: genético, colônia de formiga, enxame de abelha, etc.)	4
Algoritmos baseados em grafos	8
AutoML	4
Outros	2

93. Entre os benefícios apontados com a utilização de IA, as organizações pesquisadas informaram como principais ganhos qualitativos/quantitativos, em resumo:

a) automação de processos repetitivos com aumento de produtividade;

b) redução do tempo e custos na execução dos processos;

c) otimização na alocação de recursos humanos;

d) aumento da confiabilidade e transparência dos processos;

e) maior tempestividade na resposta aos clientes;

f) maior atuação preventiva das organizações em sua área de atuação;

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

g) uniformização de procedimentos.

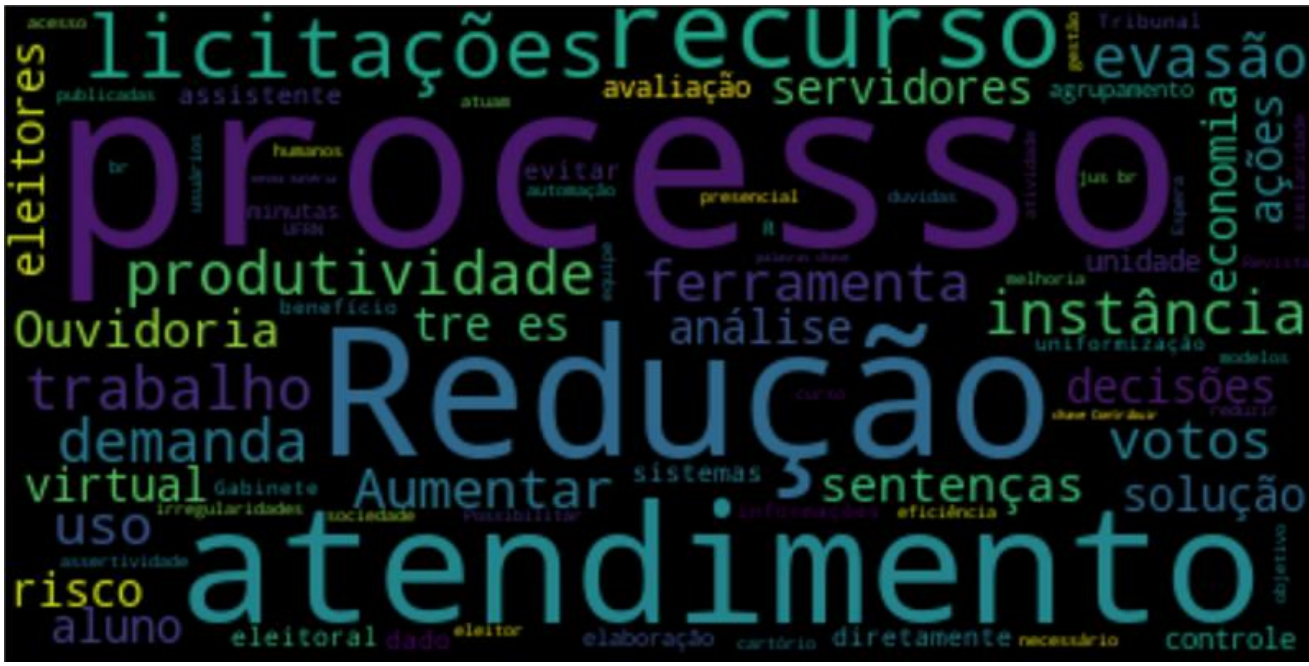


Figura 17 – *Wordcloud* - Principais benefícios obtidos com a utilização de soluções de IA identificados na pesquisa

94. Em relação às contratações realizadas pelas organizações para implementação de soluções de IA, dentre as 29 organizações que informaram a respeito, a pesquisa evidenciou que ainda são poucas as iniciativas que envolvem contratação específica para o desenvolvimento de soluções de TI. A grande maioria dos casos levantados utiliza colaboradores da própria organização no desenvolvimento das soluções e as contratações estão concentradas em insumos de infraestrutura. Importante observar que o Ministério da Saúde e o Incra informaram que utilizaram a modalidade Encomenda Tecnológica para desenvolvimento das soluções de IA:

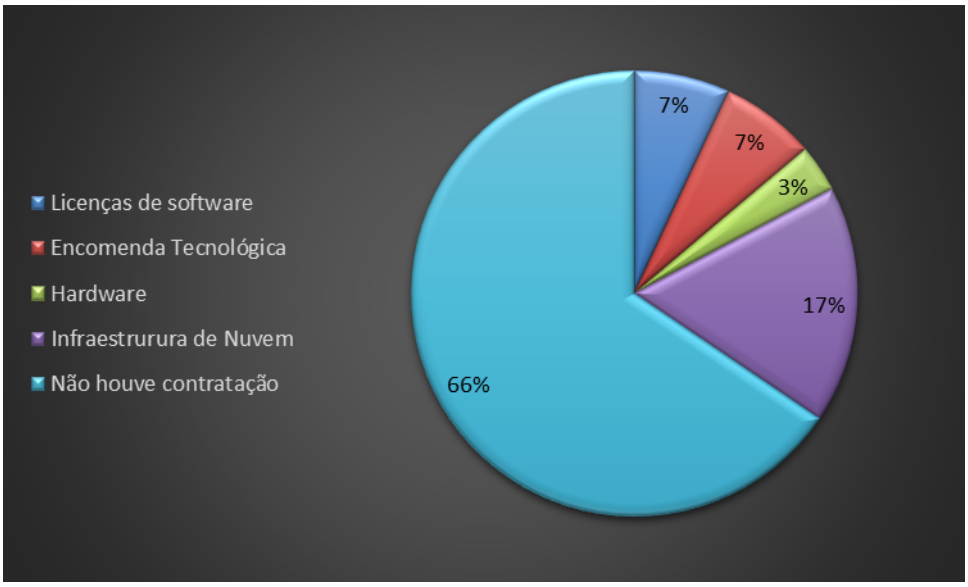


Figura 18 - Perfil de contratações para implementação de soluções de IA

95. Um ponto a ser a ser destacado é a possibilidade da realização de eventos denominados *Hackatons* para o desenvolvimento de soluções de IA. O *Hackathon* é um evento que reúne programadores, designers e outros profissionais externos ligados ao desenvolvimento de software para uma maratona de programação, cujo objetivo é desenvolver um software ou solução tecnológica que atenda a um fim específico. Ao final do evento, as melhores soluções desenvolvidas, geralmente em forma de protótipo,

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

são premiadas e podem ser posteriormente refinadas e internalizadas pelas instituições. Organizações como o Tribunal Regional do Trabalho – 5ª Região (TRT5) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) já utilizaram com sucesso esse recurso.

96. Nas respostas aos questionários, foram detalhadas 36 soluções de IA desenvolvidas pelas organizações públicas pesquisadas. O Apêndice X traz uma consolidação das respostas obtidas.

22 Estratégia Brasileira de Inteligência Artificial (EBIA)

23 Conjuntura internacional

97. Atualmente, verifica-se que há uma corrida global entre as nações para se tornarem líderes na utilização de inteligência artificial (IA). Nesse contexto, uma das principais ações desenvolvidas é a elaboração e implementação de uma estratégia nacional para a utilização de inteligência artificial. Tal processo passa, além da elaboração da estratégia, pela implementação de legislação que dê suporte às ações planejadas em nível governamental.

98. Nesse sentido, verifica-se que as estratégias nacionais desenvolvidas abordam temas como ganhos de produtividade, impactos no mercado de trabalho, na educação e demais áreas de atuação governamental, bem como questões referentes à pesquisa e à ética e ao incentivo da utilização da IA como instrumento para otimização de processos e diminuição de custos.

99. O Instituto de Tecnologia e Sociedade do Rio (ITS) produziu em 2020 um resumo detalhado dos planos estratégicos de desenvolvimento de IA de 15 países (Estados Unidos, Canadá, Japão, Singapura, China, Emirados Árabes Unidos, Finlândia, Dinamarca, Taiwan, França, Comissão Europeia, Reino Unido, Coreia do Sul, Austrália, Suécia, Índia, México, Alemanha, Portugal e Itália)[15]. Entre os países avaliados, descreve-se abaixo as considerações apresentadas sobre a estratégia de IA algumas nações:

Tabela 5 - Objetivos das estratégias nacionais de IA[16]

País	Objetivo estratégico principal
Alemanha	O governo pretende que a "IA fabricada na Alemanha" se torne um selo de qualidade reconhecido mundialmente. A estratégia é garantir uma posição de destaque para a Alemanha como centro de pesquisa em IA, aumentando sua competitividade e promovendo os vários usos da IA na sociedade, a fim de alcançar um progresso tangível no interesse de seus cidadãos. A estratégia foi resultado de uma parceria entre diversos ministérios e foi sujeita também a uma consulta pública (Disponível em: https://www.ki-strategie-deutschland.de/home.html?file=files/ download /Nationale _KI-Strategie_engl.pdf
Austrália	Em junho de 2021, O governo federal australiano divulgou seu primeiro plano de ação dedicado a estabelecer a Austrália como líder global no desenvolvimento e adoção de inteligência artificial responsável. O plano estabelece quatro áreas-chave para o governo se concentrar, incluindo: i) desenvolver e adotar IA para transformar as empresas australianas; ii) criar um ambiente para crescer e atrair os melhores talentos de IA do mundo; iii) utilizar tecnologias de IA de ponta para resolver os desafios nacionais da Austrália; e iv) tornar a Austrália uma líder global em IA responsável e inclusiva. Entre os benefícios esperados na implementação do plano estão proteger o meio ambiente, melhorar os resultados da saúde, promover cidades inteligentes e impulsionar a economia. Disponível em https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-action-plan

[15] Disponível em https://itsrio.org/wp-content/uploads/2020/03/RelatorioAI.pdf

[16] Fonte: resumo detalhado dos Planos estratégicos de desenvolvimento de Inteligência Artificial – 2020 (ITS-2020). Disponível em https://itsrio.org/wp-content/uploads/2020/03/RelatorioAI.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

País	Objetivo estratégico principal
China	O Plano estratégico chinês é bastante compreensivo e detalhado. Tem sua implementação prevista em três fases, com o objetivo de tornar a China líder mundial em IA até 2030. O primeiro passo é alinhar a indústria de IA da China com os concorrentes até 2020; o segundo, alcançar a "liderança mundial" em alguns campos de IA até 2025; e terceiro, tornar-se o centro "primário" da inovação em IA até 2030.
Estados Unidos da América	A estratégia americana não possui, efetivamente, um plano traçado para coordenar suas atividades. Os estudos acerca do tema vêm sendo feitos desde o governo Obama no ano de 2016, quando foram elaborados três relatórios. Com o governo Trump assumiu-se uma abordagem diferente, orientada para o livre mercado de Inteligência Artificial. Em março de 2019, o governo federal dos EUA lançou o AI.gov para facilitar o acesso a todas as iniciativas governamentais de IA atualmente em andamento. O site consolida todas as iniciativas na área. Disponível em: http://ai.gov
França	A estratégia francesa visa o fortalecimento do ecossistema de IA no país para garantir sua atuação no contexto global. O foco é permanecer independente e criar uma frente de desenvolvimento. O governo se comprometeu a: i) a fomentar e apoiar o ecossistema de IA francesa; ii) melhor explorar seus bancos de dados centralizados para o desenvolvimento da IA; e iii) estabelecer uma estrutura ética abordando as questões levantadas pela IA. Disponível em: https://www.aiforhumanity.fr/pdfs MissionVillani_Report_ ENG-VF.pdf.
Reino Unido	O governo britânico lançou o Acordo Setorial para IA em 2018. Esse documento faz parte da estratégia industrial mais ampla do governo e tem como objetivo posicionar o Reino Unido como líder global em IA. Os esforços apontam políticas para melhorar pesquisas nos campos público e privado, investimento em educação (Ciência Tecnologia, Engenharia e Matemática), melhorar a infraestrutura digital e, por fim liderar o diálogo global no campo ético. Disponível em https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sectordeal
México	O governo mexicano validou como estratégia o documento "*Towards an AI Strategy in Mexico: Harnessing the AI Revolution*" feito pelo Oxford Insights, C-Minds e encomendado pela Embaixada Britânica no México. De acordo com esse documento o México busca ser um líder global em IA e digitalização como forma de promover desenvolvimento, tanto econômico como social, para seus cidadãos. O documento aponta uma série de recomendações organizadas nas categorias: i) governo e serviços públicos; ii) dados e infraestrutura digital; iii) pesquisa e desenvolvimento; iv) capacitação, habilidade e educação; e v) ética. Disponível em https://www.gob.mx/mexicodigital/articulos/estrategiade-inteligencia-artificial-mx-2018

100. Em decorrência da relevância que tem tido a IA no contexto global, verifica-se que, além das estratégias individuais de cada nação, também há esforço para a construção de iniciativas regionais para o alinhamento em torno do tema. Nesse contexto, destacam-se as seguintes ações:

a) **União Europeia**: há uma Aliança Europeia em IA que estabeleceu uma abordagem abrangente e um acordo de cooperação entre os países europeus. Vinte e cinco países europeus assinaram uma "Declaração de Cooperação" em Inteligência Artificial em abril de 2018. Embora alguns desses países também tenham iniciativas nacionais de IA, eles enfatizaram a importância de atuarem juntos para aprimorar a pesquisa e a implantação, enquanto lidam coletivamente com questões sociais, econômicas, éticas e legais;

b) **Região Nórdico-báltica**: os Ministros da Dinamarca, Estônia, Finlândia, Ilhas Faroe, Islândia, Letônia, Lituânia, Noruega, Suécia e Ilhas Åland emitiram uma declaração de colaboração sobre IA em maio de 2018, "IA na região nórdico-báltica". A iniciativa tem o objetivo de desenvolver e

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

promover o uso da inteligência artificial para servir melhor os seres humanos. Os países pretendem colaborar no desenvolvimento de competências, acesso a dados, padrões e princípios, garantindo o papel da IA no Mercado Único Digital Europeu, evitando regulamentação desnecessária e utilizando o Conselho Nórdico de Ministros para facilitar a colaboração;

c) **Membros do G7:** os países líderes do G7 - Canadá, França, Alemanha, Itália, Japão, Reino Unido e Estados Unidos - se comprometeram com a "Visão Comum de Charlevoix para o Futuro da Inteligência Artificial". Essa visão compartilhada destaca o potencial de crescimento econômico ao lado de desafios sociais urgentes; afirma que "a IA que promove o crescimento econômico, a confiança da sociedade, a igualdade de gênero e a inclusão depende de um ambiente político previsível e estável que promova a inovação".

101. É nesse contexto mundial que se avalia na próxima seção a Estratégia Brasileira de Inteligência Artificial.

24 Estratégia Brasileira de Inteligência Artificial

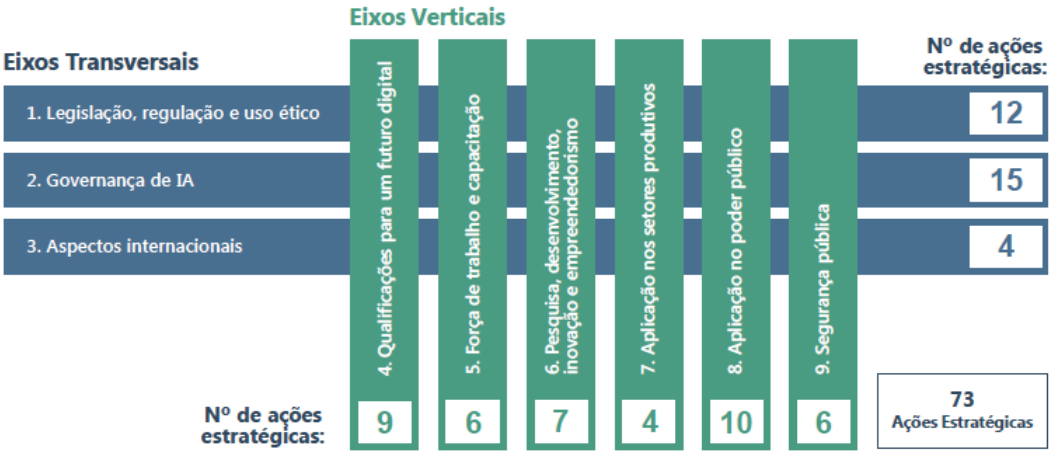
25 Descrição geral

102. Instituída pela Portaria MCTI 4.617/2021, alterada pela Portaria MCTI 4.979/2021, a Estratégia Brasileira de Inteligência Artificial (EBIA) assume o papel de nortear as ações do Estado brasileiro em prol do desenvolvimento das ações, em suas várias vertentes, que estimulem a pesquisa, inovação e desenvolvimento de soluções em Inteligência Artificial, bem como, seu uso consciente, ético e em prol de um futuro melhor.

103. Alinhada às diretrizes da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) endossadas pelo Brasil, a EBIA fundamenta-se nos cinco princípios definidos pela Organização para uma gestão responsável dos sistemas de IA, quais sejam: (i) crescimento inclusivo, o desenvolvimento sustentável e o bem-estar; (ii) valores centrados no ser humano e na equidade; (iii) transparência e explicabilidade; (iv) robustez, segurança e proteção e; (v) a responsabilização ou a prestação de contas (*accountability*).

104. A EBIA tem seis principais objetivos: (i) contribuir para a elaboração de princípios éticos para o desenvolvimento e uso de IA responsáveis; (ii) promover investimentos sustentados em pesquisa e desenvolvimento em IA; (iii) remover barreiras à inovação em IA; (iv) capacitar e formar profissionais para o ecossistema da IA; (v) estimular a inovação e o desenvolvimento da IA brasileira em ambiente internacional; e (vi) promover ambiente de cooperação entre os entes públicos e privados, a indústria e os centros de pesquisas para o desenvolvimento da Inteligência Artificial.

105. Para tanto, a Estratégia estabelece nove eixos temáticos, caracterizados como os pilares do documento, apresenta um diagnóstico da situação atual da IA no mundo e no Brasil, destaca os desafios a serem enfrentados, oferece uma visão de futuro, e apresenta um conjunto de ações estratégicas que a aproxima dessa visão:



Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Figura 19 – Eixos temáticos da EBIA

106. O processo de elaboração da EBIA foi construído por meio do desenvolvimento das seguintes etapas:

106.1. contratação de consultoria especializada em IA para realizar estudo a respeito dos potenciais impactos sociais e econômicos das tecnologias que embasam a IA, com vistas a analisar propostas de políticas que mitigassem seus efeitos negativos ao mesmo tempo em que maximizassem seus efeitos positivos;

106.2. *benchmarking* nacional e internacional;

106.3. processo de consulta pública à sociedade, realizada por meio de plataforma eletrônica do Governo Federal, entre 12 de dezembro de 2019 e 3 de março de 2020, a qual recebeu mais de mil contribuições.

107. Ressalta-se que o texto da Portaria MCTI 4.979/2021 reconhece que a EBIA é considerada uma política pública[17] e, nesse contexto, deve ser entendida como um conjunto de intervenções e diretrizes emanadas de atores governamentais, que visam tratar, ou não, problemas públicos e que requerem, utilizam ou afetam recursos públicos.

26 Metodologia utilizada para avaliação da EBIA

108. Uma vez que o Relator determinou que fosse incluído no escopo do presente levantamento a identificação dos principais riscos e oportunidades associados à implementação da EBIA, foi necessário estabelecer primeiramente quais seriam os critérios mais adequados para atender a essa demanda.

109. Ressalta-se que o "Referencial para Avaliação de Governança em Políticas Públicas", publicado pelo TCU em 2014, assim define o termo "políticas públicas" (BRASIL, TCU, 2014b, p. 21):

> Políticas públicas são um conjunto articulado e estruturado de ações e incentivos que buscam alterar uma realidade em resposta a demandas e interesses dos atores envolvidos. Dispõem usualmente sobre "o que fazer" (ações), "aonde chegar" (objetivos relacionados ao estado de coisas que se pretende alterar) e "como fazer" (princípios e macroestratégias de ação). Política Pública diz respeito, portanto, "à mobilização político-administrativa para articular e alocar recursos e esforços para tentar solucionar dado problema coletivo".

110. O objetivo do Referencial de Controle de Políticas Públicas é estabelecer uma estrutura de critérios comuns que possam compor um modelo para avaliar o nível de maturidade de políticas públicas, orientando e sistematizando ações de controle, de modo a contribuir para a melhoria do desempenho e dos resultados das políticas públicas, por meio do aprimoramento de seu processo de formulação, implementação e avaliação.

111. De acordo com esse referencial, para que os estágios de políticas sejam realizados de maneira adequada, algumas fases devem ser cumpridas com a adoção de boas práticas nacionais e internacionais para políticas públicas de qualidade. Essas fases estão organizadas com base em sete blocos de controle: (i) diagnóstico do problema e formação da agenda; (ii) análise de alternativas e tomada de decisão; (iii) desenho e institucionalização da política; (iv) estruturação da governança e gestão; (v) alocação e gestão de recursos orçamentários e financeiros; (vi) operação e monitoramento; e (vii) avaliação. A figura a seguir ilustra a interpelação entre estes componentes:

[17] De acordo com a própria Portaria MCTI 4.617/2021, a EBIA a deve ser uma política pública constantemente acompanhada, avaliada e ajustada, pois o ritmo da evolução tecnológica da Inteligência Artificial tende a se acelerar.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

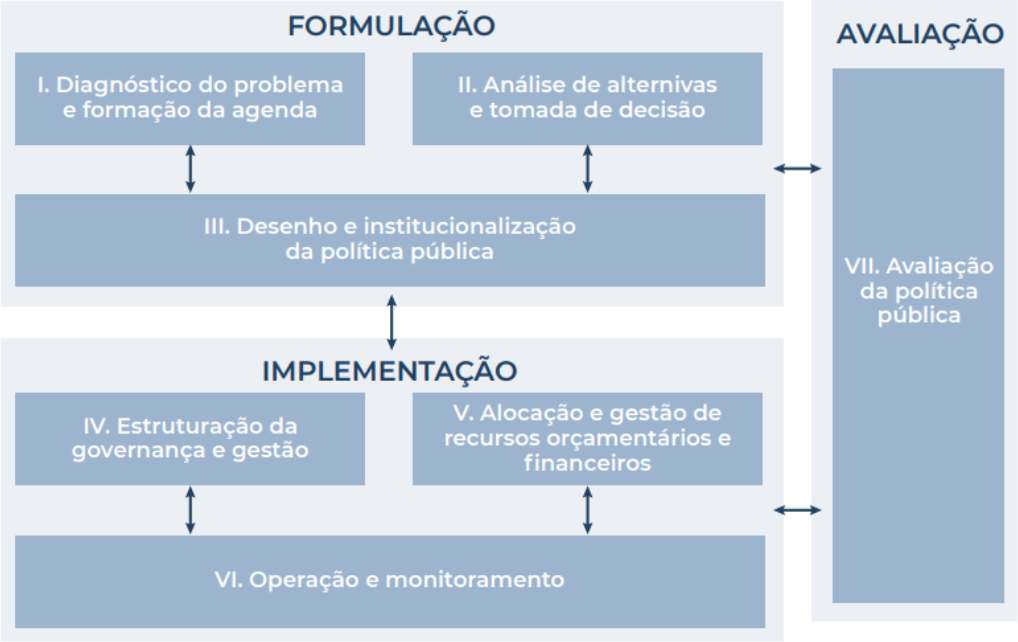


Figura 20 – Blocos de controle do Referencial de Controle de Políticas Públicas.

112. Como a EBIA foi publicada durante a execução do presente levantamento e, portanto, ainda não havia produzido completamente seus efeitos, optou-se por selecionar apenas os quatro primeiros blocos de controle para a sua avaliação, os quais são conceituados a seguir:

112.1. **Diagnóstico do problema** refere-se ao conjunto de práticas voltadas para a identificação do problema, sua delimitação e caracterização, identificação de público-alvo e de possíveis causas, efeitos e formas de tratamento. **Formação da agenda pública** envolve a escolha do conjunto de problemas ou temas considerados merecedores de intervenção pública.

112.2. **Análise de alternativas** se desenvolve por meio de escrutínios formais ou informais das consequências do problema e confronta aspectos como custo-benefício e custo-efetividade de cada alternativa disponível para apoiar o processo de tomada de decisão. **Tomada de decisão** representa o momento em que os interesses dos atores são equacionados e as intenções (objetivos e métodos) de enfrentamento de um problema são explicitadas.

112.3. **Desenho da política pública** contempla a caracterização da política com base em modelo lógico que explicita, entre outras coisas, seus objetivos, produtos, atividades, resultados e impactos. A **Institucionalização** envolve a publicação da política pública por meio de norma legal (lei, decreto, resolução etc.) apropriada, emitida por órgão dotado de legitimidade e competência para fazê-lo.

112.4. **Estruturação da governança e gestão** envolve a definição das estruturas de governança, de gestão de riscos e controles internos, de monitoramento e de avaliação da política pública, bem como do plano de implementação da política, dos processos e operações necessárias ao seu funcionamento.

113. O objetivo da aplicação de um questionário de autoavaliação de controles (CSA) é a avaliação do nível de maturidade, o que significa verificar a qualidade do processo de formulação, implementação e avaliação da EBIA frente a um conjunto de boas práticas que, reconhecidamente, contribuem para a melhoria de seu desempenho e resultados.

114. Especificamente, buscou-se avaliar os seguintes aspectos baseados no referencial citado:

a) se o planejamento identificou os resultados esperados na transformação digital;

b) se a elaboração das estratégias contou com a participação das partes interessadas;

c) se a EBIA possui uma visão estratégica de longo prazo;

d) se foi explicitado o estágio de referência inicial da política (linha de base) que servirá de subsídio para avaliação do seu resultado;

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

e) se a definição dos objetivos é precisa para orientar a execução das ações;

f) se os esforços e os resultados são passíveis de mensuração, considerando aspectos como viabilidade física, financeira e temporal;

g) se os objetivos e as metas são coerentes com a visão estratégica da EBIA;

h) se a ordem de precedência para a realização de objetivos e metas dependentes entre si está identificada;

i) se foram definidos prazos para a realização dos objetivos intermediários e finais;

j) se o sistema de monitoramento e avaliação da transformação digital dispõe de processos, procedimentos e meios suficientes (recursos financeiros, pessoas, estrutura etc.) para assegurar o monitoramento e a avaliação que forneça informações confiáveis, tempestivas e necessárias para a tomada de decisão.

115. Ressalta-se que a metodologia utilizada está alinhada com a visão da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) sobre o ciclo de uma política de Inteligência Artificial. Entende-se que os critérios selecionados para avaliação da EBIA cobrem as três primeiras fases do ciclo proposto pela OCDE, o qual é constituído pelas seguintes etapas:

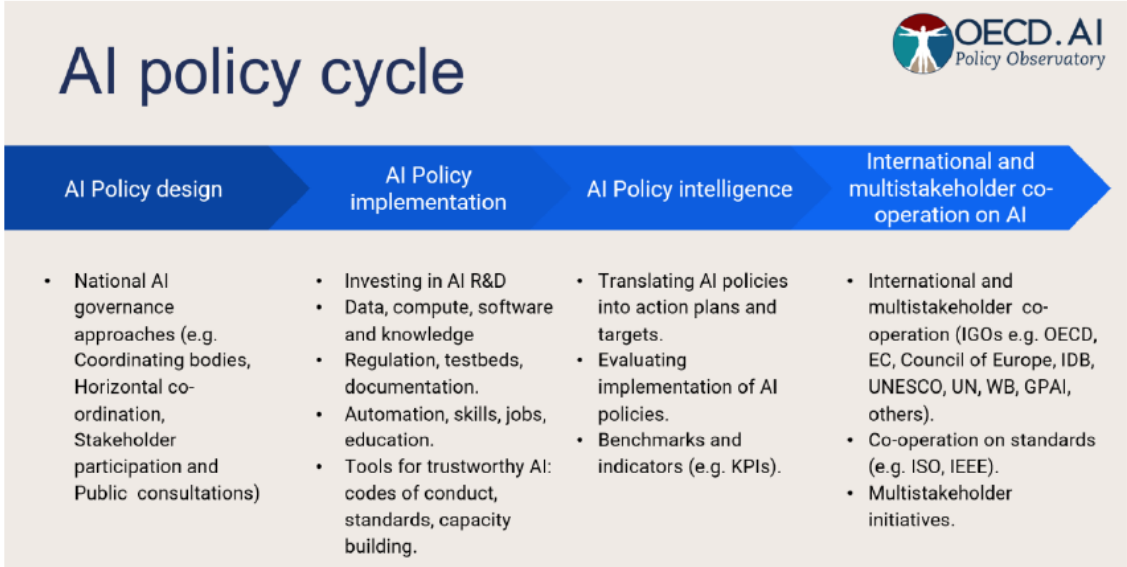


Figura 21 – Ciclo de uma Política de Inteligência Artificial. Fonte: OCDE (adaptado)

27 Achados de auditoria

Achado I – Os objetivos da EBIA não são específicos, mensuráveis, realistas (alcançáveis) e delimitados em um recorte temporal

Situação encontrada

116. A operacionalização de uma política pública é delineada em função dos objetivos propostos. De acordo com a Lei 13.971/2019, que instituiu o Plano Plurianual da União para o período 2020 a 2023, objetivo é uma declaração de resultado a ser alcançado que expressa, em seu conteúdo, o que deve ser feito para a transformação de determinada realidade.

117. Feitas essas considerações, importa dizer que a EBIA não define precisamente os objetivos para enfrentamento dos problemas públicos ou aproveitamento das oportunidades no que diz respeito ao uso da Inteligência Artificial.

118. Na introdução da EBIA são apresentados seis objetivos estratégicos, conforme ilustrado na figura abaixo[18]:

[18] Fonte: Estratégia Brasileira de Inteligência Artificial

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

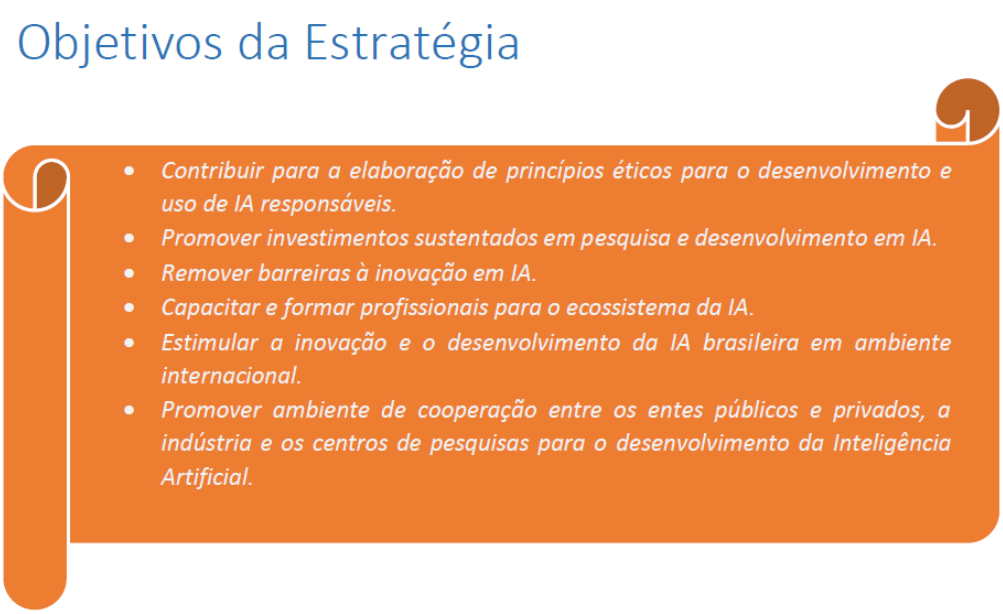


Figura 22 – Objetivos Estratégicos da EBIA

119. O TCU ("Referencial para Avaliação de Governança em Políticas Públicas do TCU"[19] e Acórdão 2.359/2018-TCU-Plenário) e o Instituto de Pesquisa Econômica Aplicada - IPEA (Guia Prático de Análise *Ex Ante*), alinhados às melhores práticas internacionais, enfatizam a necessidade de definir os objetivos e os respectivos indicadores de acordo com o modelo SMART, cujas propriedades desejáveis são as de serem:

a) específicos (*specific*): indicadores devem refletir informações simples e facilmente comunicáveis, objetivando um aspecto específico a ser submetido a mudanças pela intervenção;

b) mensuráveis (*measurable*): indicadores qualitativos e quantitativos devem ser mensuráveis, possibilitando aferir se os resultados propostos foram alcançados ou não;

c) alcançáveis (*achievable*): a meta estabelecida por um indicador deve ser alcançável e realista em relação ao contexto em que se insere a intervenção;

d) relevantes *(relevant*): indicadores devem refletir informações relacionadas aos componentes que medem;

e) temporalmente regulares (*time*): indicadores devem explicitar a periodicidade com que precisarão ser medidos para que possam ser úteis à medição de resultados.

120. Porém, os objetivos da EBIA não satisfazem totalmente a propriedade SMART (sigla em inglês para *Specific, Measurable, Attainable, Realistic e Time-bound*), uma vez que não são específicos, mensuráveis, realistas (alcançáveis) e delimitados em um recorte temporal. É percebido apenas a propriedade de serem apropriados (relevantes).

121. Em razão dessas falhas, na prática, os objetivos apresentados na EBIA correspondem mais a diretrizes estratégicas que orientam as ações derivadas dela. Não se sabe em que de concreto se traduzem os objetivos definidos.

122. Outro problema encontrado é que as ações estratégicas da EBIA não definem indicadores claros para a ação governamental, conforme requer a propriedade SMART. Além disso, embora seja perceptível que as ações estratégicas estão relacionadas com seus eixos temáticos, a EBIA não diz explicitamente a quais objetivos as ações estratégicas estão associadas, dificultando a avaliação do nível de coerência entre eles.

[19] Disponível em: https://portal.tcu.gov.br/governanca/governancapublica/politicas-publicas/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

123. Diante dessas inconsistências, cumpre registrar que, para evitar o comprometimento de suas ações, o TCU tem o entendimento de que não é viável ou aceitável que os objetivos definidos em uma política pública não sejam expressos em termos de resultados.

124. Assim, o MCTI deveria estabelecer objetivos (ou indicadores) claros para balizar a ação governamental, à exemplo do que fez o Ministério da Educação por meio da Lei 13.005/2014, que instituiu o Plano Nacional de Educação (PNE). A figura abaixo exemplifica um dos vinte objetivos definidos pelo PNE, os quais possuem as características de específico, mensurável, apropriados, realista e com prazo delimitado:

META 9 Elevar a taxa de alfabetização da população com 15 (quinze) anos ou mais para 93,5% (noventa e três inteiros e cinco décimos por cento) até 2015 e, até o final da vigência deste PNE, erradicar o analfabetismo absoluto e reduzir em 50% (cinquenta por cento) a taxa de analfabetismo funcional.

Figura 23 – Objetivo 9 do Plano Nacional de Educação

125. Outro exemplo do Governo Federal que pode ser citado é o Decreto 10.332/2020, que institui a Estratégia de Governo Digital para o período de 2020 a 2022, o qual na Iniciativa 8.2 estabelece que devem ser implementados recursos de inteligência artificial em, no mínimo, 12 serviços públicos federais, até 2022.

126. Portanto, uma vez que a definição genérica de objetivos não é suficiente para orientar a execução das ações, entende-se que o MCTI deveria ter elaborado os objetivos ou as ações estratégicas conforme o modelo SMART.

Critérios

a) Acórdão 2.359/2018-TCU-Plenário, relatoria do Min. Vital do Rêgo;

b) Avaliação de Políticas Públicas - Guia Prático de Análise *Ex Ante*, IPEA;

c) Lei 13.971/2019.

Evidências

a) Seção "Objetivos da Estratégia" da EBIA (Portaria MCTI 4.979/2021).

Causas

a) Não identificadas.

Efeitos

a) Dificuldade de avaliar o nível de coerência entre as ações estratégicas e os objetivos definidos pela EBIA;

b) Dificuldade em estabelecer riscos específicos ao alcance de resultados da estratégia;

c) Impossibilidade de avaliar se os objetivos fixados guardam nítida relação com a finalidade última da EBIA ou com os problemas/oportunidades previamente identificados;

d) Insuficiência dos objetivos para orientarem a execução de ações.

Proposta de Encaminhamento

a) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

Achado II – **Inexistência da explicitação do estágio de referência inicial ("marco zero", "situação inicial") e da definição de indicadores e metas para subsidiar a avaliação do desempenho e resultados da EBIA**

Situação encontrada

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

127. A elaboração de indicadores quantitativos sobre o problema permite dimensioná-lo no tempo anterior à implementação da política, traçando o seu panorama. Nota-se que, sem valores da linha de base acerca da realidade à qual a política visa promover mudanças, não se poderá desenvolver futuramente nenhuma avaliação relevante do atingimento de seus resultados.

128. Nesse sentido, cumpre dizer que a EBIA falhou em responder questões triviais como: Quais são os indicadores-chave e metas para os principais objetivos? Qual é a situação inicial observada desses indicadores? Em qual magnitude vão ser tratados, ou seja, qual é o grau das mudanças geradas pela política que serão consideradas satisfatórias?

129. Em que pese a EBIA ter uma seção específica destinada ao diagnóstico da situação atual, apresentando dados qualitativos e quantitativos sobre do uso de IA no país, convém relatar que não foram explicitamente definidos e priorizados quais indicadores a EBIA pretende medir, com o intuito de subsidiar posterior avaliação do desempenho da intervenção pública.

130. Importante salientar que é recomendada a definição de indicadores que acompanhem o fluxo de operacionalização da política, como: insumos/recursos, processos, produtos, resultados e impactos, conforme ilustrado na figura abaixo:

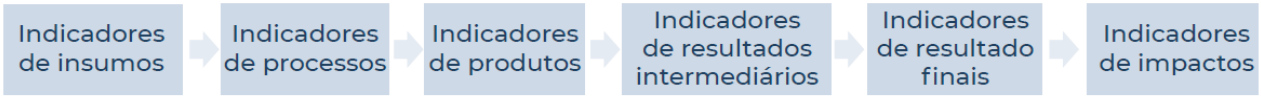


Figura 24 – Cadeia de indicadores desejável

131. Nada obstante, entende-se que, no mínimo, a EBIA deveria ter os indicadores de resultado associados aos seis objetivos estratégicos, ou às 73 ações estratégicas, conforme previsto no modelo SMART (Acórdão 2.359/2018-TCU-Plenário e Guia Prático de Análise *Ex Ante)*.

132. Com efeito, verifica-se que os indicadores poderiam ser derivados de informações apresentadas na própria EBIA, tais como "O Brasil está situado na 66ª posição do Índice de Inovação Global de 2019" e "O número de pós-graduações em programas de tecnologias digitais no Brasil (i.e. que envolvam robótica, Inteligência Artificial / aprendizado de máquina ou big data) em 2016 era de 224, sendo 72 mestrados e 152 doutorados". Por exemplo, um indicador básico que a EBIA poderia medir seria a posição no ranking de inovação global e a quantidade de cursos de pós-graduação ofertados. Além disso, os indicadores deveriam ser associados às ações estratégicas e, consecutivamente, aos objetivos estratégicos da EBIA para acompanhar o andamento e medir o cumprimento dos objetivos de uma política.

133. Dessa forma, seria possível definir o conjunto de indicadores (associados aos eixos temáticos) a serem medidos pela EBIA com os respectivos dados sobre a situação atual, o que representaria o "estágio de referência inicial" ou "marco zero" da Estratégia, à exemplo do que foi feito pelo PNE no documento intitulado "Linha de Base" (http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkW1/document/id/493812).

134. Entretanto, isso não foi feito, o que quer dizer também que não foram definidas as metas para cada um dos indicadores. Nesse contexto, faz-se necessário esclarecer a diferença entre metas e indicadores. De acordo com a Lei 13.971/2019, indicador é instrumento gerencial que permite a mensuração de desempenho de programa em relação à meta declarada, enquanto meta é a declaração de resultado a ser alcançado, de natureza quantitativa ou qualitativa, que contribui para o alcance do objetivo.

135. Na elaboração de uma política pública, é esperado que sejam estabelecidas metas de entregas de produtos ou benefícios para um horizonte temporal, com vistas ao alcance dos objetivos pretendidos. Isso quer dizer que no caso da EBIA, deveriam ser estabelecidas metas precisas e objetivamente caracterizadas, que concorressem para a consecução dos propósitos mais gerais da intervenção pública, de modo a orientar as ações governamentais e assegurar a transparência sobre metas e resultados.

136. Uma meta necessita, pelo menos, de um quantitativo explícito e de um prazo máximo de execução para ser clara e mensurável. Dessa forma, quanto à possibilidade de mensuração dos resultados da

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

EBIA, considerando aspectos como viabilidade física e temporal, a equipe de fiscalização verificou que, das 73 ações estratégicas, poucas apresentaram um quantitativo claro ou o ano máximo de execução (sendo raras exceções).

137. Diante dessas falhas, um exemplo do que poderia ser feito para atender ao Modelo SMART é, por exemplo, reescrever a ação estratégica "instituir programas de formação tecnológica para professores e educadores", do eixo vertical "Qualificações para um futuro digital" adicionando informações, tais como: "instituir, **na modalidade EAD**, **três** programas de formação tecnológica, **nas áreas de estatística, ciência de dados e inteligência artificial** para atender a no mínimo **2.000** professores e educadores, **até o ano de 2023**".

138. Entende-se que a definição de metas quantitativas é essencial para a avaliação da eficácia, eficiência e efetividade da política. O uso de indicadores quantitativos e/ou qualitativos apropriados deveria fazer parte do desenho da política, de modo que esses parâmetros iniciais possam ser comparados com os resultados efetivos e as metas alcançadas. Contudo, faltou clareza da EBIA em relação aos resultados finais que pretende atingir com a sua implementação.

Critérios

a) Acórdão 2.359/2018-TCU-Plenário, relatoria do Min. Vital do Rêgo;

b) Avaliação de Políticas Públicas - Guia Prático de Análise *Ex Ante*, IPEA;

c) Lei 13.971/2019, art. 3º, inciso II.

Evidências

a) Estratégia Brasileira de Inteligência Artificial (Portaria MCTI 4.979/2021).

Causas

a) Não identificadas.

Efeitos

a) Falta de clareza acerca dos produtos e efeitos esperados da política;

b) Impossibilidade de monitorar o desempenho da política em termos de alcance de objetivos e produção de resultados de curto e médio prazo;

c) Ausência de entendimento dos resultados pretendidos com a EBIA pela sociedade.

Proposta de Encaminhamento

d) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

Achado III – **Falhas formais na apresentação do modelo lógico e sobre como a EBIA incide sobre os problemas/oportunidades, projetando seus resultados e impactos de longo prazo**

Situação encontrada

139. Uma das formas de se entender a EBIA é analisá-la como uma solução estruturada para capturar benefícios, solucionar problemas e remover barreiras do uso da IA no Brasil.

140. Nesse sentido, a literatura sobre planejamento no setor público considera como boa prática o uso de técnicas que auxiliam no desenho da política pública. Nada obstante, importa dizer que foram observadas falhas formais na elaboração da EBIA, devido à ausência de artefatos requeridos por normativos legais, relacionados à qualidade do processo de formulação de uma política pública.

141. Verificou-se que a EBIA não atendeu a dois requisitos importantes na fase de diagnóstico do problema e desenho, quais sejam: (i) apresentar de forma clara e objetiva a teoria que sustenta a política, explicitando a forma como a política incide sobre as causas e os efeitos do problema; e (ii) elaborar modelo lógico que detalha entradas, atividades, produtos, impactos e resultados.

142. Em relação ao primeiro item, vale salientar que, para uma melhor compreensão da política, o primeiro passo é delimitar qual oportunidade se quer aproveitar ou qual o problema se quer mitigar ou

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

resolver com a intervenção da política proposta. Problema público pode ser definido como a diferença entre a situação existente (realidade) e a situação desejada.

143. O Guia Prático de Análise Ex Ante[20] do IPEA, em consonância ao Decreto 9.191/2017, art. 27, inciso I, sugere que a caracterização do problema/oportunidade pode ser decomposta da seguinte maneira:

1) Identificação do problema/oportunidade.
2) Causas potenciais dos problemas/benefícios das oportunidades.
3) Dados quantitativos acerca do problema/oportunidade:
   a) evidências na realidade brasileira; e
   b) comparação internacional.

Figura 25 – Decomposição dos problemas/oportunidades. Fonte: Guia Prático de Análise *Ex Ante* (adaptado)

144. Após a criação de uma lista de problemas, pode-se utilizar um método específico, como Árvore de Problema, para verificar se as possíveis causas e consequências para a ocorrência do problema foram levantadas, no momento de sua identificação e se existe consistência lógica e coerência entre o problema, as causas e os efeitos apresentados. Essa árvore contém o problema central a ser atacado; as causas potenciais e o relacionamento entre elas e as consequências. A sua elaboração deixa a análise do problema mais didática e transparente, permitindo posteriormente à sociedade verificar se a política realmente ataca as causas centrais dos problemas identificados. Entende-se também que a construção de árvore similar pode ser utilizada para identificar os efeitos positivos do aproveitamento das oportunidades de inovação com o uso de inteligência artificial.

145. Ressalta-se que a seção de diagnóstico e dos eixos temáticos da EBIA e o documento produzido pela consultoria contratada pelo MCTI para elaboração da EBIA (peça 626) abordam sobre dados quantitativos dos problemas/oportunidades discutidos. Porém, devido à ausência do uso de metodologia que identificasse explicitamente os problemas e as oportunidades, não é possível dizer se os dados quantitativos coletados são adequados e suficientes.

146. Ainda sobre o primeiro item, as melhores práticas orientam que "a teoria de como a política incide sobre as causas do problema, projetando seus resultados e impactos de longo prazo" seja descrita da seguinte forma: se (descreva o desenho do programa), então (descreva os resultados), o que então levará a (descreva os impactos).

147. A segunda falha formal identificada na elaboração da EBIA é que não foi apresentado explicitamente um modelo lógico que demonstre como os recursos e atividades geram produtos, resultados e seus respectivos impactos.

148. O modelo lógico geralmente é composto por um fluxo, exibindo a racionalidade de conexão entre as atividades propostas através da política e os objetivos que se pretende atingir, conforme demonstrado na figura abaixo:

Figura 26 – Etapas do modelo lógico de uma política pública. Fonte: Guia Prático de Análise *Ex Ante*

149. Como se pode observar, essa é uma estrutura racional (passo a passo estruturado) que demonstra como recursos e ações mudam comportamentos, geram produtos e produzem resultados e impactos.

[20] O Guia Prático de Análise *Ex Ante* aborda apenas os problemas. Como uma tecnologia emergente abre a possibilidade para um país ganhar vantagem competitiva e capturar benefícios, o modelo proposto pelo referido guia foi adaptado para também incluir as oportunidades de inovação.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

Trata-se, portanto, de um importante fluxo que estabelece conexão entre os resultados pretendidos e os recursos e ações necessárias ao alcance daqueles.

150. Ao não apresentar o referido modelo, entende-se que a EBIA não apresentou transparência adequada quanto à identificação e definição de resultados esperados da política.

Critérios

a) Decreto 9.191/2017, art. 27, inciso I;

b) Avaliação de Políticas Públicas - Guia Prático de Análise *Ex Ante*, IPEA;

Evidências

a) Estratégia Brasileira de Inteligência Artificial (Portaria MCTI 4.979/2021).

b) Documento produzido pela Consultoria contratada no Âmbito da elaboração da Estratégia Brasileira de IA (Peça 626).

Causas

a) Não identificadas.

Efeitos

a) Dificuldade de validar os materiais que embasaram a elaboração da EBIA pela sociedade;

b) Dificuldade de avaliar se as soluções propostas são adequadas aos problemas/oportunidades relacionados ao uso de IA;

c) Dificuldade de entender como a EBIA organiza os processos organizacionais e o fluxo de insumos e processos de maneira a gerar produtos, resultados e impactos;

a) Falta de clareza acerca dos produtos e efeitos esperados da política.

Proposta de Encaminhamento

a) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

Achado IV – **As estruturas de governança e gestão necessárias à implementação da política não estão formalmente institucionalizadas na EBIA**

Situação encontrada

151. Políticas lançadas sem uma clara definição de quem faz o quê, com que grau de autonomia, com que formas de monitoramento e controle – e sem muitas vezes especificar seu passo a passo operacional e sem planejamento de seus riscos – geram incertezas, brechas e pontos de travamento ("gargalos"). Sendo assim, para mitigar tais riscos, a política pública deve vir acompanhada pela criação de instâncias de gestão e governança.

152. De acordo com o Guia Prático de Análise *Ex Ante,* a gestão diz respeito à maneira como os atores determinam e operam os insumos e processos que geram produtos, de forma a alcançar resultados positivos. Ela também se debruça sobre os aspectos que foram especificados no modelo lógico, organizando soluções técnicas e tecnológicas capazes de garantir as entregas dos produtos com qualidade e no tempo esperado.

153. Já os modelos de governança definem quem comanda, quem coordena, quem acompanha (supervisiona) e quem gera a política. A governança de uma política define as competências e as atribuições de sua execução e a responsabilidade por seus ajustes. Um modelo estabelece quem pode o quê (competências) e quem faz o quê (atribuições).

154. Importante ressaltar que parte significativa das definições sobre a gestão e a governança devem vir previamente institucionalizadas, ou seja, inscritas nos próprios normativos dos programas, contendo, no mínimo, conforme previsto no Decreto 9.203/2017, informações sobre: (i) se a atuação governamental conjunta é formalmente definida; (ii) se os papéis e responsabilidades das partes interessadas estão claramente definidos; e (iii) se os processos decisórios da atuação conjunta estão definidos.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

155. Por essa perspectiva, analisando o conteúdo da EBIA, verifica-se que não está estabelecida nenhuma estrutura de governança e gestão que satisfaça os requisitos acima. Quanto a essa ausência, vale dizer que a publicação da EBIA sem a definição de arranjos institucionais que condicionam a forma pela qual as políticas são formuladas, implementadas e avaliadas prejudica o funcionamento de instrumentos como a *accountability* (responsabilização e prestação de contas) e transparência da política pública.

156. Nada obstante, de acordo com a resposta do MCTI ao questionário de autoavaliação (peça 627), tal estrutura de governança será construída mediante a elaboração de um regimento interno, disciplinando a competência do MCTI referente ao art. 2º, inciso II da Portaria 4.617/2021, que instituiu a Estratégia Brasileira de Inteligência Artificial e seus eixos temáticos:

> Art. 2º Caberá ao Ministério de Ciência, Tecnologia e Inovações, no âmbito de suas atribuições:
> I - criar instâncias e práticas de governança para priorizar, implantar, monitorar e atualizar as ações estratégicas estabelecidas na Estratégia Brasileira de Inteligência Artificial;

157. Sendo assim, entende-se que a elaboração do regimento supramencionado deverá envolver as estruturas, os processos, os mecanismos, os princípios, as regras e as normas que influenciam a governança e a gestão de uma política pública.

Critérios

a) Decreto 9.203/2017, art. 4º, inciso IV, VIII e X;

b) Avaliação de Políticas Públicas - Guia Prático de Análise *Ex Ante*, IPEA.

Evidências

a) Estratégia Brasileira de Inteligência Artificial (Portaria MCTI 4.979/2021).

Causas

a) Não identificadas.

Efeitos

a) Ausência de estrutura atuante de coordenação das políticas federais de fomento à inovação sob uma perspectiva integrada de governo;

b) Inoperância de instâncias responsáveis pela política;

c) Os atores governamentais envolvidos com a política pública podem não atuar de maneira coerente e coordenada entre si;

d) Ausência de mecanismos de avaliação e monitoramento de resultados.

Proposta de Encaminhamento

e) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

Achado V – **A estrutura de monitoramento e avaliação da EBIA não está formalmente definida em norma**

Situação encontrada

158. Monitoramento é uma atividade gerencial interna, realizada sistematicamente durante o período de execução e operação, para se saber como uma intervenção evolui ao longo do tempo, através de dados da gerência da política pública sobre metas iniciais, indicadores e resultados associados aos programas. O sistema de monitoramento e avaliação, por sua vez, consiste no conjunto de estruturas, processos, tecnologias e recursos desenhados e operacionalizados por atores públicos e/ou privados, os quais são responsáveis por apresentar evidências quantificáveis do problema; de sua evolução, frente as intervenções; e de seus resultados e impactos ao longo do tempo.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

159. Nada obstante, necessário consignar que, embora o MCTI tenha informado que tal estrutura será construída mediante um regimento interno (peça 627), o fato é que a EBIA não definiu formalmente em norma uma estrutura de monitoramento e avaliação, contrariando o disposto no Decreto 9.203/2017, art. 4º, inciso III.

160. Prever a estrutura de monitoramento no corpo do documento que instituí a política é necessário, pois a avaliação das políticas públicas deve assegurar o aprendizado e aperfeiçoamento contínuos, e, também, criar condições para que haja *accountability*, envolvendo as dimensões de transparência da ação pública e responsabilização perante a sociedade.

161. Considera-se boa prática que complementa o sistema de monitoramento: (i) o estabelecimento de formas de revisão, fixando a periodicidade da avaliação de desempenho da política; (ii) a definição de uma cadeia de indicadores de: insumos/recursos, processos, produtos, resultados e impactos para acompanhar o desempenho da política; e (iii) o estabelecimento de um cronograma, ou marcos temporais, de verificação do desempenho da política.

162. Além disso, é importante verificar se o sistema de monitoramento prevê a elaboração de relatórios de acompanhamento de implementação e progresso, a fim de se constatar se foi efetivamente implementado um plano de monitoramento e avaliação para acompanhar o progresso e o alcance das metas e objetivos estabelecidos para a política pública. Nesse contexto, merece registro a iniciativa do Governo Alemão, que após dois anos da publicação de sua estratégia nacional de IA em 2018, publicou um documento de acompanhamento e atualização da estratégia, em 2020[21].

163. Por fim, pressupõe-se que o monitoramento da política deve ocorrer até o seu término. Sobre este assunto, importa dizer que a EBIA não estabeleceu um prazo final para implementação de suas ações. Dessa maneira, entende-se que a ausência de uma data final para cumprimento das ações e iniciativas previstas, somada à ausência de estrutura de monitoramento implica na dificuldade em compreender os resultados pretendidos com a implementação da EBIA e a identificação de interdependências e obstáculos, o que deveria ter sido evitado pelo MCTI.

Critérios

a) Decreto 9.203/2017, art. 4º, inciso III.

Evidências

a) Estratégia Brasileira de Inteligência Artificial (Portaria MCTI 4.979/2021).

Causas

a) Não identificadas.

Efeitos

a) Impossibilidade de verificar se a política pública opera de forma consistente, em conformidade com o que foi definido durante sua estruturação;

b) Impossibilidade de verificar se o desempenho é satisfatório, em termos de eficiência e alcance de objetivos;

c) Impossibilidade de verificar se há problemas de formulação e/ou implementação que precisam ser tempestivamente tratados.

Proposta de Encaminhamento

d) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

[21] Disponível em https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

28 Inteligência Artificial no Controle

29 Utilização de IA na atividade de controle

164. As tecnologias de IA, em especial as referentes à aprendizagem de máquina, já vêm sendo utilizadas e proporcionando maior produtividade e precisão às atividades de controle. Historicamente, a auditoria dependia de amostragem para identificar os riscos. Essa abordagem não leva em consideração a maioria das informações disponíveis e depende do processamento manual. Com o aprendizado de máquina, os auditores podem tratar e analisar grandes volumes de dados para identificar anomalias e transações de alto risco que podem ser posteriormente analisadas mais detalhadamente.

165. Conforme argumenta a Intosai, a IA pode facilitar o processo de auditoria pública, pois ajuda a realizar muitas tarefas demoradas com mais eficiência. Nesse contexto, a IA pode ser aplicada em áreas como[22]:

a) Automação de processos robóticos (RPA): muitas tarefas de auditoria iterativa podem ser executadas com mais eficiência usando a RPA. Depois que os dados das entidades auditadas são carregados no sistema do órgão fiscalizador, a RPA pode identificar inconsistências e *outliers* que os auditores humanos podem então resolver;

b) Ferramentas de otimização de pesquisa: os algoritmos de IA podem ser usados para classificar e agrupar entidades, enquanto um mecanismo de risco projetado por IA pode calcular uma pontuação na natureza crescente do risco;

c) Rede Neural Artificial: as redes neurais artificiais reconhecem e memorizam dados ou padrões de transação, de forma que com base em problemas ou observações identificadas em auditorias anteriores, podem-se detectar casos semelhantes;

d) Extração de informações e mineração de dados: a IA pode ser usada para coletar, agrupar, extrair e analisar receitas e despesas do governo armazenadas em vários sistemas. Os algoritmos também podem confrontar dados de importação e exportação de sistemas integrados para identificar anomalias rapidamente;

e) Processamento de linguagem natural: esses algoritmos podem automatizar as classificações de observação de auditoria com base em dados históricos para aprender critérios e aplicar regras semelhantes para automação de tarefas.

166. Dessa forma, o uso de tecnologias de IA permite que o auditor se concentre nas áreas de maior risco na auditoria, reduzindo os custos envolvidos na atividade e indo ao encontro do momento atual por que passam os órgãos de controle da Administração Pública no tocante à diminuição de seus quadros de colaboradores.

167. Nesse contexto, o TCU vem desenvolvendo soluções que visam apoiar as diversas atividades de auditoria como os sistemas Alice (Análise de Licitações e Editais), Monica (Monitoramento Integrado para o Controle de Aquisições), Adele (Análise de Disputa em Licitações Eletrônicas), Sofia (Sistema de Orientação sobre Fatos e Indícios para o Auditor), Carina (Crawler e Analisador de Registros da Imprensa Nacional), e Ágata (Aplicação Geradora de Análise Textual com Aprendizado).

168. O Apêndice XI apresenta o detalhamento de algumas dessas soluções, bem como iniciativas de outros órgãos de controle na utilização de IA no suporte às suas atividades fim.

30 Soluções de IA para a atividade de controle

169. Se por um lado a inteligência artificial tem o potencial de aumentar os ganhos em eficiência, melhorar serviços e reduzir custos, por outro, ela também apresenta novos desafios e riscos, conforme apresentado na seção 19 – Considerações sobre o uso de IA. Tais desafios envolvem questões como a segurança de dados, a possibilidade de tratamento automatizado desigual e a produção em massa de decisões institucionais incorretas e/ou discriminatórias.

170. Nesse contexto, é necessário o uso de controles pelos gestores públicos para garantir a confiabilidade dos sistemas de IA, assim como é necessário que os órgãos de controle auditem as

[22] Conteúdo disponível em http://intosaijournal.org/artificial-intelligence-preparing-for-the-future-of-audit/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

aplicações baseadas em algoritmos de IA. Dessa forma, várias entidades que atuam no segmento de controle já desenvolveram propostas e frameworks que visam endereçar as questões relativas à asseguração da utilização ética e segura da IA.

171. Nesse sentido, dado o potencial de desenvolvimento e implementação de soluções de IA pela APF, é importante que o corpo técnico do TCU esteja devidamente capacitado para o adequado acompanhamento e fiscalização das referidas ações. Assim, entende-se que seja necessário a elaboração de um programa de capacitação específico sobre o tema, o qual poderá ter como base os diversos referenciais já existentes, a seguir descritos.

31 GAO - An Accountability Framework for Federal Agencies and Other Entities

172. A Instituição Superior de Auditoria dos Estados Unidos (*U.S. Government Accountability Office* - GAO) publicou um framework para governança em Inteligência Artificial intitulado "Artificial Intelligence: *An Accountability Framework for Federal Agencies and Other Entities*"[23]. O documento apresenta uma estrutura de responsabilidade de IA projetada para garantir a prestação de contas e o uso responsável da IA em processos de negócio do governo.

173. A publicação define uma estrutura que (i) serve como um guia prático para avaliar e auditar sistemas de IA; (ii) é flexível o suficiente para se adaptar às tecnologias em evolução; (iii) permite que não especialistas auditem sistemas de IA; e (iv) estabelece um *benchmark* para avaliar a segurança, justiça e eficácia em sistemas de IA implantados pelo governo.

174. A estrutura é organizada em torno de quatro princípios complementares que tratam sobre governança, dados, desempenho e monitoramento. Cada um destes componentes inclui itens de verificação, procedimentos de auditoria e os tipos de evidências a serem coletadas. Também são descritas as principais práticas-chaves que as entidades de governo devem assegurar na implementação e utilização dos sistemas de IA. Cada prática inclui um conjunto de questões e controles que devem ser verificados pelos auditores.

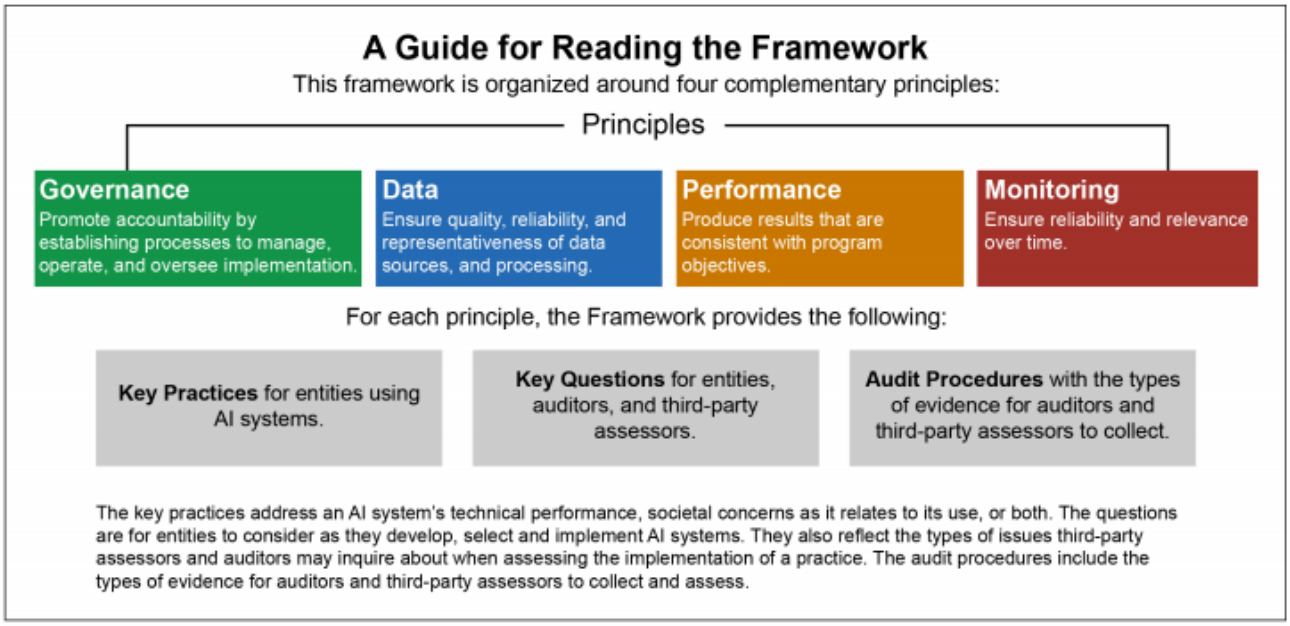


Figura 27 – Detalhamento da estrutura do Framework do GAO

32 ISSAI - Auditing machine learning algorithms

175. Foi publicado no âmbito do *The International Standards of Supreme Audit Institutions* (ISSAI) um artigo[24] que propõe uma abordagem metodológica à realização de auditoria pelas Instituições Superiores de Auditoria (SAIs) sobre algoritmos de aprendizado de máquina desenvolvidos pelas entidades

[23] Disponível em https://www.gao.gov/assets/gao-21-519sp.pdf

[24] Disponível em https://www.auditingalgorithms.net/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

governamentais. O objetivo é ajudar as SAIs a auditarem aplicações de IA, abrangendo os riscos relacionados ao uso de modelos de aprendizagem de máquina, além de sugerir possíveis testes para obter evidências.

176. Ressalta-se que o modelo proposto requer conhecimento específico dos auditores em ciência de dados e métodos quantitativos. Além disso, é desejável que o auditor tenha entendimento básico sobre linguagem de programação e infraestrutura de TI. Essa limitação ao público-alvo é em razão do artigo focar em questões que são específicas para aplicações de IA, como a transparência, a autonomia, a igualdade e a justiça das decisões de um algoritmo de ML.

177. O *framework* é estruturado em cinco áreas de auditoria e baseado no modelo de processo de mineração de dados amplamente conhecido como *Cross Industry Standard Process for Data Mining* (CRISP-DM)[25], que inclui todas as fases do ciclo de vida de uma aplicação de IA - desde a compreensão do negócio até a implantação e operação contínua.

178. O modelo vem acompanhado ainda de uma lista de verificação de auditoria , a qual resume os requisitos mínimos que uma organização auditada deve atender. A lista inclui controles e testes de auditoria sugeridos para mitigar possíveis riscos. Registra-se que o modelo pode ser utilizado de forma flexível, uma vez que os auditores, com base em sua avaliação de risco, podem selecionar a(s) etapa(s) que querem auditar ao longo do ciclo CRISP-DM.

33 ICO - Guidance on the AI auditing framework

179. O *Information Commissioner's Office* (ICO), autoridade independente do Reino Unido criada para proteger os dados e informações no que tange ao interesse público, criou um *framework* de auditoria para sistemas de IA. A premissa é que, embora a IA traga diversos benefícios para organizações e indivíduos, ela também apresenta diversos riscos que podem ameaçar o direito à privacidade dos cidadãos.

180. O *framework* tem como principal objetivo garantir que a solução de IA (própria ou de terceiros) processe os dados pessoais de maneira lícita, justa e transparente e é destinado tanto aos especialistas em TI (especialistas em aprendizado de máquina, cientistas de dados, desenvolvedores e engenheiros de software) como aos gestores que lidam com a conformidade dos dados.

181. O *framework* tem como premissa que um gestor deve: (i) ser responsável pela conformidade do sistema; (ii) avaliar o impacto de proteção aos dados e mitigar os riscos; e (ii) documentar e demonstrar as escolhas que foram feitas no projeto.

182. O modelo proposto pelo ICO subdivide-se em duas grandes áreas: i) governança e accountability; e ii) áreas de riscos específicas de IA. Os riscos de IA são tratados em oito áreas específicas:

[25] CRISP-DM: https://pt.wikipedia.org/wiki/Cross_Industry_Standard_Process_for_Data_Mining

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

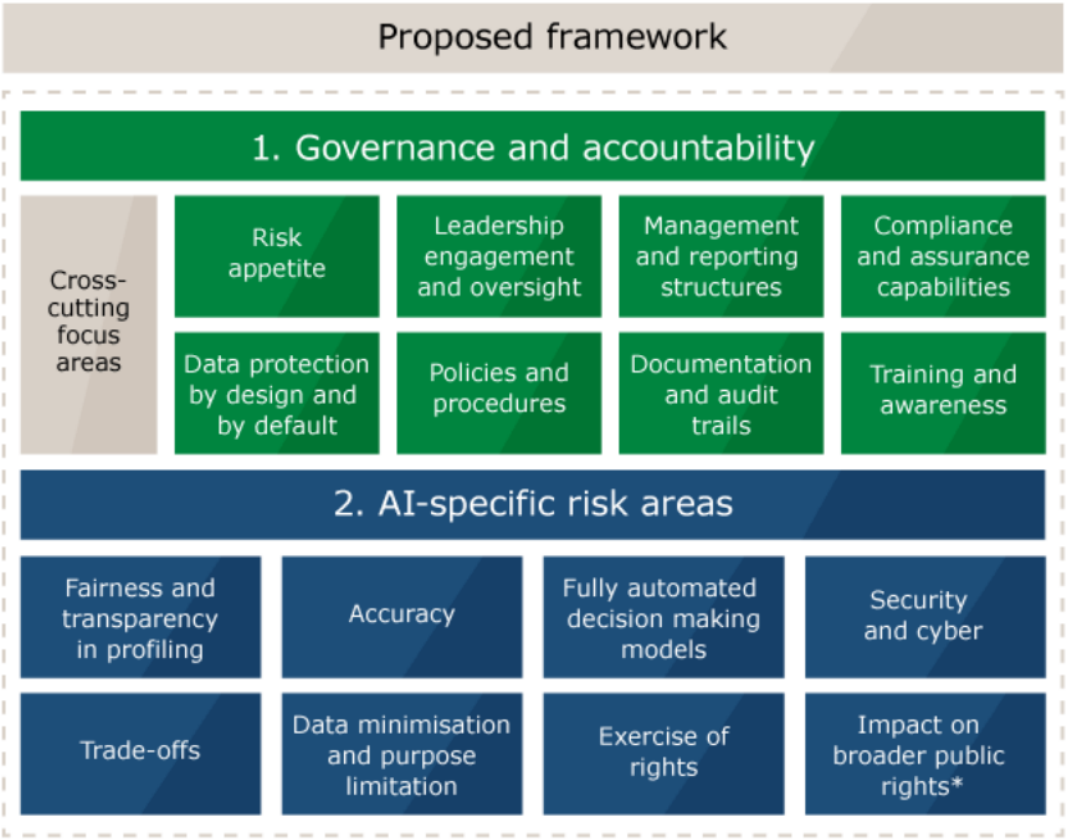


Figura 28 - Detalhamento da estrutura do Framework do ICO

183. Para cada área de risco, é descrito: (i) os riscos envolvidos; (ii) como a IA pode aumentar sua probabilidade e / ou impacto; e (iii) algumas medidas possíveis que a organização pode usar para identificar, avaliar, minimizar, monitorar e controlar esses riscos.

184. É ainda disponibilizado um *toolkit* projetado para fornecer suporte prático para auditoria, que consiste na avaliação de 61 controles relacionados às oito áreas de riscos e às fases de desenvolvimento de soluções de IA (https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/about-this-guidance).

34 Comissão Europeia - Assessment List for Trustworthy Artificial Intelligence

185. Na mesma linha do ICO, a Comissão Europeia, por meio do *High-Level Expert Group on Artificial Intelligence* (AI HLEG) publicou o documento "*Assessment List for Trustworthy Artificial Intelligence – (ALTAI)*"[26] com o propósito de auxiliar as organizações a avaliarem se o sistema de IA que está sendo desenvolvido, implantado, adquirido ou usado cumpre os sete requisitos de Inteligência Artificial Confiável, conforme especificado pelas diretrizes de ética da Comissão Europeia.

186. O conceito de IA confiável foi introduzido nas Diretrizes de ética para Inteligência Artificial Confiável (IA) e é baseado em sete requisitos principais: i) Agência Humana e Supervisão; ii) Robustez técnica e segurança; iii) Privacidade e governança de dados; iv) Transparência; v) Diversidade, não discriminação e justiça; vi) Bem-estar ambiental e social; e vii) Prestação de contas.

187. Os princípios de IA são traduzidos em uma lista de verificação dinâmica por meio de um questionário de autoavaliação, que orienta, na prática, aos desenvolvedores de IA a cumprirem os princípios contidos nos sete requisitos básicos.

[26] Disponível em https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

188. Cumpre observar que a lista de avaliação para IA confiável pode ser utilizada de forma flexível, uma vez que as organizações podem escolher os elementos relevantes para o sistema de IA específico ou adicionar elementos que acharem relevantes.

189. O guia de autoavaliação é destinado a equipes multidisciplinares, envolvendo especialista em IA, cientistas de dados, usuários do sistema, área de negócio e controle interno. Assim, a ALTAI auxilia as organizações a entenderem o que é IA confiável, em particular quais riscos um sistema de IA pode gerar e como minimizar esses riscos e, ao mesmo tempo, maximizar os benefícios da IA.

35 Transparência Brasil - Estrutura de avaliação de riscos a direitos e de transparência

190. A organização Transparência Brasil elaborou uma "Estrutura de avaliação de riscos a direitos e de transparência"[27] para o uso de IA pelo poder público, o qual permite entender a transparência necessária em processos de desenvolvimento, aquisição e implementação de IA pelo estado, de forma a garantir o devido controle social democrático.

191. O modelo funciona como um guia para encontrar pontos críticos e, a partir disso, elaborar recomendações ao governo, exigir mais transparência pública, correções ou testes para garantir não-discriminação e erros algorítmicos, ou até a eventual descontinuidade de alguma ferramenta cujo risco seja incontornável.

192. A estrutura de avaliação é dividida da seguinte forma: (i) riscos a direitos pela natureza da ferramenta; (ii) riscos a direitos por discriminação algorítmica; (iii) riscos ao direito à privacidade; (iv) potencial abuso autoritário do espaço cívico; e (v) transparência no uso de sistemas de inteligência artificial. Cada uma dessas áreas tem itens de verificação associados. Assim, avalia-se o nível de transparência existente como instrumento necessário para acompanhar todo o processo de utilização pública de sistemas de IA, tendo em vista os riscos a direitos elencados:

Tabela 6 - Avaliações a serem realizadas

I - No fluxo de utilização da ferramenta, há supervisão humana em todas as decisões sugeridas ou tomadas pelo algoritmo?	VIII - Existem evidências de que este algoritmo funcione no ambiente em que ele está sendo utilizado? As evidências são baseadas em experimentos científicos relevantes?
II - Em caso de erro do algoritmo corrigido por humano, essa informação é usada para aprimoramento do algoritmo?	IX - Existe regulamentação específica sobre o uso deste algoritmo na área em que ele está sendo aplicado? Quais são? Se não, existe apoio de uma equipe jurídica especializada para garantir que haja respaldo jurídico?
III - A ferramenta, por sua natureza, pode impactar direitos fundamentais, seja por erro ou por design do seu algoritmo, seja direta ou indiretamente? Se sim, quais?	X - Existe uma equipe técnica que acompanha e monitora a implementação deste algoritmo? Esta equipe contém funcionários do órgão capazes de analisar criticamente os caminhos tomados?
IV - Quais grupos ou populações serão afetadas por esse algoritmo? Esses segmentos foram considerados no processo de treinamento da ferramenta?	XI - A equipe responsável pelo desenvolvimento do algoritmo inclui especialistas da área na qual o algoritmo será aplicado?
V - Existem órgãos ou pessoas dentro da entidade que utilizam o algoritmo que podem prestar informações sobre seu uso às autoridades competentes?	XII - Um comitê de ética acompanha/acompanhou o desenvolvimento do algoritmo e os ritos envolvidos na coleta e uso de dados?
VI - O impacto negativo é criado ou acentuado a partir do algoritmo?	

[27] Disponível em https://www.transparencia.org.br/downloads/publicacoes/Estrutura_Avaliacao_Risco.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

VII - Esse algoritmo é imprescindível para atingir o objetivo apontado? Se ele tem o potencial de afetar o exercício de direitos fundamentais ou de se colocar como intermediário para acesso a eles, existem formas alternativas para exercício de tal direito? Se sim, quais?	

36 LIAA-3R - Diretrizes de auditabilidade e conformidade no desenvolvimento e testes de soluções de IA

193. O Laboratório de Inteligência Artificial Aplicada do Tribunal Regional Federal da 3ª Região (LIAA-3R), com o objetivo de incentivar a pesquisa e contribuir para o aprimoramento dos serviços judiciais e administrativos, elaborou o documento[28] que consolida as diretrizes de auditabilidade e conformidade, para orientar o desenvolvimento de modelos de inteligência artificial.

194. O documento consolida os critérios legais que devem ser seguidos no âmbito do LIAA-3R e é organizado em três partes. Na primeira delas, intitulada "Diretrizes Gerais", procurou-se seguir a divisão de matérias e os preceitos da Resolução CNJ 332/2020, da Portaria do CNJ 271/2020, assim como do documento intitulado *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*.

195. Na segunda parte, intitulada "Diretrizes Específicas", abordam-se algumas das regras aplicáveis especificamente à Justiça Federal da 3ª Região e a seus laboratórios de inovação. A terceira e última parte cuida de algumas diretrizes relacionadas à observância da LGPD.

196. Além disso, o documento apresenta duas listas de verificação, na forma de questões técnicas, destinadas à equipe de desenvolvimento e ao Grupo de Validação Ética e Jurídica do LIAA-3R (GVE), com a finalidade de avaliar se requisitos específicos foram observados, os quais envolvem as seguintes áreas:

Tabela 7 – Diretrizes LIAA-3R – Questões Técnicas

Equipe de desenvolvimento	GVE
1. Questões relacionadas ao escopo e à finalidade do projeto	1. Ação e supervisão humanas
2. Questões relacionadas aos usuários e ao contexto de uso:	2. Solidez técnica e segurança
3. Questões relacionadas aos *datasets* utilizados	3. Privacidade e gestão de dados
4. Questões relacionadas à arquitetura e tecnologias adotadas	4. Transparência
5. Questões relacionadas aos testes da solução	5. Diversidade, não discriminação e equidade (*fairness*)
	6. Bem-estar social e ambiental
	7. Prestação de Contas (*Accountability*)

[28] Disponível em https://www.trf3.jus.br/documentos/adeg/Inova/LIAA-3R/Aquivos/Manual_de_Conformidade_-_final_13.abril.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

	8. Questões relacionadas às aprovações obtidas
	9. Questões relacionadas a possíveis conflitos de interesses
	10. Questões relacionadas à equipe do projeto

37 Conclusão

197. O presente levantamento foi proposto com o objetivo geral de conhecer a atual utilização das tecnologias de Inteligência Artificial na Administração Pública Brasileira, avaliar os riscos e oportunidades da Estratégia Brasileira de Inteligência Artificial e verificar os principais aspectos referentes ao uso de tecnologias de IA nas atividades de controle.

198. Nesse sentido, o trabalho buscou inicialmente descrever na Seção 2 – Visão Geral do Objeto os principais conceitos, tecnologias envolvidas, domínios de aplicação da tecnologia bem como considerações que devem ser referentes à ética, segurança e privacidade dos dados quando do desenvolvimento e implementação de soluções de IA.

199. Com relação à utilização de IA na Administração Pública Federal, a Seção 3 abrangeu os principais benefícios na utilização de soluções de IA no Setor Público, entre eles a redução de gastos em decorrência da maior eficiência operacional, o aprimoramento dos serviços disponíveis ao cidadão e a disponibilização de novos serviços, além das vantagens da utilização dos dados para tomadas de decisão pelos entes governamentais.

200. Nesse contexto, verificou-se em estudo que no Brasil, apesar do país apresentar diferenciais competitivos que incentivam a maior implementação e soluções de IA, como a grande disponibilidade e a representatividade dos dados, o Setor Público brasileiro ainda carece de melhores políticas e investimentos no sentido de aprimorar sua infraestrutura, capital humano e capacidade de inovação.

201. Com o objetivo de se obter um panorama abrangente do estado atual de utilização da IA nas organizações da APF, realizou-se pesquisa por meio de questionário eletrônico que foi respondido por 263 organizações. Entre as principais constatações a partir da análise das respostas, tem-se que grande parte das instituições ainda não planejam utilizar soluções de IA em seus processos produtivos ou na disponibilização de serviços aos cidadãos. Entre os principais dificultadores informados para a adoção da tecnologia, foi reportado o baixo número de colaboradores disponíveis com capacitação suficiente para a implementação.

202. A pesquisa também evidenciou que a maior parte das soluções são desenvolvidas por equipe própria dos órgãos, sendo que 48% das organizações pesquisadas não realizaram nenhuma capacitação interna em relação ao desenvolvimento e implementação de tecnologias de IA.

203. Verificou-se também que o Poder Judiciário está à frente em relação à utilização de soluções de IA, que tem beneficiado e contribuído principalmente na diminuição do estoque processual daquele Poder. Também estão sendo implementadas pelo Poder Judiciário estratégias e infraestruturas que permitem o compartilhamento de dados e o intercâmbio de soluções entre os diversos órgãos que o compõem.

204. Entre os principais benefícios apontados com a utilização de IA, as organizações pesquisadas informaram a automação de processos repetitivos com aumento de produtividade, a redução do tempo e custos na execução dos processos, a otimização na alocação de recursos humanos, o aumento da confiabilidade e transparência dos processos e maior tempestividade na resposta aos clientes.

205. Em relação à Estratégia Brasileira de Inteligência Artificial (EBIA), apresentou-se como algumas nações têm tratado o tema, tanto em nível nacional como em nível regional, com o objetivo de se tornarem líderes no segmento.

206. Nesse sentido, a avaliação realizada da EBIA, realizada tendo como base o Referencial para Controle de Políticas Públicas elaborado pelo TCU, evidenciou falhas em sua elaboração que podem

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

comprometer a sua adequada implementação, monitoramento e atingimento dos objetivos. Foram identificados os seguintes achados em relação à EBIA:

a) os objetivos da EBIA não são específicos, mensuráveis, realistas (alcançáveis) e delimitados em um recorte temporal;

b) inexistência da explicitação do estágio de referência inicial ("marco zero", "situação inicial") e da definição de indicadores e metas para subsidiar a avaliação do desempenho e resultados da EBIA;

c) falhas formais na apresentação do modelo lógico e sobre como a EBIA incide sobre os problemas/oportunidades, projetando seus resultados e impactos de longo prazo;

d) as estruturas de governança e gestão necessárias à implementação da política não estão formalmente institucionalizadas na EBIA;

e) a estrutura de monitoramento e avaliação da EBIA não está formalmente definida em norma.

207. Cabe ressaltar que, apesar do disposto nos achados, não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020. Entretanto, com o objetivo de contribuir para o aperfeiçoamento da EBIA, as informações contidas nos achados devem ser encaminhadas ao MCTI, órgão gestor da referida política pública.

208. O Levantamento abordou a utilização a utilização de IA na atividade de controle, identificando casos de uso da solução nas organizações da Administração Pública.

209. O trabalho abordou ainda o controle de soluções de IA pelo Poder Público, abordando os principais referenciais já produzidos a respeito. Em decorrência das várias características específicas que as soluções de IA apresentam em relação aos demais sistemas computacionais, entende-se que deve ser elaborado plano de capacitação específico sobre o tema para o corpo técnico do TCU, em especial da Secretaria de Fiscalização de Tecnologia da Informação.

210. Por fim, de forma a contribuir para o aprimoramento da IA no Brasil, propõe-se encaminhar o presente relatório ao MCTI, para que este, caso julgue necessário, adote as providências pertinentes.

38 Proposta de Encaminhamento

211. Ante todo o exposto, propõe-se o encaminhamento dos autos ao gabinete do Exmo. Relator Aroldo Cedraz com as seguintes propostas:

a) levantar o sigilo dos autos, por conter informações relevantes à Administração Pública;

b) autorizar a Secretaria de Fiscalização de Tecnologia da Informação (Sefti) a divulgar as informações consolidadas constantes deste levantamento, preferencialmente por intermédio de fichas síntese, sumário executivo e infográfico;

c) encaminhar ao Ministério da Ciência, Tecnologia e Inovações (MCTI), à Comissão de Ciência e Tecnologia, Comunicação e Informática da Câmara dos Deputados, à Comissão de Ciência, Tecnologia, Inovação, Comunicação, Inovação e Informática do Senado Federal, bem como aos órgãos discriminados no Apêndice IX do presente relatório, cópia do acórdão que vier a ser proferido, bem como cópia do relatório da unidade técnica e seus respectivos apêndices, destacando que o relatório e o voto que fundamentaram a deliberação ora encaminhada podem ser acessados por meio do endereço eletrônico www.tcu.gov.br/acordaos;

d) arquivar o presente processo, com fulcro no art. 169, inciso V, do Regimento Interno do TCU.

É o relatório.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360192.

VOTO

Trata-se de levantamento com o objetivo de avaliar o estágio atual e perspectivas de utilização de Inteligência Artificial (IA) na Administração Pública Federal (APF), identificar os riscos associados e avaliar a proposta para uma Estratégia Brasileira de Inteligência Artificial (EBIA).

2. A Organização para a Cooperação e Desenvolvimento Econômico (OCDE) define IA como um sistema baseado em máquina que pode, para um determinado conjunto de objetivos definidos pelo homem, fazer previsões, recomendações ou decisões que influenciam ambientes reais ou virtuais.

3. Em termos históricos, considera-se que a inteligência artificial passou a ser pesquisada e desenvolvida de maneira formal em 1956, a partir do trabalho conjunto de cientistas da época que se reuniram no Dartmouth College, nos EUA, para estabelecer os principais conceitos e objetivos relacionados ao tema.

4. Desde então, pode-se dizer que a IA passou por três eras distintas: a simbólica (1956-1974), em que se buscava aplicações que fossem capazes de representar o conhecimento e raciocinar a partir dele; a dos sistemas especialistas (1980-1987 e 1993-2011), em que a lógica humana de raciocínio era embutida diretamente nos algoritmos, sob a forma de regras estáticas; e a do aprendizado de máquina (a partir de 2011), que utiliza técnicas estatísticas para encontrar padrões em quantidades suficientes de dados e, dessa forma, realizar tarefas repetitivas sem a necessidade de constante orientação humana.

5. O uso de IA na administração pública é uma fronteira relativamente nova. Isso pode ser explicado, em certa medida, pela ocorrência de períodos de grande frustração com os resultados da tecnologia, conhecidos como "invernos da IA" – entre 1974 e 1980, com relação aos sistemas simbólicos, e de 1987 a 1993, com os sistemas especialistas. Foi somente a partir de 2011, com o aumento exponencial da capacidade de processamento dos computadores e a criação de modelos poderosos de redes neurais, que a IA destacou-se ao ponto de chamar a atenção de governos ao redor do mundo.

6. À medida que a IA e o aprendizado de máquina evoluem, mais tarefas administrativas e orientadas a processos podem ser automatizadas, com ganhos de eficiência e liberação de profissionais públicos para outras prioridades. Entre as oportunidades reconhecidas, temos a disponibilização de novos serviços, o aprimoramento dos já disponíveis, bem como o amadurecimento e ampliação de tomadas de decisão baseadas em dados.

7. Além disso, o setor público é responsável por definir prioridades, investimentos e regulamentações nacionais a respeito do assunto. Afinal, a IA tem potencial para integrar o processo de formulação de políticas públicas e de implantação de serviços e, de igual modo, proporcionar ganhos exponenciais para empresas no setor privado, desde que seja utilizada de forma ética e responsável.

8. Nesse sentido, foi instalada no Senado Federal, no último dia 30 de março, comissão dedicada a elaborar um projeto para regulação da IA no Brasil, tendo como fundamentos o texto do PL 21/2020, aprovado recentemente pela Câmara dos Deputados, bem como outros dois projetos de lei em tramitação no Senado sobre o mesmo assunto.

9. Entendo que tal providência é especialmente necessária, considerando que a publicação *Government AI Readiness Index* 2020, elaborada pela Oxford Insights e pelo Governo do Canadá, avaliou 172 países quanto à capacidade dos governos nacionais para implementar soluções de IA no fornecimento de serviços públicos aos cidadãos, tendo o Brasil alcançado apenas a 63ª posição no ranking resultante, sendo o 6º colocado entre os países da América Latina e Caribe.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360191.

10. Como a posição no ranking acima sugere, Brasil ainda está em seus primeiros passos. Assim, como dito inicialmente, o presente trabalho buscou identificar e conhecer o estado atual de uso da tecnologia pela Administração Pública.

11. A esse respeito, cabe ressaltar que, embora existam inúmeros exemplos de sistemas de IA baseados em regras em uso no setor público (a exemplo do próprio robô Alice, deste Tribunal), o escopo deste levantamento foi limitado a aplicações de última geração, baseadas em aprendizagem de máquina. A motivação para tal escolha foi o entendimento, por parte da unidade técnica, de que sistemas especialistas (baseados em regras) não poderiam ser classificados como "tecnologias emergentes", temática que vem direcionando trabalhos mais recentes da Sefti, como o que relatei no ano de 2021 acerca do **blockchain**.

12. Embora tal restrição esteja perfeitamente alinhada à estratégia de fiscalização adotada pela unidade instrutiva nos últimos anos, e tenha sido útil para permitir que a equipe se aprofundasse em aspectos específicos de modelos e técnicas de aprendizagem de máquina em uso na administração pública, entendo que será necessário que este Tribunal realize, no futuro, uma avaliação mais abrangente sobre o tema, contemplando sistemas de IA de primeira, segunda e terceira gerações, uma vez que mesmo aplicações mais simples, baseadas em regras e tipologias (a exemplo do já mencionado robô Alice) podem trazer ganhos expressivos de eficiência e eficácia para os órgãos de governo.

13. Registro que as providências nesse sentido serão apresentadas ao final deste voto, como parte da proposta de estratégia para a indispensável continuidade da atuação do Tribunal de Contas da União sobre o assunto, dada sua extrema relevância para o processo de transformação digital de nosso país.

***

14. Passo, portanto, a apresentar breve resumo das principais conclusões do presente trabalho, desenvolvido com grande competência pela equipe da Sefti. Nesse contexto, para maior clareza acerca da abrangência do levantamento em questão, será utilizada a sigla IA/ML sempre que as assertivas forem relacionadas exclusivamente a aplicações de aprendizagem de máquina (*machine learning*).

15. A Sefti aplicou um questionário eletrônico sobre utilização de IA/ML, com o objetivo principal de obter informações sobre a utilização dessa tecnologia pelos órgãos da APF em 2021. Em resposta à solicitação do Tribunal, 263 organizações encaminharam as informações tempestivamente.

16. Como resultado, essa pesquisa identificou 44 ferramentas usadas por órgãos governamentais, sendo 36 delas direcionadas para o público interno e apenas 8 voltadas para o atendimento direto às necessidades do cidadão.

17. Além disso, a equipe também juntou aos autos trabalho de autoria da Fundação Getúlio Vargas no âmbito do Poder Judiciário, que identificou diversos projetos com o uso de IA em tribunais, sendo 29 em desenvolvimento, 7 em fase de testes e 27 em operação. Entretanto, em torno da metade das aplicações eram oriundas de tribunais de justiça estaduais.

18. Tais números apenas confirmam o temos afirmado há tempos sobre a vocação da atividade judicial para o uso da IA, incluído nesse contexto o Tribunal de Contas da União. Afinal, em grande parte dos processos tomamos decisões fundamentadas em entendimentos consolidados, repetidos centenas ou até milhares de vezes nos colegiados, e ainda assim votos e acórdãos são produzidos de forma quase exclusiva com o uso da boa e velha, porém mais lenta e cara, inteligência natural.

19. Por óbvio, esse combate não se mostra insipiente apenas em nossa casa. O questionário aplicado pela Sefti revelou que apenas 28% das organizações pesquisadas declararam usar alguma solução baseada em inteligência artificial, outras 45% responderam que "ainda não", sugerindo a intenção de fazê-lo, e os 27% restantes declararam que não têm previsão de uso da tecnologia.

***

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360191.

20. Para melhor sistematização das respostas ao questionário aplicado, a equipe optou por utilizar o Modelo de Maturidade em Inteligência Artificial proposto pelo Gartner Group. Como resultado, temos que dois terços das organizações consultadas (188) possuem maturidade baixa ou nula em IA, visto que não tratam do assunto ou o fazem em conversas meramente especulativas, sem realizar uma única Prova de Conceito (POC).

21. Nesses casos, verificou-se como impeditivo o simples desconhecimento em relação às oportunidades e benefícios que podem ser alcançados com o uso das referidas tecnologias, o que sugere uma lacuna que perpassa as limitações na força de trabalho e alcança as lideranças dessas organizações.

22. Já com relação às 75 organizações que pontuaram em níveis intermediários e acima no modelo do Gartner, constatou-se que 49 desenvolveram soluções com força de trabalho própria, embora 80% contem com no máximo 5 colaboradores aptos no uso de IA/ML. Em outras 14 instituições o desenvolvimento se deu por meio de profissionais terceirizados, e 12 recorreram a soluções prontas.

23. Ora, em meio a uma situação fiscal grave, e tão permanente que nem pode mais ser chamada de crise, em que a contratação de pessoal preparado por meio de novos concursos é um luxo raro, as lideranças devem estar atentas aos recursos que possibilitam não só o "não retrocesso" de serviços públicos, como a melhoria e expansão daqueles à disposição da sociedade.

24. Nessa linha, de acordo com artigo "*How much time and money can AI save government?*", elaborado pela Deloitte University Press, a automação das tarefas dos funcionários do governo federal americano poderia economizar entre 96,7 milhões e 1,2 bilhão **de horas** anuais.

25. No mesmo sentido, o grupo Singularity classifica a IA – juntamente com o **blockchain**, a realidade aumentada e o metaverso – na categoria de tecnologias exponenciais, assim denominadas por dois motivos: a capacidade de gerar resultados várias ordens de grandeza superiores aos recursos investidos para sua implementação; e o tempo de maturação um pouco mais lento, caracterizado por uma etapa inicial de seis meses a um ano até que os primeiros resultados expressivos comecem a ser percebidos.

26. Diante de tamanho potencial, não é exagero afirmar que a Inteligência Artificial constitui um dos caminhos mais promissores, senão o principal deles, para que a sociedade brasileira receba melhores serviços públicos, apesar da limitação orçamentária que pesa sobre a força de trabalho.

27. No entanto, temo que se repita com a IA fenômeno aparentado ao ocorrido no processo de adoção da computação em nuvem ou até na incorporação de metodologias ágeis, quando a percepção dos riscos da inovação por gestores públicos impediu, por bastante tempo, os tão necessários avanços no setor público.

28. Outra característica verificada pela Sefti é a predominância de processos administrativos no uso de IA/ML, as chamadas áreas meio, em relação aos processos finalísticos. Embora isso não seja um problema em si, uma vez que essa tecnologia pode e deve ser usada para ganhos diversos, a questão reforça a necessidade do envolvimento das lideranças para identificar aplicações que possam oferecer saltos na qualidade e agilidade dos serviços oferecidos diretamente ao cidadão.

29. Por tudo isso, entendo que o TCU deve, mais uma vez, cumprir um papel indutor. Nesse sentido, para dar maior segurança ao gestor proativo e, ao mesmo tempo, estimular aqueles mais reticentes a buscar ativamente os potenciais benefícios ao seu alcance, considero pertinente orientar à Sefti que elabore guia com a definição de diretrizes, parâmetros e eventuais riscos, a fim de auxiliar líderes e gestores públicos no processo de implementação ou contratação de serviços que envolvam o uso de inteligência artificial.

***

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360191.

30. Outra questão abordada no presente levantamento foi a Estratégia Brasileira de Inteligência Artificial (EBIA). A iniciativa de elaborar uma estratégia para o desenvolvimento dessa tecnologia no país está alinhada aos grandes atores globais no assunto, como demonstra o Instituto de Tecnologia e Sociedade do Rio (ITS) que produziu, em 2020, uma compilação dos planos estratégicos de desenvolvimento sobre o assunto em 15 países, entre eles Estados Unidos, Japão, Singapura, China, Taiwan, França, Reino Unido, Índia e Coreia do Sul.

31. A EBIA tem seis principais objetivos: (i) contribuir para a elaboração de princípios éticos para o desenvolvimento e uso de IA responsáveis; (ii) promover investimentos sustentados em pesquisa e desenvolvimento em IA; (iii) remover barreiras à inovação em IA; (iv) capacitar e formar profissionais para o ecossistema da IA; (v) estimular a inovação e o desenvolvimento da IA brasileira em ambiente internacional; e (vi) promover ambiente de cooperação entre os entes públicos e privados, a indústria e os centros de pesquisas para o desenvolvimento da Inteligência Artificial.

32. O texto da Portaria MCTI 4.979/2021 registra que a EBIA é considerada uma política pública, e deve ser constantemente acompanhada, avaliada e ajustada, pois o ritmo da evolução tecnológica da Inteligência Artificial tende a se acelerar. Com base em tal enquadramento, a presente avaliação foi realizada segundo os critérios constantes do Referencial para Controle de Políticas Públicas elaborado pelo TCU.

33. No entanto, como a EBIA foi publicada durante a execução do presente levantamento e, portanto, ainda não produziu seus efeitos, a Sefti optou por verificar apenas os quatro primeiros blocos de controle: diagnóstico do problema, análise de alternativas, desenho da política pública e a estruturação da governança e gestão, que envolve a definição das estruturas de gestão de riscos, monitoramento e avaliação da política pública.

34. A fiscalização revelou lacunas que comprometem a implementação e o monitoramento das ações, bem como o alcance dos objetivos. Dos pontos verificados pela Sefti, destaco o que segue.

35. Primeiro, os objetivos da EBIA não são específicos, mensuráveis, realistas (alcançáveis) e delimitados em um recorte temporal. Na prática, podem ser considerados apenas diretrizes, que são mais abertas por natureza. Em consequência, as ações estratégicas não declaram explicitamente a quais objetivos estão associadas, dificultando a avaliação do nível de coerência entre tais elementos.

36. Segundo, não está claro o estágio de referência inicial, o chamado “marco zero”, nem os indicadores e metas para avaliação do desempenho. Sem valores da linha de base acerca da realidade a ser aprimorada, não se poderá desenvolver qualquer análise relevante dos resultados alcançados. O diagnóstico da situação atual, presente na EBIA, apresenta apenas dados gerais sobre do uso de IA no país, sem definir o que se pretende medir ou priorizar.

37. Por último, há falhas na apresentação do modelo lógico da EBIA e sobre como ela incide sobre os problemas identificados ou as oportunidades priorizadas, uma vez que não há discriminação de entradas, atividades, produtos, impactos e resultados.

38. Em que pesem as lacunas identificadas acima na Estratégia Brasileira de Inteligência Artificial, deixo de propor recomendações neste momento, em respeito ao disposto no artigo 13 da Resolução-TCU 315/2020, que veda esse tipo de encaminhamento em deliberações proferidas em processos de fiscalização autuados sob o instrumento levantamento.

***

39. Um último aspecto tratado no presente levantamento e que merece destaque refere-se aos impactos da inteligência artificial sobre a atividade de controle.

40. Como se sabe, o TCU desenvolveu diversas soluções baseadas em IA ao longo dos últimos anos que suportam diretamente o planejamento e a execução de fiscalizações, bem como a instrução de processos nesta Corte, a exemplo dos robôs Alice, Mônica, Adele e Ágata, ou das ferramentas Detecta

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360191.

e Sofia. Tais soluções têm sido utilizadas de forma cada vez mais intensa, proporcionando maior produtividade e precisão às atividades de controle.

41. Para além do uso interno, é importante que este Tribunal esteja atento à crescente adoção de algoritmos e aplicações de inteligência artificial pela Administração Pública Federal, com a consequente necessidade de que tais soluções sejam, eventualmente, auditadas de modo a assegurar sua confiabilidade e segurança, além de avaliar aspectos de especial relevância para a IA, tais como o uso ético das informações, a ausência de preconceitos ou distorções embutidas nos algoritmos e a transparência de sua operação.

42. Nesse sentido, a equipe da Sefti identificou diversos guias, modelos e **frameworks** criados por organizações públicas e do terceiro setor, no Brasil e no mundo, que poderiam servir de referência para o desenvolvimento de metodologia do TCU, a ser aplicada em futuras fiscalizações que tenham como objetivo avaliar sistemas e algoritmos baseados em inteligência artificial e aprendizagem de máquina.

43. Diante da constatação de que, quando da conclusão do levantamento ora sob análise, em setembro de 2021, já havia pelo menos 75 organizações federais utilizando ativamente soluções baseadas em IA, e que esse número tende a crescer exponencialmente nos próximos anos, entendo ser urgente o desenvolvimento e a validação de metodologia para auditoria de tais sistemas por esta Corte, dando continuidade ao trabalho iniciado pela Sefti.

44. Por esse motivo, proponho que seja determinada a realização de nova etapa do presente trabalho, tendo agora como objetivos o estabelecimento e a validação de referencial teórico e metodológico próprio do Tribunal de Contas da União para auditoria de aplicações e algoritmos de inteligência artificial, abrangendo tanto sistemas especialistas baseados em regras como soluções de aprendizagem de máquina.

45. Uma vez construído e validado tal referencial, por meio de auditoria-piloto sobre solução de IA a ser definida posteriormente, registro desde já minha sugestão para que este Tribunal realize Fiscalização de Orientação Centralizada sobre as aplicações de IA mais relevantes ou críticas do Governo Federal e do Poder Judiciário, a serem identificadas com base em novo levantamento sobre a situação do uso de IA pelo setor público.

46. Por fim, considero pertinente que o presente processo seja convertido em acompanhamento, de modo a viabilizar as ações que acabo de propor, bem como o indispensável controle concomitante sobre a implementação da EBIA, inclusive com a proposição de eventuais ajustes necessários para que a Estratégia possa cumprir a nobre missão a que se propõe.

Ante o exposto, Voto para que o Tribunal adote a minuta de Acórdão que ora submeto à apreciação deste Colegiado.

TCU, Sala das Sessões, em 25 de maio de 2022.

AROLDO CEDRAZ
Relator

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360191.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360191.

ACÓRDÃO Nº 1139/2022 – TCU – Plenário

1. Processo TC 006.662/2021-8.
2. Grupo I – Classe de Assunto: V – Levantamento de Auditoria.
3. Interessados/Responsáveis: não há.
4. Órgãos/Entidades: Advocacia -Geral da União; Agência Brasileira de Desenvolvimento Industrial; Agência Brasileira de Inteligência; Agência Brasileira Gestora de Fundos Garantidores e Garantias S.a.; Agência Espacial Brasileira; Agência Nacional de Águas; Agência Nacional de Assistência Técnica e Extensão Rural; Agência Nacional de Aviação Civil; Agência Nacional de Energia Elétrica; Agência Nacional de Mineração; Agência Nacional de Saúde Suplementar; Agência Nacional de Telecomunicações; Agência Nacional de Transportes Aquaviários; Agência Nacional de Transportes Terrestres; Agência Nacional de Vigilância Sanitária; Agência Nacional do Cinema; Agência Nacional do Petróleo, Gás Natural e Biocombustíveis; Amazonas Geração e Transmissão de Energia S.a.; Amazônia Azul Tecnologias de Defesa S.a.; Associação das Pioneiras Sociais; Autoridade Portuaria de Santos S.a; Banco Central do Brasil; Banco da Amazônia S.a.; Banco do Brasil S.a.; Banco do Nordeste do Brasil S.a.; Caixa Econômica Federal; Casa da Moeda do Brasil; Centrais de Abastecimento de Minas Gerais S. A.; Centrais Eletricas Brasileiras Sa; Centrais Elétricas do Norte do Brasil S.a.; Centro de Tecnologia da Indústria Química e Têxtil do Senai/rj; Centro Federal de Educação Tecnológica Celso Suckow da Fonseca; Centro Federal de Educação Tecnológica de Minas Gerais; Centro Nacional de Tecnologia Eletrônica Avançada S.a; Colégio Pedro Ii; Comando da Aeronautica; Comando da Marinha; Comando do Exército; Comissão de Valores Mobiliários; Comissão Nacional de Energia Nuclear; Companhia Brasileira de Trens Urbanos; Companhia de Desenvolvimento dos Vales do São Francisco e do Parnaíba; Companhia de Geração e Transmissão de Energia Elétrica do Sul do Brasil; Companhia de Pesquisa de Recursos Minerais; Companhia Docas do Ceará; Companhia Docas do Espírito Santo; Companhia Docas do Pará; Companhia Docas do Rio de Janeiro; Companhia Docas do Rio Grande do Norte; Companhia Hidro Elétrica do São Francisco; Companhia Nacional de Abastecimento; Conselho Administrativo de Defesa Econômica; Conselho da Justiça Federal; Conselho Nacional de Desenvolvimento Científico e Tecnológico; Conselho Nacional de Justiça; Conselho Nacional do Ministério Público; Conselho Superior da Justiça do Trabalho; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Defensoria Pública da União; Departamento de Polícia Federal; Departamento de Polícia Rodoviária Federal; Departamento Nacional de Infraestrutura de Transportes; Departamento Nacional de Obras Contra As Secas; Eletrobrás Termonuclear S.a.; Empresa Brasileira de Correios e Telegrafos - ECT; Empresa Brasileira de Infraestrutura Aeroportuária; Empresa Brasileira de Pesquisa Agropecuária; Empresa Brasileira de Serviços Hospitalares; Empresa de Pesquisa Energética; Empresa de Tecnologia e Informações da Previdência - Dataprev; Empresa Gerencial de Projetos Navais; Financiadora de Estudos e Projetos; Fundação Alexandre de Gusmão; Fundação Biblioteca Nacional; Fundação Cultural Palmares; Fundação Escola Nacional de Administração Pública; Fundação Instituto Brasileiro de Geografia e Estatística; Fundação Instituto de Pesquisa Econômica Aplicada; Fundação Jorge Duprat Figueiredo, de Segurança e Medicina do Trabalho; Fundação Nacional de Artes; Fundação Nacional de Saúde; Fundação Nacional do Índio; Fundação Oswaldo Cruz; Fundação Universidade de Brasília; Fundação Universidade do Amazonas; Fundação Universidade Federal da Grande Dourados; Fundação Universidade Federal de Ciências da Saúde de Porto Alegre; Fundação Universidade Federal de Mato Grosso do Sul; Fundação Universidade Federal de Ouro Preto; Fundação Universidade Federal de Rondônia; Fundação Universidade Federal de São Carlos; Fundação Universidade Federal de São João Del Rei; Fundação Universidade Federal de Sergipe; Fundação Universidade Federal de Uberlândia; Fundação Universidade Federal de Viçosa; Fundação Universidade Federal do Abc; Fundação Universidade Federal do Acre; Fundação Universidade Federal do Amapá; Fundação Universidade Federal do Maranhão; Fundação Universidade Federal do Pampa; Fundação Universidade Federal do

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360190.

Piauí; Fundação Universidade Federal do Rio Grande; Fundação Universidade Federal do Tocantins; Fundação Universidade Federal do Vale do São Francisco; Fundo Nacional de Desenvolvimento da Educação; Furnas Centrais Elétricas S.a.; Hospital Nossa Senhora da Conceição S.a.; Indústrias Nucleares do Brasil S.a.; Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis; Instituto Chico Mendes de Conservação da Biodiversidade; Instituto do Patrimônio Histórico e Artístico Nacional - IPHAN; Instituto Federal de Educação, Ciência e Tecnologia Farroupilha; Instituto Federal de Educação, Ciência e Tecnologia Baiano; Instituto Federal de Educação, Ciência e Tecnologia Catarinense; Instituto Federal de Educação, Ciência e Tecnologia da Bahia; Instituto Federal de Educação, Ciência e Tecnologia da Paraíba; Instituto Federal de Educação, Ciência e Tecnologia de Alagoas; Instituto Federal de Educação, Ciência e Tecnologia de Brasília; Instituto Federal de Educação, Ciência e Tecnologia de Goiás; Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso; Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais; Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco; Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina; Instituto Federal de Educação, Ciência e Tecnologia de São Paulo; Instituto Federal de Educação, Ciência e Tecnologia do Acre; Instituto Federal de Educação, Ciência e Tecnologia do Amapá; Instituto Federal de Educação, Ciência e Tecnologia do Amazonas; Instituto Federal de Educação, Ciência e Tecnologia do Ceará; Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo; Instituto Federal de Educação, Ciência e Tecnologia do Maranhão; Instituto Federal de Educação, Ciência e Tecnologia do Pará; Instituto Federal de Educação, Ciência e Tecnologia do Paraná; Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro; Instituto Federal de Educação, Ciência e Tecnologia do Sertão Pernambucano; Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais; Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais; Instituto Federal de Educação, Ciência e Tecnologia do Tocantins; Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro; Instituto Federal de Educação, Ciência e Tecnologia Goiano; Instituto Nacional da Propriedade Industrial; Instituto Nacional de Colonização e Reforma Agrária; Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira; Instituto Nacional de Metrologia, Qualidade e Tecnologia; Instituto Nacional de Tecnologia da Informação; Instituto Nacional do Seguro Social; Ministério da Ciência, Tecnologia e Inovações; Ministério da Economia; Ministério Público da União; Petrobras Transporte S.a. - MME; Petróleo Brasileiro S.a.; Procuradoria Geral da Fazenda Nacional; Secretaria de Controle Interno do Ministério das Relações Exteriores; Secretaria de Empreendedorismo e Inovação; Secretaria Especial da Receita Federal do Brasil; Secretaria -Executiva da Casa Civil da Presidência da República; Secretaria -Executiva do Ministério da Agricultura, Pecuária e Abastecimento; Secretaria -Executiva do Ministério da Cidadania; Secretaria -Executiva do Ministério da Ciência, Tecnologia e Inovações; Secretaria -Executiva do Ministério da Economia; Secretaria -Executiva do Ministério da Educação; Secretaria -Executiva do Ministério da Infraestrutura; Secretaria -Executiva do Ministério da Justiça e Segurança Pública; Secretaria -Executiva do Ministério da Mulher, da Família e dos Direitos Humanos; Secretaria -Executiva do Ministério da Saúde; Secretaria -Executiva do Ministério da Transparência, Fiscalização e Controladoria-Geral da União; Secretaria -Executiva do Ministério das Comunicações; Secretaria -Executiva do Ministério de Minas e Energia; Secretaria -Executiva do Ministério do Desenvolvimento Regional; Secretaria -Executiva do Ministério do Meio Ambiente; Secretaria -Executiva do Ministério do Turismo; Secretaria -Geral do Ministério da Defesa; Superintendência da Zona Franca de Manaus; Superintendência de Desenvolvimento da Amazônia; Superintendência de Desenvolvimento do Nordeste; Superintendência de Seguros Privados; Superintendência do Desenvolvimento do Centro -Oeste; Superintendência Nacional de Previdência Complementar; Superior Tribunal Militar; Telecomunicações Brasileiras S.a.; Tribunal de Justiça do Distrito Federal e dos Territórios; Tribunal Regional do Trabalho da 1ª Região/rj; Tribunal Regional do Trabalho da 10ª Região/df e TO; Tribunal Regional do Trabalho da 11ª Região/am e RR; Tribunal Regional do Trabalho da 12ª Região/sc; Tribunal Regional do Trabalho da 13ª Região/pb; Tribunal Regional do Trabalho da 14ª Região/ac e RO; Tribunal Regional do Trabalho da 15ª Região -

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360190.

Campinas/SP; Tribunal Regional do Trabalho da 16ª Região/ma; Tribunal Regional do Trabalho da 17ª Região/es; Tribunal Regional do Trabalho da 18ª Região/go; Tribunal Regional do Trabalho da 19ª Região/al; Tribunal Regional do Trabalho da 2ª Região/sp; Tribunal Regional do Trabalho da 20ª Região/se; Tribunal Regional do Trabalho da 21ª Região/rn; Tribunal Regional do Trabalho da 22ª Região/pi; Tribunal Regional do Trabalho da 23ª Região/mt; Tribunal Regional do Trabalho da 24ª Região/ms; Tribunal Regional do Trabalho da 3ª Região/mg; Tribunal Regional do Trabalho da 4ª Região/rs; Tribunal Regional do Trabalho da 5ª Região/ba; Tribunal Regional do Trabalho da 6ª Região/pe; Tribunal Regional do Trabalho da 7ª Região/ce; Tribunal Regional do Trabalho da 8ª Região/pa e AP; Tribunal Regional do Trabalho da 9ª Região; Tribunal Regional Eleitoral da Bahia; Tribunal Regional Eleitoral da Paraíba; Tribunal Regional Eleitoral de Alagoas; Tribunal Regional Eleitoral de Goiás; Tribunal Regional Eleitoral de Mato Grosso do Sul; Tribunal Regional Eleitoral de Minas Gerais; Tribunal Regional Eleitoral de Pernambuco; Tribunal Regional Eleitoral de Rondônia; Tribunal Regional Eleitoral de Roraima; Tribunal Regional Eleitoral de Santa Catarina; Tribunal Regional Eleitoral de São Paulo; Tribunal Regional Eleitoral do Acre; Tribunal Regional Eleitoral do Amapá; Tribunal Regional Eleitoral do Amazonas; Tribunal Regional Eleitoral do Ceará; Tribunal Regional Eleitoral do Distrito Federal; Tribunal Regional Eleitoral do Espírito Santo; Tribunal Regional Eleitoral do Maranhão; Tribunal Regional Eleitoral do Mato Grosso; Tribunal Regional Eleitoral do Pará; Tribunal Regional Eleitoral do Paraná; Tribunal Regional Eleitoral do Piauí; Tribunal Regional Eleitoral do Rio de Janeiro; Tribunal Regional Eleitoral do Rio Grande do Norte; Tribunal Regional Eleitoral do Rio Grande do Sul; Tribunal Regional Eleitoral do Sergipe; Tribunal Regional Eleitoral do Tocantins; Tribunal Regional Federal da 1ª Região; Tribunal Regional Federal da 2ª Região; Tribunal Regional Federal da 3ª Região; Tribunal Regional Federal da 4ª Região; Tribunal Regional Federal da 5ª Região; Tribunal Superior do Trabalho; Tribunal Superior Eleitoral; Universidade da Integração Internacional da Lusofonia Afro -Brasileira; Universidade Federal da Fronteira Sul; Universidade Federal da Integração Latino -Americana; Universidade Federal da Paraíba; Universidade Federal de Alagoas; Universidade Federal de Alfenas; Universidade Federal de Campina Grande; Universidade Federal de Goiás; Universidade Federal de Itajubá; Universidade Federal de Juiz de Fora; Universidade Federal de Lavras; Universidade Federal de Minas Gerais; Universidade Federal de Pelotas; Universidade Federal de Pernambuco; Universidade Federal de Roraima; Universidade Federal de Santa Catarina; Universidade Federal de Santa Maria; Universidade Federal de São Paulo; Universidade Federal do Cariri; Universidade Federal do Ceará; Universidade Federal do Espírito Santo; Universidade Federal do Estado do Rio de Janeiro; Universidade Federal do Oeste da Bahia; Universidade Federal do Oeste do Pará; Universidade Federal do Pará; Universidade Federal do Paraná; Universidade Federal do Recôncavo da Bahia; Universidade Federal do Rio de Janeiro; Universidade Federal do Rio Grande do Norte; Universidade Federal do Rio Grande do Sul; Universidade Federal do Sul da Bahia; Universidade Federal do Sul e Sudeste do Pará; Universidade Federal do Triângulo Mineiro; Universidade Federal dos Vales do Jequitinhonha e Mucuri; Universidade Federal Fluminense; Universidade Federal Rural da Amazônia; Universidade Federal Rural de Pernambuco; Universidade Federal Rural do Rio de Janeiro; Universidade Federal Rural do Semiárido; Universidade Tecnológica Federal do Paraná; Valec Engenharia Construções e Ferrovias S/A.

5. Relator: Ministro Aroldo Cedraz.

6. Representante do Ministério Público: não atuou.

7. Unidade Técnica: Secretaria de Fiscalização de Tecnologia da Informação (Sefti).

8. Representação legal: Álcio Antônio Lopes Guimarães (29.946/OAB-RS), Vitto Giancristoforo dos Santos (58566/OAB-RS) e outros, representando Hospital Nossa Senhora da Conceição S.A.; Leonor Chaves Maia de Sousa, Arnaldo de Morais Moreira Fernandes Vieira, Danielle Gonçalves e Silva, Ari Barbosa Ferreira e Ernesto Lima Cruz, representando Banco do Nordeste do Brasil S.A.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360190.

9. Acórdão:

VISTOS, relatados e discutidos estes autos de Levantamento de Auditoria, com o objetivo de avaliar o estágio atual e perspectivas de utilização de Inteligência Artificial (IA) na Administração Pública Federal (APF), identificar os riscos associados, conhecer os impactos para o controle e avaliar a proposta para uma Estratégia Brasileira de Inteligência Artificial (EBIA),

ACORDAM os ministros do Tribunal de Contas da União, reunidos em sessão do Plenário, ante as razões expostas pelo relator, em:

9.1. encaminhar cópia desta deliberação, bem como do relatório, do voto, do relatório da unidade técnica e dos seus respectivos Apêndices ao Ministério da Ciência, Tecnologia e Inovações (MCTI), à Comissão de Ciência e Tecnologia, Comunicação e Informática da Câmara dos Deputados, à Comissão de Ciência, Tecnologia, Inovação, Comunicação, Inovação e Informática do Senado Federal, bem como aos demais interessados (constam no Apêndice IX do relatório à peça 628);

9.2. levantar o sigilo dos autos, por conter informações relevantes à Administração Pública;

9.3. autorizar a Secretaria de Fiscalização de Tecnologia da Informação (Sefti) a divulgar as informações consolidadas constantes deste levantamento, preferencialmente por intermédio de fichas síntese, sumário executivo e infográfico;

9.4. determinar a conversão do presente processo em Acompanhamento, com os seguintes objetivos:

9.4.1. avaliar periodicamente o nível de maturidade dos órgãos da Administração Pública Federal no uso de inteligência artificial, abrangendo tanto sistemas especialistas baseados em regras como soluções de aprendizagem de máquina;

9.4.2. desenvolver, validar e aplicar referencial teórico e metodológico próprio do Tribunal de Contas da União para auditoria de aplicações e algoritmos de inteligência artificial, com abrangência idêntica à citada no item anterior;

9.4.3. avaliar a implementação da Estratégia Brasileira de Inteligência Artificial e indicar eventuais medidas necessárias para assegurar a efetividade dessa política pública;

9.5. orientar à Sefti que elabore guia, ou instrumento congênere, com a definição de diretrizes, parâmetros e eventuais riscos, a fim de auxiliar líderes e gestores públicos no processo de implementação ou contratação de serviços que envolvam o uso de inteligência artificial.

9.6. restituir os presentes autos à Sefti para continuidade da próxima etapa do trabalho.

10. Ata n° 19/2022 – Plenário.
11. Data da Sessão: 25/5/2022 – Ordinária.
12. Código eletrônico para localização na página do TCU na Internet: AC-1139-19/22-P.
13. Especificação do quórum:
13.1. Ministros presentes: Ana Arraes (Presidente), Walton Alencar Rodrigues, Benjamin Zymler, Augusto Nardes, Aroldo Cedraz (Relator), Jorge Oliveira e Antonio Anastasia.
13.2. Ministro-Substituto convocado: Augusto Sherman Cavalcanti.
13.3. Ministro-Substituto presente: André Luís de Carvalho.

(Assinado Eletronicamente)
ANA ARRAES
Presidente

(Assinado Eletronicamente)
AROLDO CEDRAZ
Relator

Fui presente:

(Assinado Eletronicamente)
CRISTINA MACHADO DA COSTA E SILVA
Procuradora-Geral

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69360190.

RELATÓRIO DE LEVANTAMENTO SOBRE A UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL NA ADMINISTRAÇÃO PÚBLICA FEDERAL

TC 006.662/2021-8 **Fiscalis: 22/2021**
Relator: Aroldo Cedraz

DA FISCALIZAÇÃO

Modalidade: Levantamento

Ato originário: Despacho do Ministro Aroldo Cedraz, de 10/2/2021 (TC 036.797/2020-0, peça 5)

Objetivo: conhecer as principais tecnologias de Inteligência Artificial; avaliar o estágio atual e perspectivas de sua utilização na Administração Pública Federal (APF); identificar os principais riscos de decisão, contratação, implementação e gestão; internalizar competências necessárias ao corpo técnico do TCU para acompanhamento das ações governamentais que utilizam tais tecnologias; e avaliar o estágio atual e conhecer o estágio de desenvolvimento e o conteúdo da proposta para uma Estratégia Brasileira de Inteligência Artificial (EBIA), identificando os principais riscos e oportunidades associados à sua implementação.

Ato de designação: Portaria de Fiscalização 81, de 26 de fevereiro de 2021, e outras (peças 2, 7, 619, 624 e 625)

Período da fiscalização: 1/3/2021 a 10/9/2021

Composição da equipe: Diego Oliveiras Farias, matr. 10642-9
Eric Hans Messias da Silva, matr. 9652-0
Erick Muzart Fonseca dos Santos, matr. 7666-0
Monique Louise de Barros Monteiro, matr. 8677-0
Tibério César Jocundo Loureiro, matr. 6520-0 (coordenador)

DOS ÓRGÃOS/ENTIDADES FISCALIZADOS

Nome	Vinculação
Ministério da Ciência e Tecnologia	Presidência da República
Órgãos da Administração Pública Federal	

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Sumário

1 Introdução 4
1.1 Deliberação que originou a fiscalização 4
1.2 Identificação do objeto 4
1.3 Objetivo, Escopo e Questões de Auditoria 4
1.4 Metodologia utilizada 4
1.5 Limitações 5
1.6 Benefícios Estimados da Fiscalização 5
2 Visão Geral do Objeto 5
2.1 O que é Inteligência Artificial 6
2.2 IA Forte e IA Fraca 6
2.3 Aprendizagem de Máquina 6
2.4 Tipos de Aprendizagem de Máquina 8
2.4.1 Aprendizagem Supervisionada 8
2.4.2 Aprendizagem não Supervisionada 8
2.4.3 Aprendizagem por reforço 8
2.5 Técnicas de aprendizagem de máquina 9
2.6 Principais domínios de aplicação da IA 10
2.7 Ciclo de desenvolvimento de soluções de IA 11
2.8 O que considerar no desenvolvimento e uso de IA 12
3 Utilização de Inteligência Artificial no governo 14
3.1 Questionário eletrônico sobre utilização de IA na Administração Pública Federal 18
4 Estratégia Brasileira de Inteligência Artificial (EBIA) 28
4.1 Conjuntura internacional 28
4.2 Estratégia Brasileira de Inteligência Artificial 30
4.2.1 Descrição geral 30
4.2.2 Metodologia utilizada para avaliação da EBIA 31
4.2.3 Achados de auditoria 34
5 Inteligência Artificial no Controle 42
5.1 Utilização de IA na atividade de controle 42
5.2 Soluções de IA para a atividade de controle 43
5.2.1 GAO - An Accountability Framework for Federal Agencies and Other Entities 44
5.2.2 ISSAI - Auditing machine learning algorithms 45
5.2.3 ICO - Guidance on the AI auditing framework 45

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

5.2.4 Comissão Europeia - Assessment List for Trustworthy Artificial Intelligence.....46

5.2.5 Transparência Brasil - Estrutura de avaliação de riscos a direitos e de transparência.....47

5.2.6 LIAA-3R - Diretrizes de auditabilidade e conformidade no desenvolvimento e testes de soluções de IA.....48

6 Conclusão.....49

7 Proposta de Encaminhamento.....51

APÊNDICE I – Principais algoritmos de Aprendizagem de Máquina.....52

APÊNDICE II – Redes Neurais Artificiais.....54

APÊNDICE III – Principais domínios de aplicação de IA.....60

APÊNDICE IV – Ciclo de desenvolvimento de soluções de IA.....69

APÊNDICE V – Considerações do desenvolvimento de soluções de IA.....72

APÊNDICE VI – Privacidade de Dados.....76

APÊNDICE VII – Algumas soluções para interpretabilidade de modelos de IA.....77

APÊNDICE VIII – Detalhamento de Soluções de IA desenvolvidas na APF.....80

APÊNDICE IX – Organizações participantes da pesquisa.....101

APÊNDICE X – Consolidação das informações referentes às 36 soluções de IA que foram detalhadas nas respostas ao questionário eletrônico.....107

APÊNDICE XI – Exemplos de aplicações de IA no Controle.....111

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

1 Introdução

1.1 Deliberação que originou a fiscalização

1. O Levantamento em tela foi autorizado por meio de despacho do Ministro Relator Aroldo Cedraz, de 10/2/2021, decorrente de proposta de fiscalização apresentada pela Sefti, no âmbito do TC 036.797/2020-0 (Administrativo, peça 5).

1.2 Identificação do objeto

2. O objeto do presente levantamento são as tecnologias de Inteligência Artificial (IA) nas organizações da Administração Pública Federal (APF) em suas diversas formas de utilização.

1.3 Objetivo, Escopo e Questões de Auditoria

3. O objetivo deste trabalho é fazer um levantamento sobre o estágio atual da utilização de tecnologias de IA por diversas organizações que compõem a Administração Pública, sob vários aspectos.

4. O primeiro deles diz respeito a conhecer as principais tecnologias utilizadas no desenvolvimento de soluções de IA, e avaliar o estágio atual e as perspectivas de sua utilização nas organizações públicas, além de identificar os principais riscos relacionados ao tema.

5. O segundo objetivo refere-se ao conhecimento do estágio atual de desenvolvimento e o conteúdo da proposta para uma Estratégia Brasileira de Inteligência Artificial, e identificar os principais riscos e oportunidades associados à sua implementação.

6. O escopo do trabalho abrangeu a utilização de IA nas organizações dos três poderes da República, tendo sido analisados somente aspectos em relação à IA enquanto tecnologias de aprendizado de máquina, conforme explicado mais detalhadamente adiante, sendo que não foram consideradas as soluções desenvolvidas por organizações da área privada.

7. Por fim, o levantamento buscou verificar de que forma o controle das soluções de IA desenvolvidas no âmbito da APF pode ser exercido, e como as tecnologias de IA podem auxiliar a atividade de controle.

8. Nesse sentido, elaboraram-se as seguintes questões de auditoria:

a. Quais as principais tecnologias de IA e seus benefícios para aplicação pela APF?

b. Qual o cenário atual de utilização da IA na APF?

c. Quais dificultadores e riscos devem ser considerados pelos gestores na implementação de soluções de IA na APF?

d. Qual o estágio atual da implementação da Estratégia Brasileira de IA e os riscos e oportunidades referentes à sua implementação?

e. Quais metodologias podem ser aplicadas pelas equipes de auditoria para fiscalizar soluções de IA na APF?

f. Como a IA pode ser utilizada para apoiar a atividade de controle?

1.4 Metodologia utilizada

9. Primeiramente, registra-se que os trabalhos foram realizados em conformidade com as Normas de Auditoria do Tribunal de Contas da União (NAT), definidas na Portaria-TCU 168/2011, e com o Roteiro de Levantamento, estabelecido na Portaria-Segecex 5/2021.

10. A presente auditoria foi realizada na modalidade de teletrabalho, na forma que foi direcionada pelas boas práticas desenhadas pelo TCU para auditorias virtuais utilizando as ferramentas já disponíveis (https://tcucloud.sharepoint.com/sites/praticasremotasauditoria).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

11. Para refinar a matriz de planejamento que foi elaborada pela equipe de auditoria, foram realizadas entrevistas com especialistas, em que foram colhidas sugestões para as questões de auditoria, além de identificar organizações relevantes no uso de IA no país.

12. A fiscalização baseou-se em entrevistas com gestores para conhecer os principais casos de uso e funcionalidades da tecnologia no setor público. Além disso, verificou-se a existência de trabalhos realizados por vários organismos internacionais que abrangem diversos dos objetivos presentes neste levantamento, servindo assim de importante fonte para o trabalho. Também se utilizou a literatura referente ao tema, bem como fontes acadêmicas, conforme elencadas na seção "Referências".

13. Além disso, a equipe da fiscalização elaborou questionário eletrônico que foi encaminhado a 293 organizações abrangidas no presente escopo, com o objetivo de colher detalhes sobre como a IA é utilizada para resolver problemas de negócio e agregar valor nos serviços prestados aos cidadãos.

14. A equipe também entrevistou especialistas da área, tanto do meio acadêmico como da inciativa privada, com o objetivo de subsidiar a elaboração das informações constantes neste Levantamento.

1.5 Limitações

15. Como limitações do trabalho, citam-se o grande número de instituições abrangidas no levantamento e a dificuldade de obtenção de resposta ao questionário eletrônico por todos os envolvidos. Além disso, referente à avaliação da Estratégia Brasileira de Inteligência Artificial (EBIA), a recriação do Ministério das Comunicações em 2020, a qual motivou a remoção de vários dos servidores do MCTI que estavam encarregados da elaboração da referida estratégia para o novo ministério, trouxe dificuldades na obtenção de informações referentes ao tema.

1.6 Benefícios Estimados da Fiscalização

16. Consideram-se os benefícios da presente fiscalização:

a. Fornecer subsídios para aprimorar o conhecimento dos gestores públicos federais em relação aos benefícios e critérios para a utilização de tecnologias de IA na disponibilização de serviços à sociedade e melhoria dos processos internos nas organizações públicas;

b. Estabelecer uma linha de base da situação atual de utilização de tecnologias de IA na Administração Pública Federal, de forma a subsidiar trabalhos futuros que busquem aferir a evolução do tema;

c. Dar subsídios para o aprimoramento da Estratégia Brasileira de Inteligência Artificial (EBIA), contribuindo para que alcance seus objetivos, de maneira legítima, econômica, eficaz e eficiente e, assim, possa atender à sociedade;

d. Aquisição de conhecimento pelo TCU nas metodologias e *frameworks* aplicáveis à fiscalização de soluções de IA.

2 Visão Geral do Objeto

17. Esta seção tem o intuito de conceituar o que é IA e apresentar suas principais técnicas, além de citar as principais áreas de aplicação e explicar as principais questões relativas ao desenvolvimento de soluções de IA.

18. Tendo em vista que existe farto material que discorre sobre o assunto em fontes públicas e na internet, não será feito um aprofundamento maior nas subseções a seguir, ficando o detalhamento sobre o tema a cargo dos apêndices deste relatório.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

2.1 O que é Inteligência Artificial[1]

19. Ao longo do tempo, muitas definições foram dadas para o termo Inteligência Artificial, e a associação do mesmo a outros termos como aprendizado de máquina (*machine learning*) e aprendizado profundo (*deep learning*) tem gerado dificuldades para uma melhor compreensão sobre o tema.

20. O aspecto artificial da IA é bastante simples: refere-se a qualquer coisa não natural e, neste caso, feito pelo homem. Também pode ser representado por meio do uso de termos como máquinas, computadores ou sistemas. Já inteligência é um conceito muito mais amplamente contestado, explicando por que ainda não há consenso sobre como definir IA (Miaihle and Hodes, 2017).

21. Inteligência Artificial pode ser definida como o uso de tecnologia digital para criar sistemas capazes de realizar tarefas comumente pensadas para exigir inteligência.

22. Nesse contexto, cita-se a definição da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), que trata IA como um sistema baseado em máquina que pode, para um determinado conjunto de objetivos definidos pelo homem, fazer previsões, recomendações ou decisões que influenciam ambientes reais ou virtuais.

23. Dessa forma, pode-se falar que a IA atual envolve principalmente máquinas usando estatísticas para encontrar padrões em grandes quantidades de dados e realizar tarefas repetitivas com dados sem a necessidade de constante orientação humana. Assim, a IA não consiste numa solução tecnológica aplicável para todos os casos, uma vez que, em geral, não apresenta boa performance sem que haja uma quantidade de dados significativa, relevante e com boa qualidade.

2.2 IA Forte e IA Fraca

24. O termo IA Forte refere-se à ideia de que a inteligência humana geral - que abrange diferentes domínios e capacidades - pode ser igualada ou até superada por máquinas. A partir desse conceito, várias suposições são levantadas no sentido de que as crenças desse sistema possam não estar alinhadas às humanas. Tal representação de IA na mídia de massa e na cultura popular capturaram a imaginação do público e podem contribuir para a desconfiança da tecnologia impulsionada por IA, embora haja opiniões mistas entre pesquisadores de IA quanto à probabilidade de tal cenário.

25. Já a IA fraca está menos preocupada com a criação de uma superinteligência unificada, mas aceita e utiliza a noção de que humanos e computadores têm diferentes pontos fortes e competências. A IA tira proveito do fato de que os computadores são excelentes no desempenho quando se trata de processamento rápido e consistente de grandes quantidades de dados, bem como na execução de tarefas com base em dados lógicos e regras explícitas, enquanto os humanos ainda são mais eficientes em lidar com situações ambíguas ou aquelas que requerem intuição, criatividade, emoção, julgamento e empatia. Atualmente, na prática, praticamente toda a IA utilizada é a IA fraca.

26. O presente levantamento tem seu escopo limitado às aplicações da IA Fraca, e mais especificamente no campo das tecnologias associadas à Aprendizagem de Máquina, apresentadas a seguir.

2.3 Aprendizagem de Máquina[2]

27. Nos primeiros dias da Inteligência Artificial, rapidamente foram resolvidos problemas que são intelectualmente difíceis para humanos, mas relativamente simples para computadores – problemas

[1] O resumo apresentado nesta seção é baseada no documento "*Hello, World: Artificial intelligence and its use in the public sector*", publicado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE) (2019) e disponível em https://www.oecd.org/gov/innovative-government/working-paper-hello-world-artificial-intelligence-and-its-use-in-the-public-sector.htm?utm_content=buffer2b1a0&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer

[2] O resumo apresentado nesta seção é baseado em tradução livre do livro (Goodfellow, Bengio, & Courville, *Deep Learning*, 2016).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

que podem ser descritos como uma lista de regras formais, matemáticas. O desafio real para a IA provou-se ser resolver as tarefas que são fáceis para as pessoas, porém difíceis para que elas as descrevam formalmente – problemas que resolvemos de forma intuitiva e automática, como reconhecer palavras pronunciadas ou rostos em imagens.

28. Muitos dentre os primeiros sucessos da Inteligência Artificial ocorreram em ambientes relativamente formais e não requeriam que computadores tivessem muito conhecimento sobre o ambiente. Um exemplo é o jogo de xadrez – ainda na década de 1990, máquinas já eram capazes de vencer humanos. Porém, derivar uma estratégia bem-sucedida para jogar xadrez requer tão somente que se descreva completamente uma grande lista de regras formais, o que pode ser codificado previamente por um programador. Nesse caso, por exemplo, o desafio não está em ser capaz de descrever o ambiente – composto basicamente por um conjunto limitado de peças que podem mover-se apenas de formas rigidamente circunscritas.

29. As tarefas diárias de uma pessoa requerem um conhecimento imenso sobre o mundo ao seu redor. Computadores precisam capturar esse mesmo conhecimento para se comportarem de forma inteligente. Um dos desafios na IA é como representar tal conhecimento formalmente para um computador.

30. Vários projetos de inteligência artificial buscaram codificar o conhecimento sobre o mundo em linguagens formais. Um computador pode raciocinar automaticamente em cima de linguagens formais utilizando regras lógicas de inferência, o que é conhecido como abordagem de base de conhecimento. Nenhum desses projetos conduziu a um maior sucesso.

31. As dificuldades enfrentadas por sistemas ao confiarem em conhecimento codificado sugerem que sistemas de IA precisam da habilidade de adquirirem seu próprio conhecimento, extraindo padrões a partir de dados brutos. Tal capacidade é conhecida como aprendizagem de máquina, ou *machine learning*. A introdução da aprendizagem de máquina permitiu que computadores tratassem problemas envolvendo conhecimento do mundo real e tomassem decisões que pareciam subjetivas. A figura abaixo ilustra a diferença entre a programação clássica e a aprendizagem de máquina.

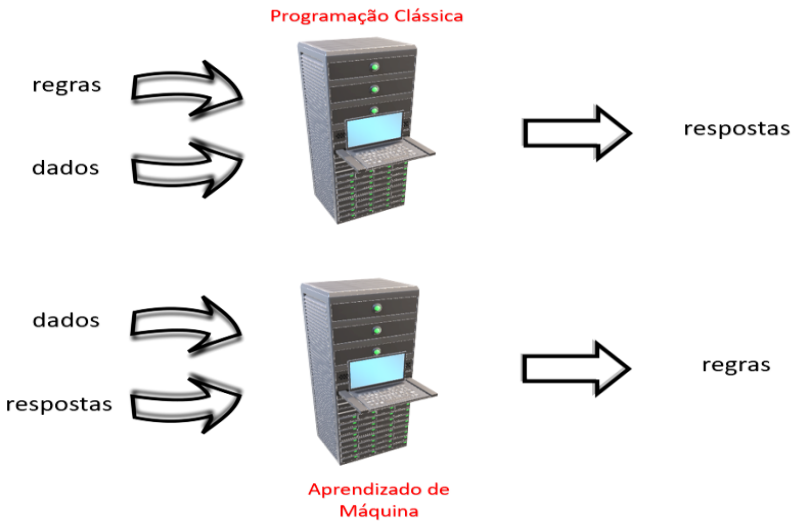


Figura 1 - Diferença entre programação clássica e aprendizado de máquina

32. Como visto acima, o conceito de aprendizado de máquina é complexo e engloba diversos aspectos técnicos. Para fins desse trabalho, será utilizada a definição formal da OCDE:

> O aprendizado de máquina é uma abordagem em que as máquinas aprendem a fazer previsões em novas situações baseadas em dados históricos. O aprendizado de máquina consiste em um conjunto de técnicas para permitir que as máquinas aprendam de maneira automatizada, sem instruções

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

explícitas de um ser humano, por confiar em padrões e inferências. Aprendizado de máquina se aproxima frequentemente de ensinar as máquinas a chegarem a um resultado, mostrando-lhes muitos exemplos de resultados - chamados de "treinamento". Outra abordagem é os humanos definirem um conjunto de amplas regras e geralmente permitem que a máquina aprenda por conta própria por tentativa e erro.

2.4 Tipos de Aprendizagem de Máquina

2.4.1 Aprendizagem Supervisionada

33. Tendo em vista o paradigma de aprendizagem a partir de dados, considera-se aprendizagem supervisionada quando a máquina, para que seja capaz de fazer predições a partir de informações de entrada, precisa ter sido previamente treinada com exemplos nos quais, além das informações de entrada, as informações de saída esperadas também são especificadas, em geral, por um supervisor humano. Trata-se, por exemplo, do aprendizado a partir de dados históricos. Assim, quanto maior a quantidade de dados ou exemplos de treinamento, melhor o desempenho preditivo esperado do algoritmo.

34. Ao mesmo tempo, o algoritmo ou modelo resultante da etapa de treinamento deverá ter também capacidade de generalização, ou seja, deverá ser capaz de fornecer resultados corretos, próximos do esperado no mundo real, para dados que não tenham sido vistos anteriormente. Esses dados poderão, por exemplo, ser parecidos com os dados de treinamento, porém não serão idênticos. Aqui, o termo "parecidos" refere-se, tecnicamente, a pertencerem a distribuições de probabilidade semelhantes.

2.4.2 Aprendizagem não Supervisionada

35. Em situações em que há uma ampla quantidade de dados (ex.: informações de perfil de clientes de uma loja virtual), porém não há clareza sobre um valor-alvo a ser previsto, ou mesmo quando não há histórico de dados suficiente para que se possa definir o valor-alvo correto, ainda assim é possível extrair algum grau de informação útil a partir dos dados. É neste contexto que surge a aprendizagem não supervisionada, voltada para a identificação de padrões nos dados. Exemplos de padrões incluem agrupamentos (ex.: grupos de usuários semelhantes entre si, com algumas características em comum), detecção de anomalias (ex.: *outliers*) e outros tipos de padrão que podem ser úteis ao menos em uma etapa de análise exploratória dos dados.

36. Em outras palavras, na aprendizagem de máquina não supervisionada não há a necessidade de uma supervisão humana para dizer à máquina o que especificamente procurar.

2.4.3 Aprendizagem por reforço

37. Entre todas as formas de aprendizado de máquina, o aprendizado por reforço é o que mais se aproxima da maneira com que os humanos e os outros animais aprendem. A abordagem neste tipo de aprendizado é de que não há uma resposta sobre como atingir determinado objetivo e, portanto, o sistema de IA terá que observar o estado do ambiente ao seu redor, agir sobre ele e colher a recompensa sobre sua ação para que possa ir aprendendo de forma interativa a descobrir qual a melhor maneira de maximizar a soma de suas recompensas no longo prazo e, com isso, atingir seu objetivo (Sutton & Barto, 2018).

38. A interação do sistema de IA com o ambiente faz com que aprenda a mapear estados observados em ações a serem executadas de acordo com sua probabilidade. Esse mapeamento é a política (*policy*) de comportamento do sistema de IA, uma regra de tomada de decisão.

39. Em geral, cada uma das três categorias de aprendizagem de máquina citadas tem sua aplicação voltada para a resolução de determinados tipos de problema. A figura abaixo ilustra as principais aplicações:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

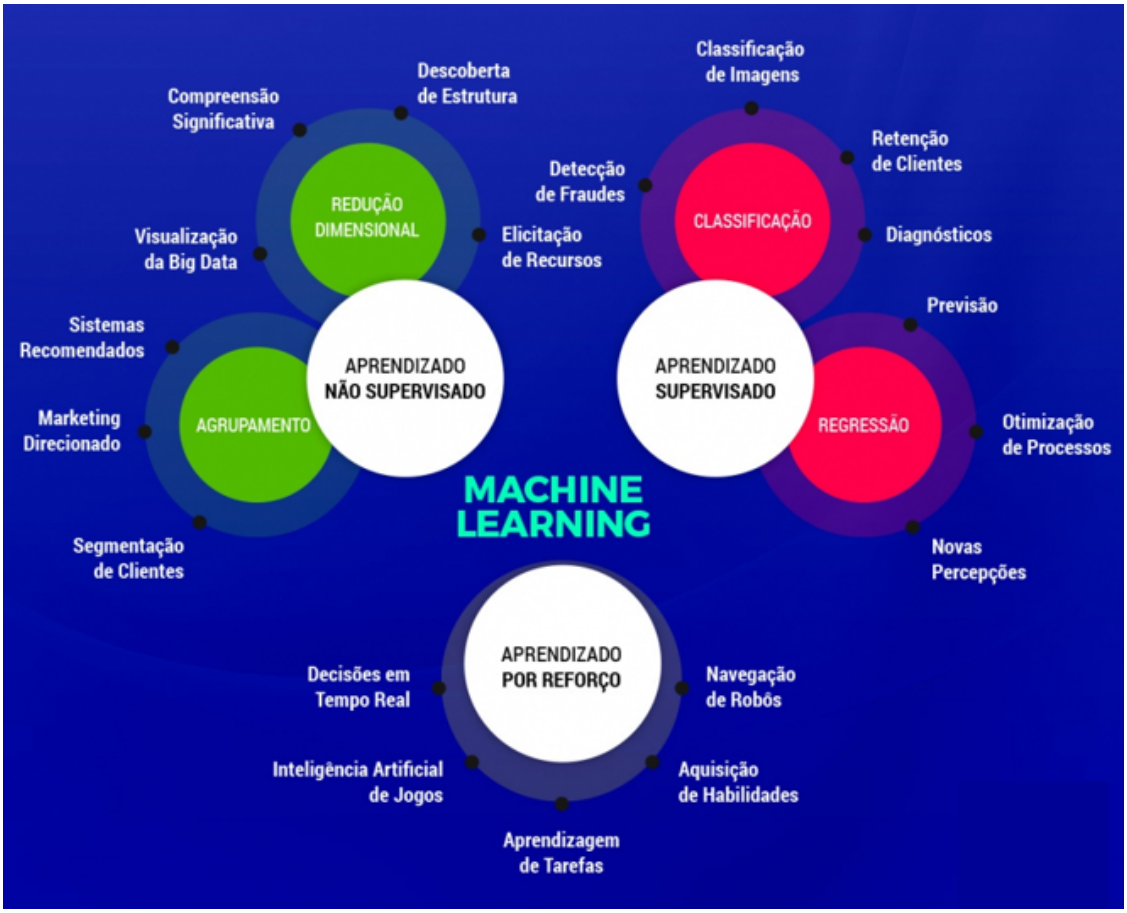


Figura 2 – Principais aplicações das tecnologias de aprendizagem de máquina[1]

2.5 Técnicas de aprendizagem de máquina

40. Não há uma tecnologia isolada que represente a IA. A exemplo, dentro de cada um dos tipos de aprendizagem de máquina identificados na secção anterior (supervisionada, não supervisionada e por reforço) existem diferenciados algoritmos e tecnologias que podem ser aplicadas, a depender dos objetivos a serem alcançados.

41. Nesse sentido, pode-se agrupar as técnicas de aprendizagem de máquina nos seguintes grupos:

Tabela 1 – Principais técnicas de Aprendizagem de Máquina[2]

Técnica	Descrição	Exemplos de aplicação
Classificação	Aprende as características de uma dada categoria, permitindo que o modelo de IA classifique dados desconhecidos em categorias pré-determinadas	• decidir se uma remessa de mercadorias passa por inspeção de fronteira; • decidir se um e-mail é *spam* ou não

[1] Fonte: https://www.programaria.org/quais-sao-os-tipos-de-aplicacoes-de-inteligencia-artificial-mais-comuns/

[2] Fonte: "*A guide to using artificial intelligence in the public sector" – Office for Artificial Intelligence – Goverment Digital Service – UK*. Disponível em: https://www.gov.uk/government/publications/a-guide-to-using-artificial-intelligence-in-the-public-sector

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Técnica	Descrição	Exemplos de aplicação
Regressão	Prevê um valor de ponto de dados desconhecido	• prever o valor de mercado de uma casa a partir de informações como tamanho, localização ou idade, prevendo as concentrações de poluentes atmosféricos nas cidades
Agrupamento	Identifica grupos de dados similares em um arquivo	• agrupar clientes de varejo para encontrar subgrupos com hábitos de consumo específicos; • agrupar dados de medidores inteligentes para identificar grupos de aparelhos elétricos e gerar contas de eletricidade discriminadas
Redução de dimensionalidade	Limita os dados às variáveis mais relevantes para tornar os modelos mais precisos ou possibilitar a visualização dos dados	• usado por cientistas de dados ao avaliar e desenvolver outros tipos de algoritmos de aprendizado de máquina
Ranqueamento	Treina um modelo de IA para classificar novos dados com base em listas vistas anteriormente	• retornar páginas por ordem de relevância quando um usuário pesquisa em um *website*

42. Para implementar cada uma das técnicas de aprendizagem de máquina citadas na Tabela 1, existem inúmeros algoritmos já desenvolvidos e implementados. O Apêndice I apresenta alguns desses algoritmos. Apesar de sua variabilidade, um fato comum a todos os algoritmos de aprendizagem de máquina é que eles necessitam essencialmente de dados para produzir seus resultados. Nesse contexto, tem-se que os dados podem estar representados de forma estruturada ou mesmo na forma de imagens ou sons. Assim, dependendo do formato dos dados, um ou outro algoritmo será mais aplicável.

43. Uma classe de algoritmos que vem ganhando cada vez mais destaque e aplicação são as Redes Neurais Artificiais e o Aprendizado Profundo (*Deep Learning*). As redes neurais são classificadas como métodos supervisionados do aprendizado de máquina e são basicamente um conjunto de funções matemáticas que simulam o funcionamento do cérebro humano, através da simulação de seus neurônios e de suas ligações. O Apêndice II traz um maior detalhamento sobre essas tecnologias.

2.6 Principais domínios de aplicação da IA

44. A IA está se tornando muito presente no dia a dia dos cidadãos e cada vez mais a tecnologia é incorporada a programas de computador, sítios na internet e dispositivos conectados e vestíveis.

45. Embora a tecnologia atual de IA ainda tenha limitações significativas, ela está evoluindo rapidamente e vem trazendo significantes benefícios em certos domínios de aplicação. Entre eles, destacam-se os modelos preditivos, o Processamento da Linguagem Natural (PLN) e a Visão Computacional, conforme ilustrado na figura abaixo:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

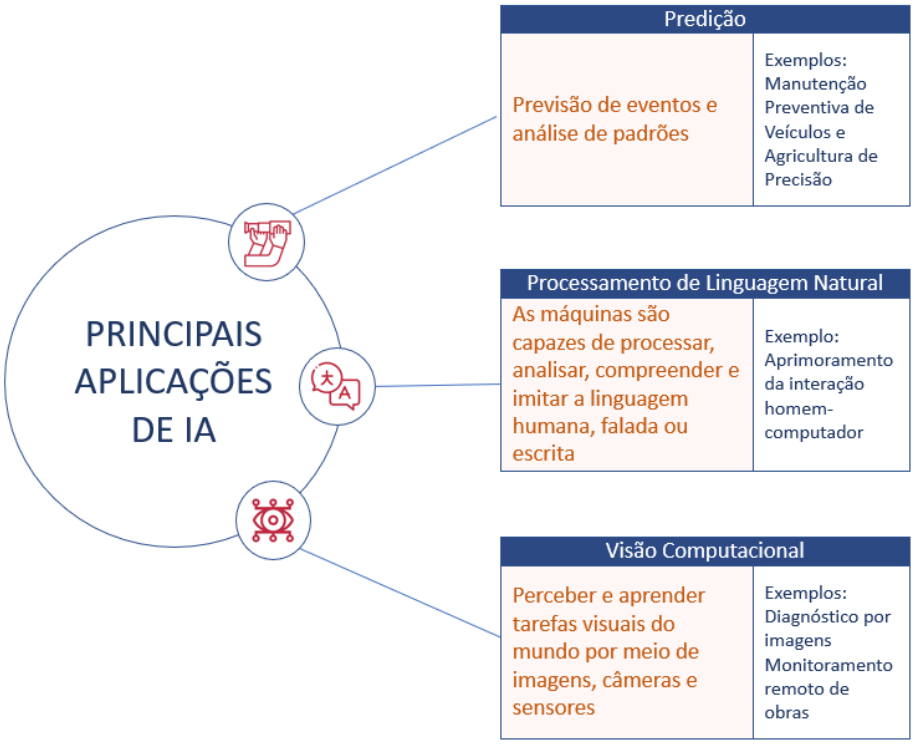


Figura 3 - Principais domínios de aplicação de IA

46. O Apêndice III apresenta um detalhamento sobre cada um desses domínios.

2.7 Ciclo de desenvolvimento de soluções de IA

47. O rápido crescimento das soluções que utilizam tecnologias de IA está tornando os sistemas cada vez maiores e complexos e os recentes avanços vêm fazendo com que estes sistemas sejam vistos como a base de uma futura geração de novos sistemas com inteligência. O fator mais importante é que se espera mais sucesso agora do que antes, pelo fato do grande interesse da comunidade científica. Com este advento, é importante que se tenham métodos e ferramentas que sirvam de auxílio nas etapas de desenvolvimento.

48. Dessa forma, já se desenvolveram várias propostas e modelos que endereçam o ciclo de vida de uma solução de IA. Dentre as metodologias existentes destacam-se as seguintes: KADS (*Knowledge Acquisition and Design Support*), MIKE (*Model-Based and Incremental Knowledge Engineering*), KOD (*Knowledge Oriented Design*), Vital, Desire, MKSM (*Methodology for Knowledge System Management*), e CYGMA (*Cycle de vie e Gestion des Métiers et des Applications*).

49. Porém, a maioria das soluções propostas para projetos que envolvem Aprendizagem de Máquina, apresenta, em resumo, as fases expostas na figura abaixo:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

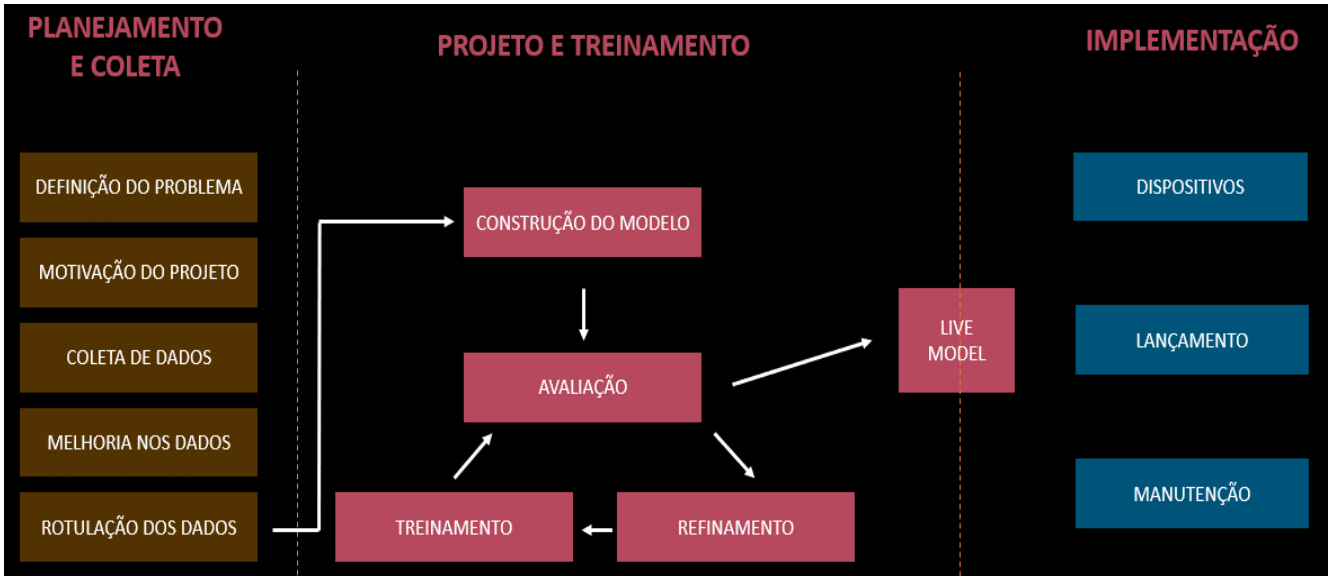


Figura 4 - Ciclo de desenvolvimento de um projeto de IA[1]

50. Em geral, todos os modelos de desenvolvimento abordam em sua fase inicial aspectos referentes à aquisição de dados a serem utilizados na elaboração dos modelos de aprendizagem de máquina. A construção do modelo é um processo no qual se busca desenvolver a solução com a maior acurácia e otimização do uso dos recursos computacionais. Por fim, após implementado o modelo num ambiente de produção, ele deve ser monitorado continuamente quanto ao seu desempenho. O Apêndice IV apresenta de forma mais detalhada as etapas do processo de desenvolvimento das soluções de IA.

2.8 O que considerar no desenvolvimento e uso de IA

51. Além de questões relativas à disponibilidade suficiente de dados, qualificação das equipes técnicas e aspectos operacionais de implementação, as soluções de IA demandam ainda aos gestores uma série de questões a serem consideradas na utilização dessa tecnologia e na busca do atingimento dos melhores resultados.

52. No documento "*A guide to using artificial intelligence in the public sector*", do Escritório para Inteligência Artificial do Serviço de Governo Digital do Reino Unido, são identificados aspectos relevantes que os gestores devem observar em relação ao desenvolvimento e à implementação de soluções de IA nas organizações públicas, destacando:

a) **qualidade de dados:** o sucesso do projeto de IA depende da qualidade dos dados utilizados no processo de aprendizagem de máquina;

b) **justiça** *(fairness)*: os modelos de IA devem ser treinados e testados utilizando conjuntos de dados relevantes, precisos e generalizáveis e os sistemas de IA devem ser implementados por usuários devidamente capacitados;

c) ***accountability:*** deve ser considerado quem é responsável por cada elemento da saída do modelo e como é atribuída a responsabilidade aos designers e implementadores dos sistemas de IA;

d) **privacidade:** a utilização dos dados deve estar em conformidade com a legislação e as políticas vigentes, a exemplo da Lei Geral de Proteção dos Dados (LGPD);

e) **interpretabilidade e transparência:** as partes interessadas devem poder entender como o modelo de IA chegou à sua decisão;

[1] Disponível em https://labelyourdata.com/articles/lifecycle-of-an-ai-project-stages-breakdown/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

f) custos: deve ser considerado quanto custará para construir, executar e manter uma infraestrutura de IA, treinar e educar a equipe e se o trabalho para instalar AI pode superar qualquer economia potencial.

53. Um outro exemplo de princípios referentes ao desenvolvimento e implementação de soluções que envolvem tecnologias de IA são elencados pelo *The Institute for Ethical AI and Machine Learning*[1]. Para cada um dos princípios, são estabelecidos compromissos a serem observados pelos responsáveis pela solução de IA, conforme abaixo:

Tabela 2 - Princípios para o desenvolvimento de soluções de IA (tradução livre)

	Princípio	Compromisso
1	**Participação humana**	avaliar o impacto de previsões incorretas e, quando razoável, projetar sistemas com processos de revisão humanos
2	**Avaliação de viés**	desenvolver continuamente processos que permitam compreender, documentar e monitorar preconceitos no desenvolvimento e na produção.
3	**Explicabilidade por justificativa**	desenvolver ferramentas e processos para melhorar continuamente a transparência e a explicabilidade dos sistemas de aprendizado de máquina onde for razoável.
4	**Operações reproduzíveis**	desenvolver a infraestrutura necessária para permitir um nível razoável de reprodutibilidade em todas as operações de sistemas de ML.
5	**Estratégia de deslocamento**	identificar e documentar informações relevantes para que os processos de mudança de negócios possam ser desenvolvidos para mitigar o impacto para os trabalhadores que estão sendo automatizados.
6	**Precisão prática**	desenvolver processos para garantir que as funções de precisão e métrica de custo estejam alinhadas aos aplicativos específicos do domínio.
7	**Confiança pela privacidade**	construir e comunicar processos que protegem e tratam os dados com as partes interessadas que podem interagir com o sistema direta e/ou indiretamente.
8	**Conscientização de risco de dados**	desenvolver e melhorar processos e infraestrutura razoáveis para garantir que os dados e a segurança do modelo sejam levados em consideração durante o desenvolvimento de sistemas de aprendizado de máquina.

54. Considerações sobre o uso de IA no Setor Público também são abordadas no documento "**Uso de Inteligência Artificial pelo Poder Público - Recomendações de Governança**"[2], elaborado pela organização Transparência Brasil no âmbito do Projeto Transparência Algorítmica, o qual traz, em resumo, as seguintes recomendações:

[1] Disponível em https://ethical.institute/principles.html#commitment-1

[2] Disponível em https://www.transparencia.org.br/projetos/transparencia-algoritmica

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

a) as bases de dados utilizadas para a construção dos modelos devem ser representativas e apropriadas para o contexto na qual serão utilizadas;

b) necessidade de supervisão humana como salvaguarda para a revisão de decisões automatizadas;

c) efetiva proteção dos dados pessoais do cidadão; e

d) transparência e explicabilidade dos sistemas.

55. Nesse contexto, o Apêndice V detalha os principais aspectos concernentes à implementação de soluções de IA que devem ser considerados pelas organizações.

3 Utilização de Inteligência Artificial no governo

56. O setor público é responsável por definir as prioridades, investimentos e regulamentações nacionais no que diz respeito à IA. Dessa forma, os governos devem buscar alavancar as tecnologias de IA para inovar e transformar o setor público, a fim de redefinir as maneiras pelas quais as políticas e serviços públicos são projetados e implementados. Essa inovação e transformação são críticas para os governos, pois eles enfrentam uma complexidade e demandas cada vez maiores de seus cidadãos e organizações públicas e privadas.

57. Assim, a IA pode ser integrada a todo o processo de formulação de políticas e desenho de serviços. À medida que a IA e o aprendizado de máquina evoluem, mais tarefas administrativas e orientadas a processos podem ser automatizadas, aumentando a eficiência do setor público e liberando os funcionários públicos para se concentrarem em um trabalho mais significativo. Com o auxílio da IA, os governos também podem entender melhor a realidade, tomar decisões dentro de suas organizações e antecipar as necessidades da população. Se bem executados, os processos automatizados podem ajudar o governo a tomar decisões mais justas e precisas do que antes.

58. Nesse sentido, o relatório "*How much time and money can AI save government?*" elaborado pela Deloitte University Press[1] aponta os principais benefícios da aplicação de IA:

a) redução de gastos em decorrência de maior eficiência operacional: de acordo com a Deloitte, a automação das tarefas dos funcionários do governo federal americano poderia economizar entre 96,7 milhões e 1,2 bilhão de horas anuais. O mesmo estudo da Deloitte também relata que a automação e a IA têm potencial para economizar entre US $ 3,3 bilhões e US $ 41,1 bilhões;

b) disponibilização de novos serviços e aprimoramento dos já existentes: aplicativos voltados para o cidadão, como ônibus autônomos e educação personalizada, melhoram a qualidade dos serviços que os governos oferecem aos cidadãos;

c) mais tomadas de decisão baseadas em dados: os governos estão coletando uma grande quantidade de dados todos os dias. No entanto, sem uma análise precisa, os dados não são adequados para *insights* acionáveis. Uma melhor tomada de decisão tem o potencial de melhorar os serviços e economizar custos.

59. A esse respeito, a publicação *Government AI Readiness Index* 2020, elaborada pela Oxford Insights e pelo Governo do Canadá[2], apresenta ranking de 172 países baseado em 33 indicadores que apontam a capacidade dos governos nacionais para implementar soluções de IA no fornecimento de serviços públicos aos cidadãos. Além disso, a publicação contém um *ranking* adicional de 32 países

[1] Disponível em https://www2.deloitte.com/content/dam/insights/us/articles/3834_How-much-time-and-money-can-AI-save-government/DUP_How-much-time-and-money-can-AI-save-government.pdf

[2] Disponível em https://www.oxfordinsights.com/government-ai-readiness-index-2020

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

quanto à possibilidade de uso responsável de IA pelos governos. Tal trabalho é uma evolução metodológica de estudos já realizados e publicados respectivamente nos anos de 2017 e 2019[1]. A figura abaixo apresenta resumidamente as dimensões que foram consideradas na avaliação realizada em 2020:

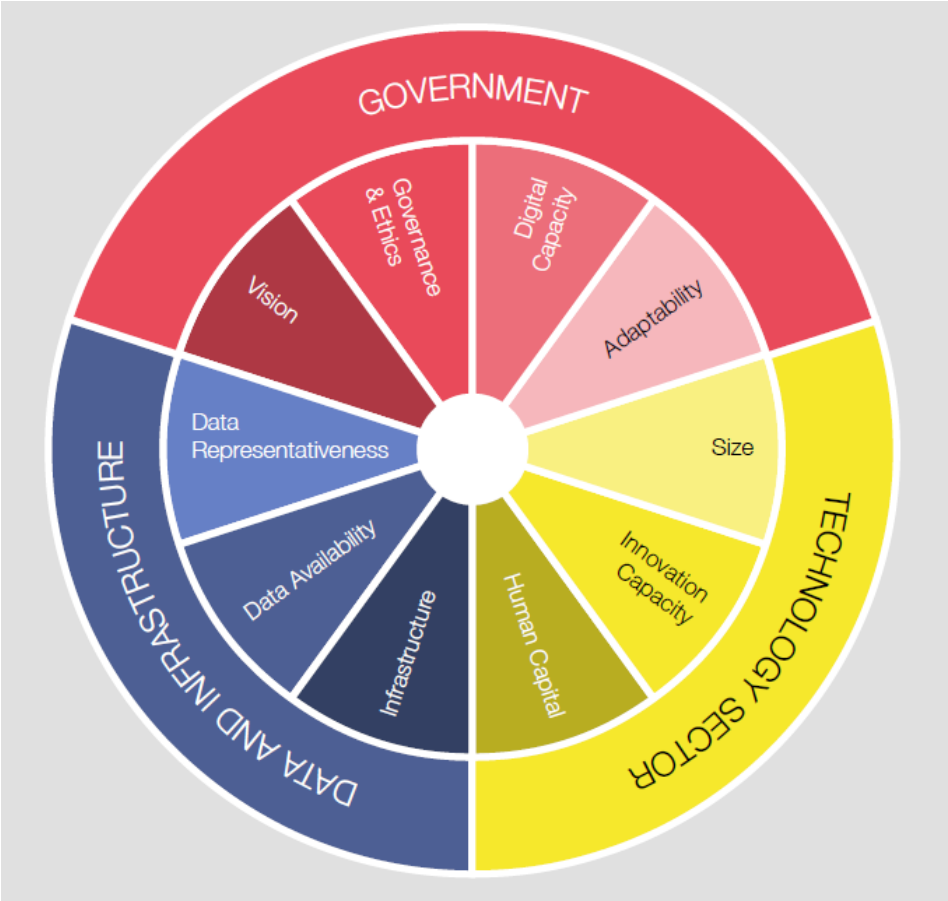


Figura 5 - Pilares e dimensões do *Government AI Readiness Index*

60. Conforme se observa, a avaliação considera três pilares (Governo; Setor de Tecnologia; e Dados e Infraestrutura), sendo que cada um dos pilares é dividido em dimensões de avaliação. O referido estudo apontou que os Estados Unidos tinham em 2020 o maior índice de capacidade de implementação de soluções de IA pelo governo, sendo que as próximas quatro posições no ranking são ocupadas por países da Europa Ocidental (Reino Unido, Finlândia, Alemanha e Suécia). As regiões de pontuação mais baixa, em média, são a África Sub-Saariana, a América Latina e o Caribe.

61. Ainda conforme o referido relatório, em 2020 o Brasil ocupava a 63ª posição no ranking mundial de capacidade do governo para implementar soluções de IA e o 6º na América Latina e Caribe. Já em relação ao uso responsável da IA, o Brasil ocupa a 30ª posição de 34 países avaliados. A figura abaixo detalha a pontuação atribuída ao Brasil nas dimensões avaliadas. Ressalta-se que o estudo é de 2020, quando o Brasil ainda não havia publicado sua estratégia para IA, o que ocorreu em 2021:

[1] Disponíveis em https://www.oxfordinsights.com/government-ai-readiness-index (2017) e https://www.oxfordinsights.com/ai-readiness2019 (2019)

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

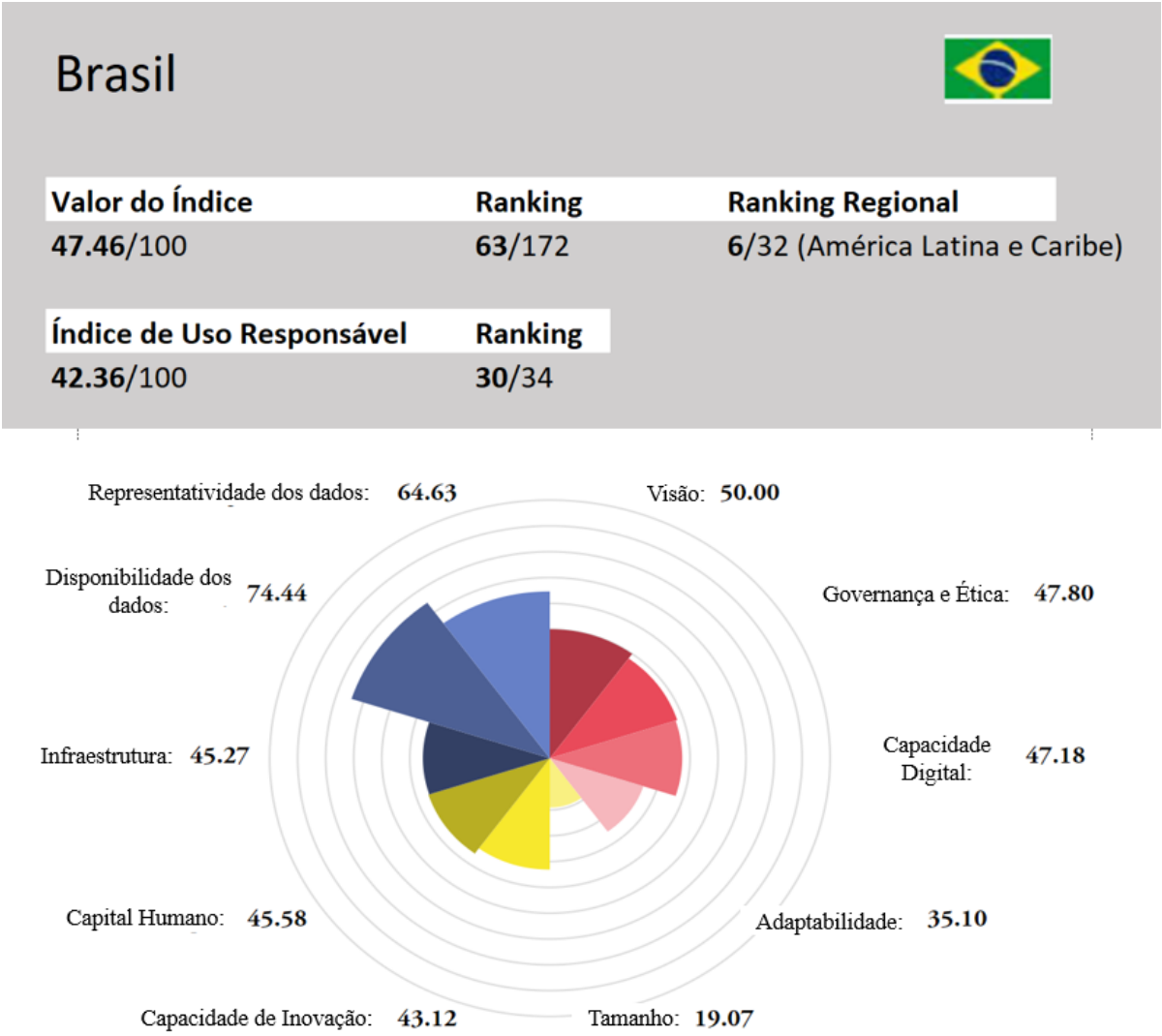


Figura 6 - Brasil – *Government IA Readiness Index*

62. A partir dos dados apresentados na figura anterior, verifica-se que os principais diferenciais para o incentivo à implementação de soluções de IA no Brasil são a disponibilidade e a representatividade de dados.

63. Nesse sentido durante o presente levantamento, buscou-se verificar os casos de uso de soluções de IA na APF, com o objetivo de se conhecer o estado atual de desenvolvimento, bem como de compartilhar no âmbito da APF as iniciativas de sucesso. Dentre as soluções implementadas, a equipe de planejamento destacou os seguintes exemplos, os quais são detalhados no Apêndice VIII:

Tabela 3 - Soluções de IA na APF

Organização	Solução	Objetivo	Principais Funcionalidades
Tribunal Superior Do Trabalho (TST)	Bem-Te-Vi	Gestão de acervo processual	• Triagem e encaminhamento de processos para analistas com base em modelo preditivo • Geração de modelos de instrução
Coordenação de Aperfeiçoamento de	Mimir	Desambiguação de dados	• Automatização do processo de preenchimento das plataformas

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Organização	Solução	Objetivo	Principais Funcionalidades
Pessoal de Nível Superior (Capes)			de governo relacionadas à pesquisa científica • Apoio à integração entre sistemas e exposição de dados • Higienização, unificação e desambiguação de dados • Aprendizado ativo
Tribunal de Contas da União (TCU)	Zello	Chatbot[1] para levar informações do TCU para o cidadão em diversos canais e redes sociais	• Interação via portal do TCU • Interação via aplicativo móvel do TCU • Interação via redes sociais • Pergunta e resposta: dúvidas sobre funcionamento do TCU e como interagir e obter informações • Consultas processuais • Consultas a contas julgadas irregulares • Emissão de certidões
Agência Nacional de Energia Elétrica (Aneel)	Águia	Monitoramento de obras em construção de usinas solares fotovoltaicas utilizando sensoriamento remoto com imagens de satélite	• Identificação dos principais marcos de implantação dos empreendimentos de geração solar (início das obras e instalação dos painéis) • Apoiar a avaliação humana e decisão dos fiscais por meio de uso de ferramentas visuais (imagens de satélite e dashboards)
Tribunal Regional do Trabalho – 5ª Região	Gemini	Agrupamento de processos similares para instrução	• Agrupamentos de processos em pilhas que contém temas semelhantes • Apresentação de processos similares já julgados
Instituto Nacional de Metrologia, Qualidade e Tecnologia (Inmetro)	Identificação de irregularidades através de imagens	Monitoramento de inspeções de equipamentos para transporte de produtos perigosos	• Avaliação da completude de registros fotográficos • Detecção de fotos clonadas
Tribunal Regional Federal – 3ª Região	Sigma e Sinara	Produtividade e padronização de instruções judiciais com reuso de modelos mais adequados às características do processo	• Seleção do modelo mais adequado para novo processo • Preenchimento de dados extraídos das peças processuais anteriores

[1] *Chatbot* é um programa de computador que tenta simular um ser humano na conversação com as pessoas. O objetivo é responder às perguntas de tal forma que as pessoas tenham a impressão de estar conversando com outra pessoa e não com um programa de computador.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Organização	Solução	Objetivo	Principais Funcionalidades
Tribunal de Contas da União (TCU)	Assistente Conjur	Sugestão de teses para resposta a questionamentos judiciais e geração de minuta de pareceres	• Seleção de teses a partir de petições iniciais ou mandados de segurança • Geração de minuta de pareceres
Secretaria do Tesouro Nacional	PropLegis	Identificação automática de padrões em proposições legislativas	• Rastreio diário de proposições na Câmara e no Senado para classificar se são de interesse do Tesouro Nacional • Classificação das proposições legislativas de interesse por área de negócio
Câmara dos Deputados	Plataforma Ulisses	Conjunto de algoritmos de inteligência artificial construídos para sustentar funcionalidades diversas necessárias à plataforma digital legislativa	• Funcionalidades diversas distribuídas entre vários sistemas, descritas em mais detalhes no Apêndice VIII

3.1 Questionário eletrônico sobre utilização de IA na Administração Pública Federal

64. Com o objetivo de obter um panorama da situação atual da utilização de soluções de Inteligência Artificial pelos diversos órgãos da Administração Pública Federal, a equipe do levantamento elaborou questionário eletrônico que tratou de vários aspectos em relação ao tema, cujo inteiro teor encontra-se no sítio https://portal.tcu.gov.br/levantamento-Inteligência Artificial.htm

65. O referido questionário foi encaminhado para 293 organizações conforme relação constante no Apêndice IX, das quais 263 organizações encaminharam as suas informações no prazo definido. Além destas, também se realizaram entrevistas com outras organizações que buscaram levantar as mesmas informações constantes no questionário eletrônico.

66. O questionário elaborado teve como principais objetivos:

a) obter informações de referência para compor uma linha de base sobre a utilização de IA pelos órgãos da APF em 2021, de modo a subsidiar futuros trabalhos do TCU que busquem acompanhar a evolução do tema na APF;

b) conhecer as principais tecnologias de IA que estão sendo utilizadas na APF, bem como seus domínios de aplicação;

c) conhecer casos de uso e experiências de sucesso na utilização de soluções de IA, com o objetivo de dar publicidade aos mesmos e incentivar o intercâmbio de experiências entre as organizações públicas;

d) levantar os principais dificultadores que os gestores públicos enfrentam para a utilização de IA na APF.

67. Ressalta-se que, durante os trabalhos de planejamento do presente levantamento, verificou-se que já haviam sido realizadas recentemente por outras organizações pesquisas que também se referiram à utilização de tecnologias de IA em organizações públicas, as quais detalham-se a seguir.

68. A primeira pesquisa identificada foi realizada em setembro de 2020 pela Transparência Brasil, em parceria com a Controladoria-Geral da União, o Ministério da Ciência, Tecnologia e Inovação (MCTI) e o Centro de Estudos sobre Tecnologias Web (Ceweb.br). O questionário foi enviado a 319

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

órgãos do Executivo Federal. Além disso, foram realizados pedidos de acesso à informação para órgãos do Legislativo e do Judiciário.

69. Tal trabalho identificou 44 ferramentas de IA usadas por órgãos governamentais, as quais foram classificadas segundo o público-alvo (interno e externo) e segundo o processo de tomada de decisão implementado nas mesmas, resultando no quadro abaixo:

	COM TOMADA DE DECISÃO	SEM TOMADA DE DECISÃO	TOTAL
Público interno	20 ferramentas	16 ferramentas	36
Externo	8 ferramentas	0 ferramentas	8
Total	28	16	44

Figura 7 - Classificação de ferramentas de inteligência artificial segundo uso para tomada de decisão e público-alvo[1]

70. Como produto, a Transparência Brasil elaborou o Relatório "Uso de Inteligência Artificial pelo Poder Público – Recomendações de Governança", o qual tem como foco principal a análise dos riscos e impactos a direitos do cidadão que as soluções de IA implementadas podem trazer. Nesse sentido, o relatório apresenta quatro considerações de governança que sintetizam os principais riscos encontrados nas soluções de IA já disponibilizadas[1]: i) bases de dados representativas e apropriadas para o contexto; ii) necessidade de supervisão humana como salvaguarda para a revisão de decisões automatizadas; iii) efetiva proteção dos dados pessoais do cidadão; e iv) transparência e explicabilidade dos sistemas.

71. Outro estudo relevante sobre a utilização de IA na APF foi publicado em 2021 pelo Centro de Inovação, Administração e Pesquisa do Judiciário da Fundação Getúlio Vargas[2], e teve como objetivo geral pesquisar o estado da arte e as iniciativas nacionais de aplicação de tecnologia com a utilização de IA nos tribunais, que busquem aperfeiçoar o sistema de justiça sob a perspectiva de seu funcionamento e de seu aparato.

72. A pesquisa da FGV evidenciou os projetos de IA já implementados e em desenvolvimento nos tribunais superiores, tribunais regionais federais, tribunais regionais do trabalho e tribunais de justiça, conforme figura abaixo:

[1] Fonte: Transparencia.Brasil. Disponível em https://www.transparencia.org.br/downloads/publicacoes/Recomendacoes_Governanca_Uso_IA_PoderPublico.pdf

[2] Disponível em https://ciapj.fgv.br/sites/ciapj.fgv.br/files/estudos_e_pesquisas_ia_1afase.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

TRIBUNAIS	EM DESENVOLVIMENTO	PROJETO-PILOTO	EM PRODUÇÃO
Tribunais Superiores (STF, STJ e TST)	4	-	5
Tribunais Regionais Federais	8	2	6
Tribunais Regionais do Trabalho	5	1	1
Tribunais de Justiça	12	4	15
Total por fase	29	7	27

Figura 8 - Fase de implementação dos projetos de IA no Poder Judiciário[1]

73. O relatório detalha todos os projetos de IA identificados nos diversos tribunais, apresentando para cada um deles a motivação do desenvolvimento, o estágio atual da solução, as funcionalidades presentes e problemas que buscam solucionar, além de apresentar os resultados obtidos. Conforme informações constantes no sítio da FGV, o projeto encontra-se em sua 2ª Fase, na qual o objeto de pesquisa foi ampliado para abranger tecnologias de "inteligência computacional" (https://ciapj.fgv.br/projetos).

74. Dessa forma, considerando-se as informações já constantes nos trabalhos mencionados acima, o questionário eletrônico aplicado pelo presente Levantamento buscou aprofundar e tem como diferencial os seguintes aspectos:

a) informações requeridas para avaliação do nível de maturidade da organização em relação à utilização de IA;

b) condições existentes referentes aos colaboradores;

c) aspectos gerais sobre o desenvolvimento de soluções de IA;

d) Informações técnicas gerais e detalhadas sobre as soluções desenvolvidas (resposta opcional).

75. Inicialmente, tem-se o seguinte quadro geral nas respostas das organizações respondentes quando perguntadas se já utilizam soluções que envolvem tecnologias de IA:

[1] Fonte: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/estudos_e_pesquisas_ia_1afase.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

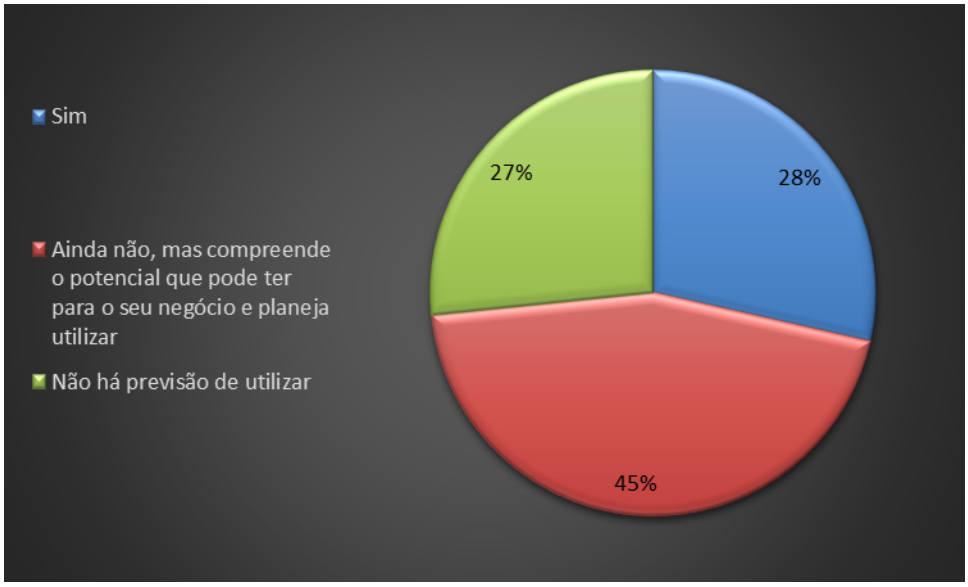


Figura 9 – Utilizam da IA pelas organizações pesquisadas

76. Partindo-se do conceito de que o presente levantamento objetiva definir uma linha de base inicial que venha a subsidiar futuros trabalhos do TCU que busquem avaliar a evolução da utilização de IA na APF, buscou-se no questionário eletrônico encaminhado às organizações públicas a inserção de questões que permitissem verificar o nível de maturidade em relação ao tema.

77. Nesse sentido, após pesquisa realizada pela equipe, optou-se por utilizar como referencial neste trabalho, com adaptações, o Modelo de Maturidade em Inteligência Artificial proposto pelo Gartner Group, publicado em 18/3/2020 (ID G00466009). O referido modelo propõe a estruturação da maturidade das organizações em relação ao uso de IA nos seguintes níveis:

Tabela 4 - Níveis de maturidade na utilização de IA - Gartner Group (ID G00466009)

Nível	Principais características
1 - Planejamento	As conversas internas sobre IA são ad hoc e especulativas. Apenas uma pequena minoria de colaboradores experimentou ou testou projetos. Nesta fase, os principais questionamentos são "O que é inteligência artificial?" e "Como a IA pode beneficiar minha linha de negócio?"
2 - Experimentação	Projetos iniciais de prova de conceito (POC) são elaborados e podem estar em fase piloto. Uma comunidade informal de prática pode ter se reunido várias vezes para compartilhamento de conhecimento e primeiros esforços de padronização.
3 - Estabilização	Os primeiros projetos de IA estão em produção. Existe um patrocinador da Alta Administração. Há orçamento exclusivo para projetos de IA. Um Centro de Especialização em IA foi criado, disponibilizando especialistas em IA, melhores práticas e tecnologia para projetos.
4 – Expansão	Todos os novos projetos digitais, incluindo revisões de processo para otimização, consideram a utilização de IA como forma de agregar valor. Existe *accountability* em torno de cada projeto. As unidades de negócio confiam nas técnicas de IA e estão prontas para usá-los. Os profissionais de DevOps / MLOps de TI são os principais responsáveis pela manutenção e atualização de soluções de IA.
5 – Transformação	A IA é rotineira e esperada como um elemento da execução de todos os processos de negócio. Todos os colaboradores no projeto e design de

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Nível	Principais características
	aplicativos são treinados e entendem os pontos fortes e riscos da IA. Os aplicativos que utilizam IA interagem de forma produtiva dentro da organização e em todo o ecossistema de negócios.

78. Além dos níveis de maturidade acima descritos, optou-se, a partir da análise das respostas ao questionário, pela criação do Nível 0, refletido pelas organizações que atualmente não utilizam e nem planejam utilizar tecnologias de IA nas soluções desenvolvidas. Dessa forma, verificou-se, no universo das 263 organizações respondentes ao questionário, a seguinte configuração em relação à maturidade para utilização de soluções de IA:

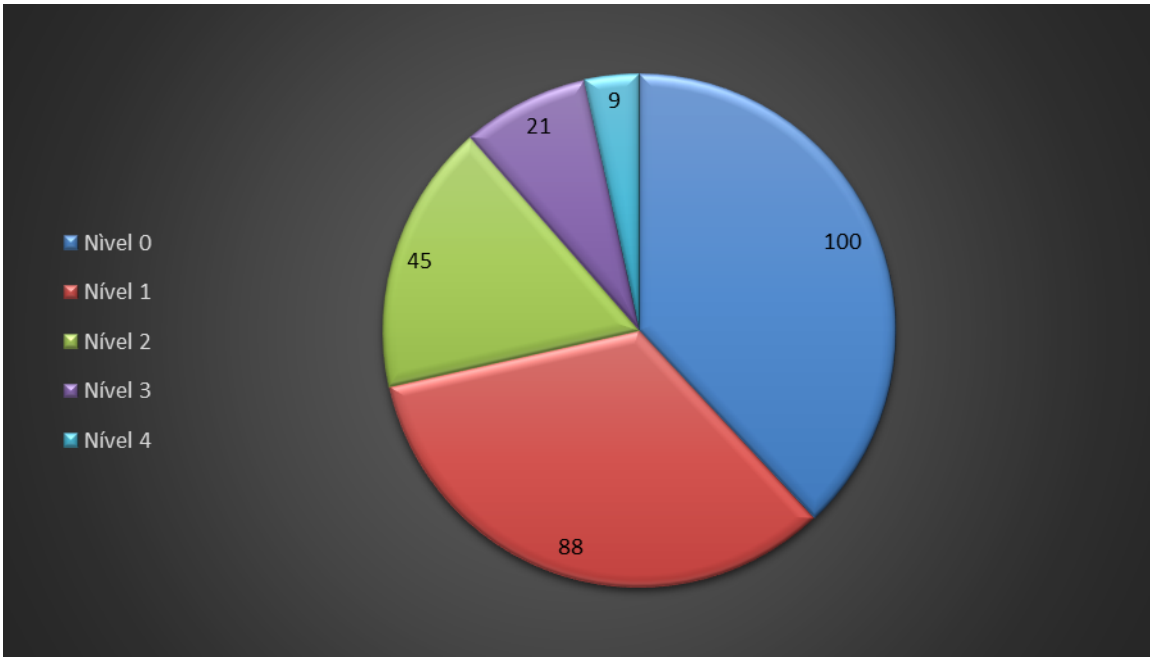


Figura 10 – Distribuição das organizações públicas federais por nível de maturidade quanto à utilização de IA

79. Conforme se observa no gráfico acima, 100 (38%) das 263 organizações públicas pesquisadas ainda não tem intenção de utilizar soluções de IA em seus processos. Dentre essas 100 organizações (nível de maturidade 0), os principais dificultadores citados para a utilização da referida tecnologia são os seguintes:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

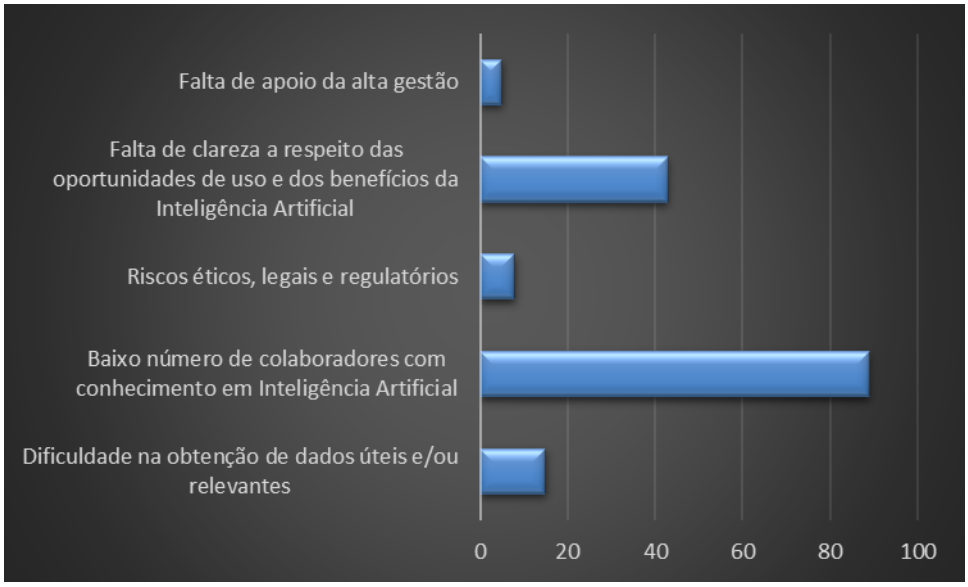


Figura 11 - Nível de Maturidade 0 - Principais Dificultadores apontados para implementação de Soluções de IA

80. Verifica-se que o ponto que mais impacta na intenção de utilização de IA nas organizações da APF é o baixo número de colaboradores com conhecimento na tecnologia. Com relação a esta questão, verificou-se no presente levantamento que, praticamente em todas as iniciativas identificadas de utilização de IA, o processo iniciou com um reduzido número de colaboradores que, em geral, trouxeram conhecimentos adquiridos externamente para dentro das organizações, antes mesmo de haver treinamentos institucionais relacionados à IA.

81. Corroborando esta afirmação, tem-se que dentre as 75 organizações que estão nos níveis de maturidade 2 a 4 e que desenvolvem internamente soluções de IA, 39 informaram dispor de no máximo cinco colaboradores que detém conhecimento em relação às tecnologias de IA.

82. Ainda com relação a este tema, tem-se que a escassez de colaboradores com conhecimento em IA pode ser suprida por meio da contratação de serviços terceirizados para o desenvolvimento de soluções ou a contratação de soluções de mercado. Neste contexto, tem-se, entre as 75 organizações com nível de maturidade 2 a 4, o seguinte cenário em relação a quem desenvolve as soluções de IA:

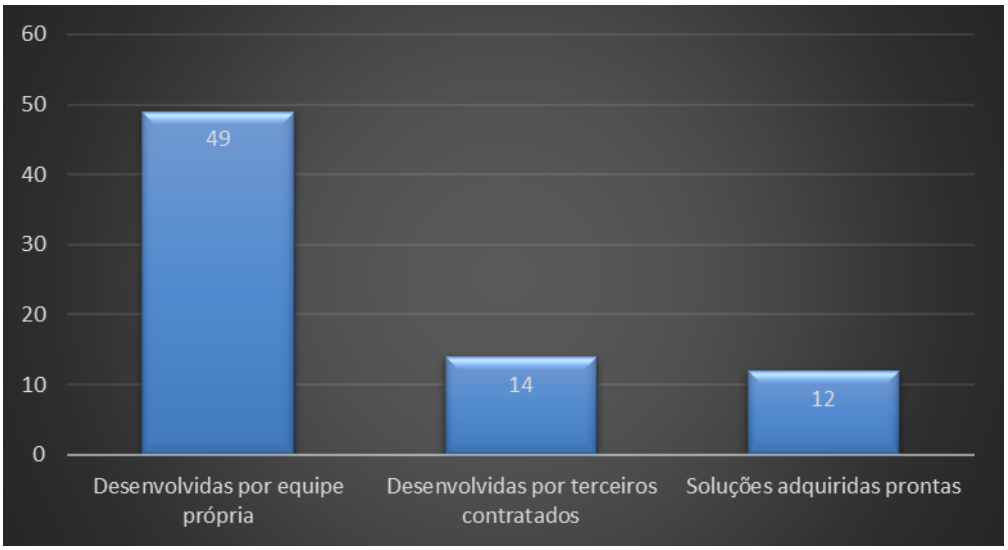


Figura 12 - Como são desenvolvidas as Soluções de IA - Níveis de Maturidade 2 a 4 (75 organizações)

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

83. Verificou-se também, em relação à capacitação de colaboradores da organização, que das 163 organizações que estão no nível de maturidade 1 a 4, 32 informaram que já dispõe de plano de capacitação institucional para desenvolvimento e sustentação de soluções de IA. Nessas mesmas organizações, a pesquisa também identificou, independentemente da existência de um plano institucional de capacitação de IA, as seguintes quantidades de treinamentos já realizados:

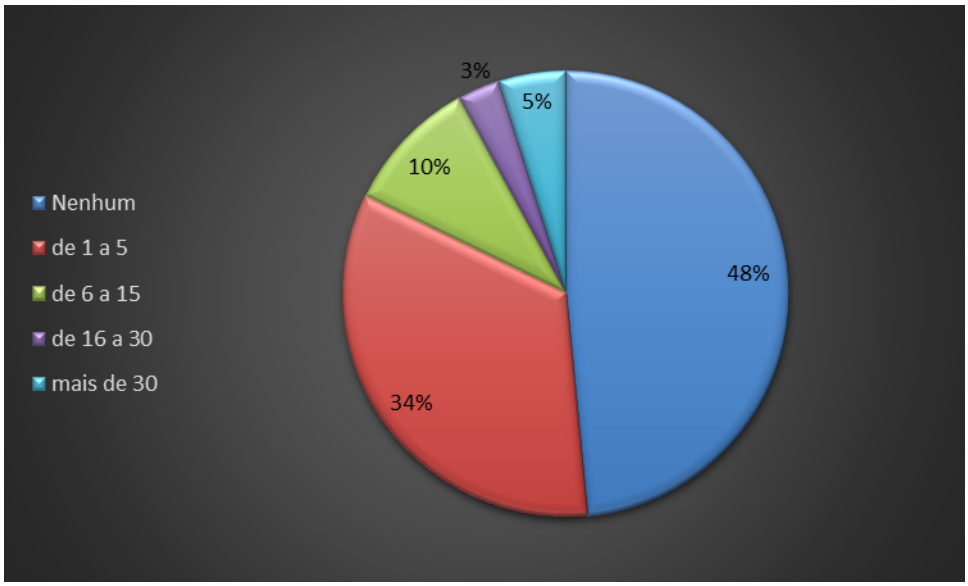


Figura 13 - Quantidade de colaboradores que receberam capacitação em desenvolvimento e sustentação de soluções de IA

84. Ainda em relação ao nível de maturidade 0, outro fator relevante que, segundo a pesquisa realizada, impede a adoção de tecnologias de IA nas organizações públicas é o desconhecimento em relação às oportunidades e benefícios que podem ser alcançados com o uso das tecnologias.

85. Se considerarmos as 163 organizações com níveis de maturidade de 1 a 4, tem-se os seguintes dificultadores apontados para implementação de soluções de IA:

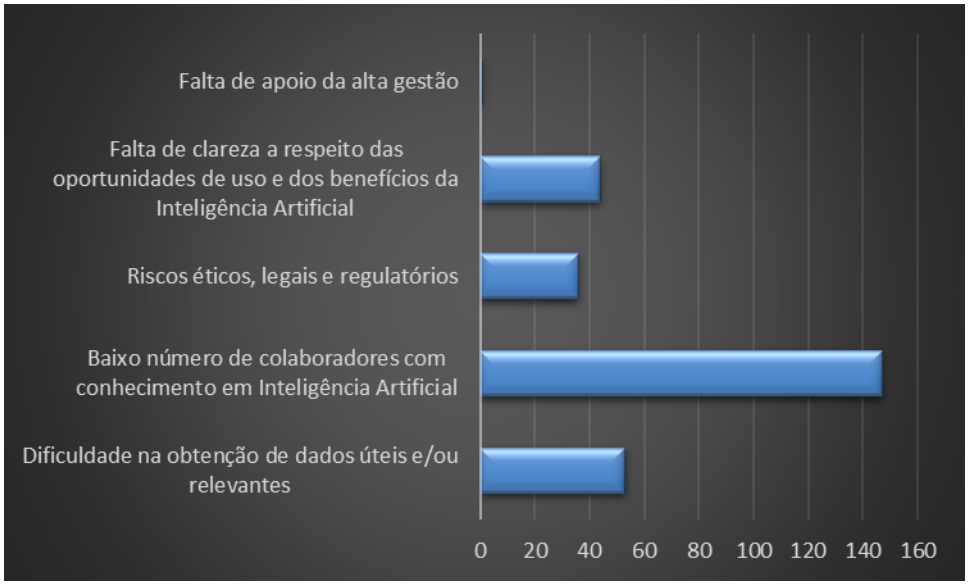


Figura 14 - Níveis de Maturidade 1 a 4 - Principais Dificultadores apontados para implementação de Soluções de IA

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

86. Verifica-se, assim, que mesmo entre as organizações que já planejam ou desenvolveram soluções de IA (níveis de maturidade 1 a 4), o baixo número de colaboradores também é o principal fator que impacta na implementação de tecnologias de IA nessas organizações. Apesar da relevância dessa questão, tem-se que apenas 32 das 163 organizações informaram que desenvolveram plano de capacitação específico para os colaboradores envolvidos no desenvolvimento ou sustentação das soluções de Inteligência Artificial. Além disso, apenas 48 (29 %) das 163 organizações com nível de maturidade 1 a 4 consideram incluir nas próximas seleções públicas conhecimentos em tecnologias de inteligência artificial como pré-requisito para contratação.

87. A pesquisa também identificou as seguintes áreas finalísticas nas quais estão sendo aplicadas as soluções de IA identificadas no presente trabalho:

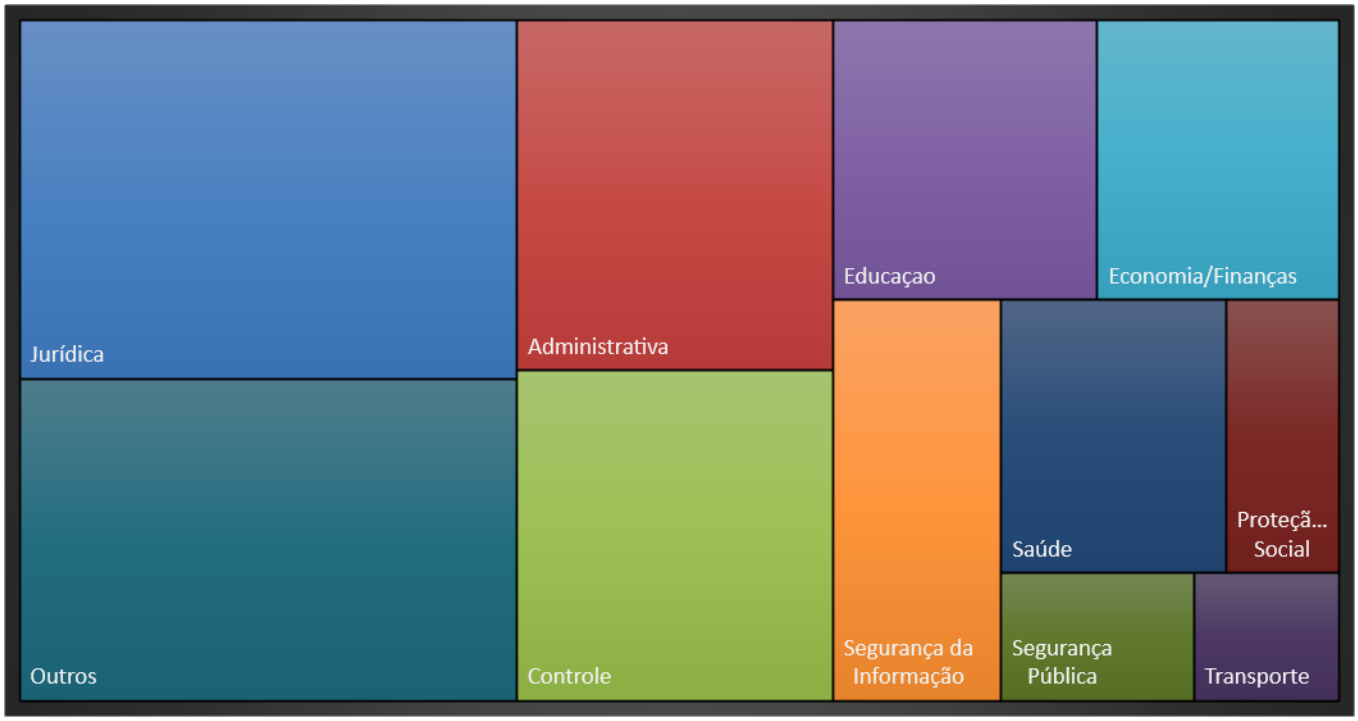


Figura 15 - Principais áreas de aplicação de soluções de IA na APF

88. Destaca-se na figura acima a predominância de utilização das soluções de IA pelos órgãos do Poder Judiciário, conforme já abordado pela pesquisa realizada pela Fundação Getúlio Vargas e detalhado nos parágrafos 71 a 73 supra. Além disso, verifica-se que há também maior utilização de tecnologias de IA nos processos administrativos internos das organizações públicas e em áreas de controle.

89. Ou seja, o questionário evidenciou que a maioria das soluções de IA implementadas até aqui não visaram o fornecimento de serviços diretos aos cidadãos, mas sim, no aprimoramento dos processos internos. Entende-se que tal quadro vai ao encontro do atual cenário pelo que passa a APF, de redução de pessoal e a consequente e necessária otimização de processos. Contudo, dado o aspecto disruptivo e inovador que IA traz, verifica-se que há espaço para crescimento do seu uso em soluções que atingem diretamente a sociedade.

90. A pesquisa também identificou que as soluções de IA em desenvolvimento, implementação ou em produção, estão distribuídas, conforme seus domínios de aplicação, conforme segue:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

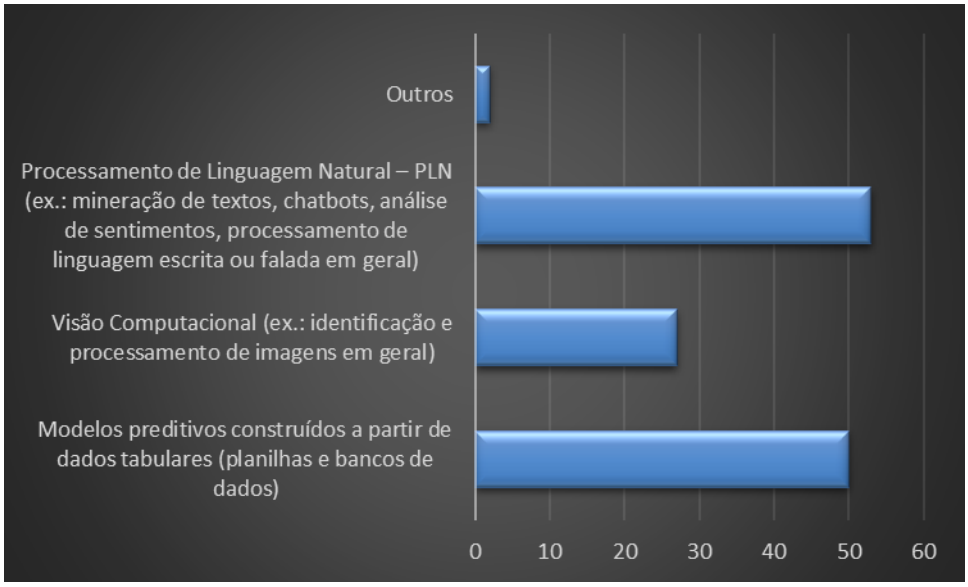


Figura 16 – Quantidade de soluções de IA na APF por domínio de aplicação

91. O domínio de processamento de linguagem natural (PLN) é amplamente utilizado nas soluções de IA que são desenvolvidas pelas organizações do Poder Judiciário, refletindo a necessidade de se tratar a grande quantidade de processos e documentos que compõem o dia a dia daquelas instituições. Além disso, destaca-se também a utilização de PLN no desenvolvimento de *chatbots* para interação com os usuários externos e o provimento de serviços. Verificou-se ainda a utilização de modelos preditivos, amplamente utilizados nas áreas de controle e fiscalização.

92. Já em relação às técnicas analíticas que são empregadas nas soluções de IA, evidenciou-se uma predominância da utilização de algoritmos tradicionais no desenvolvimento das soluções (considerando as soluções em desenvolvimento, em prova de conceito e em produção):

Técnica de IA	Qtd. Soluções implementadas
Algoritmos tradicionais de aprendizado supervisionado (Ex.: regressão linear, regressão logística, árvores de decisão, *support vector machines*, Naïve Bayes, GBM, etc.)	59
Redes neurais profundas (*Deep Learning*)	31
PLN avançado (Ex.: RNN, modelos de linguagem, *Transformers*, etc.)	32
Aprendizado não supervisionado (Ex.: agrupamentos, K-Means, DBSCAN, etc.)	32
Aprendizado por reforço (*Reinforcement Learning*)	15
Algoritmos evolucionários (Ex.: genético, colônia de formiga, enxame de abelha, etc.)	4
Algoritmos baseados em grafos	8
AutoML	4
Outros	2

93. Entre os benefícios apontados com a utilização de IA, as organizações pesquisadas informaram como principais ganhos qualitativos/quantitativos, em resumo:

a) automação de processos repetitivos com aumento de produtividade;

b) redução do tempo e custos na execução dos processos;

c) otimização na alocação de recursos humanos;

d) aumento da confiabilidade e transparência dos processos;

e) maior tempestividade na resposta aos clientes;

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

f) maior atuação preventiva das organizações em sua área de atuação;

g) uniformização de procedimentos.

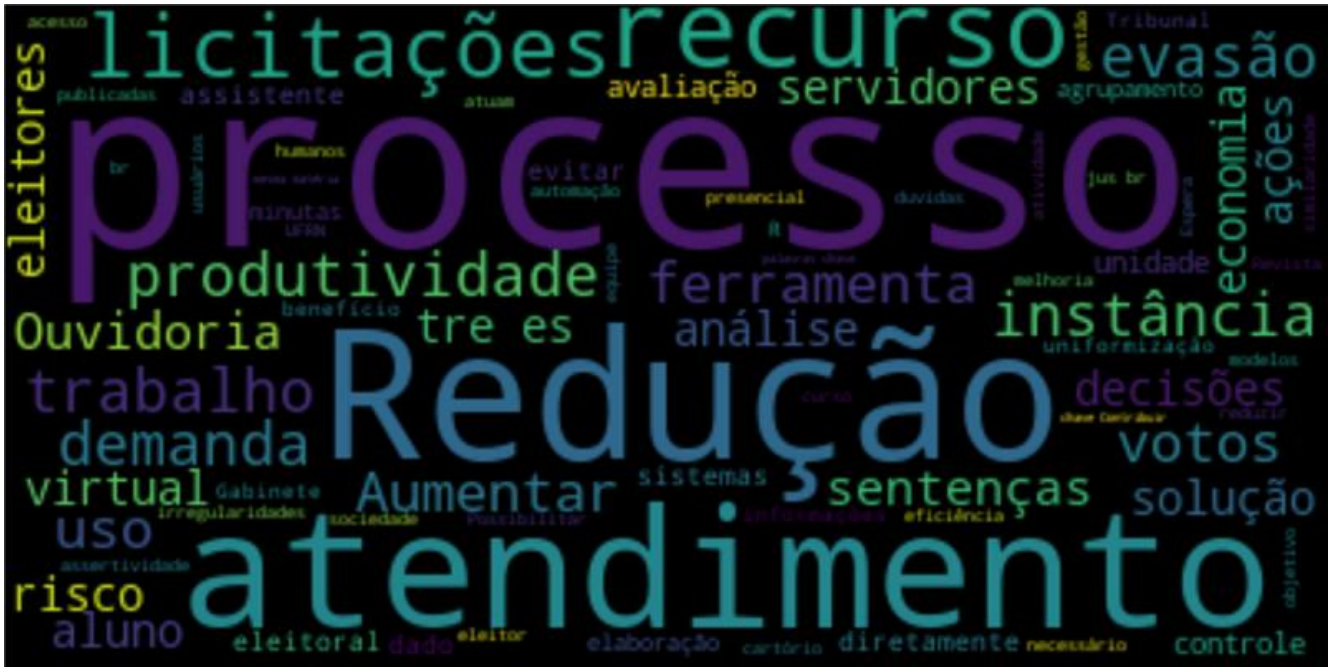


Figura 17 – *Wordcloud* - Principais benefícios obtidos com a utilização de soluções de IA identificados na pesquisa

94. Em relação às contratações realizadas pelas organizações para implementação de soluções de IA, dentre as 29 organizações que informaram a respeito, a pesquisa evidenciou que ainda são poucas as iniciativas que envolvem contratação específica para o desenvolvimento de soluções de TI. A grande maioria dos casos levantados utiliza colaboradores da própria organização no desenvolvimento das soluções e as contratações estão concentradas em insumos de infraestrutura. Importante observar que o Ministério da Saúde e o Incra informaram que utilizaram a modalidade Encomenda Tecnológica para desenvolvimento das soluções de IA:

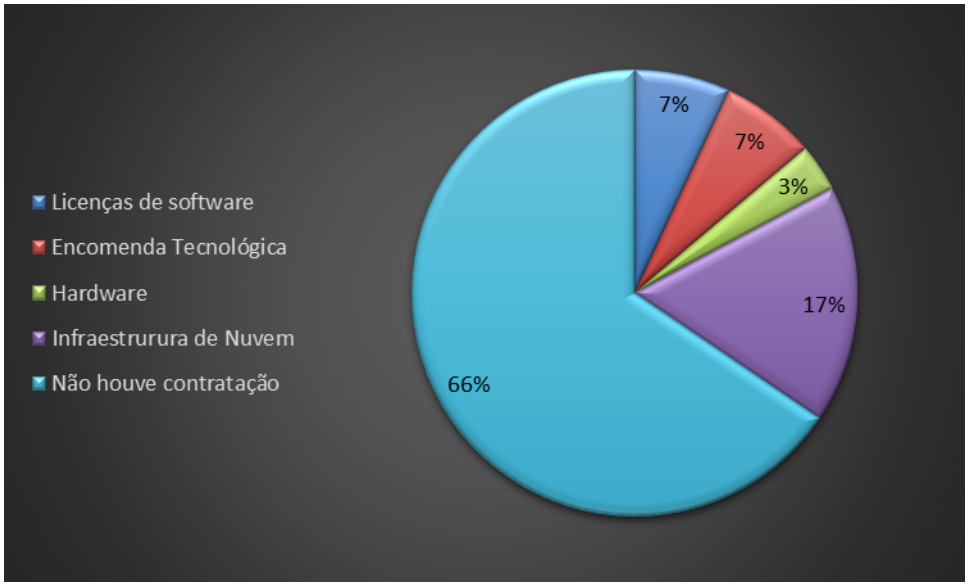


Figura 18 - Perfil de contratações para implementação de soluções de IA

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

95. Um ponto a ser a ser destacado é a possibilidade da realização de eventos denominados *Hackatons* para o desenvolvimento de soluções de IA. O *Hackathon* é um evento que reúne programadores, designers e outros profissionais externos ligados ao desenvolvimento de software para uma maratona de programação, cujo objetivo é desenvolver um software ou solução tecnológica que atenda a um fim específico. Ao final do evento, as melhores soluções desenvolvidas, geralmente em forma de protótipo, são premiadas e podem ser posteriormente refinadas e internalizadas pelas instituições. Organizações como o Tribunal Regional do Trabalho – 5ª Região (TRT5) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) já utilizaram com sucesso esse recurso.

96. Nas respostas aos questionários, foram detalhadas 36 soluções de IA desenvolvidas pelas organizações públicas pesquisadas. O Apêndice X traz uma consolidação das respostas obtidas.

4 Estratégia Brasileira de Inteligência Artificial (EBIA)

4.1 Conjuntura internacional

97. Atualmente, verifica-se que há uma corrida global entre as nações para se tornarem líderes na utilização de inteligência artificial (IA). Nesse contexto, uma das principais ações desenvolvidas é a elaboração e implementação de uma estratégia nacional para a utilização de inteligência artificial. Tal processo passa, além da elaboração da estratégia, pela implementação de legislação que dê suporte às ações planejadas em nível governamental.

98. Nesse sentido, verifica-se que as estratégias nacionais desenvolvidas abordam temas como ganhos de produtividade, impactos no mercado de trabalho, na educação e demais áreas de atuação governamental, bem como questões referentes à pesquisa e à ética e ao incentivo da utilização da IA como instrumento para otimização de processos e diminuição de custos.

99. O Instituto de Tecnologia e Sociedade do Rio (ITS) produziu em 2020 um resumo detalhado dos planos estratégicos de desenvolvimento de IA de 15 países (Estados Unidos, Canadá, Japão, Singapura, China, Emirados Árabes Unidos, Finlândia, Dinamarca, Taiwan, França, Comissão Europeia, Reino Unido, Coreia do Sul, Austrália, Suécia, Índia, México, Alemanha, Portugal e Itália)[1]. Entre os países avaliados, descreve-se abaixo as considerações apresentadas sobre a estratégia de IA algumas nações:

Tabela 5 - Objetivos das estratégias nacionais de IA[2]

País	Objetivo estratégico principal
Alemanha	O governo pretende que a "IA fabricada na Alemanha" se torne um selo de qualidade reconhecido mundialmente. A estratégia é garantir uma posição de destaque para a Alemanha como centro de pesquisa em IA, aumentando sua competitividade e promovendo os vários usos da IA na sociedade, a fim de alcançar um progresso tangível no interesse de seus cidadãos. A estratégia foi resultado de uma parceria entre diversos ministérios e foi sujeita também a uma consulta pública (Disponível em: https://www.ki-strategie-deutschland.de/home.html?file=files/ downloads /Nationale _KI-Strategie_engl.pdf
Austrália	Em junho de 2021, O governo federal australiano divulgou seu primeiro plano de ação dedicado a estabelecer a Austrália como líder global no desenvolvimento e adoção de inteligência artificial responsável. O plano estabelece quatro áreas-chave para o

[1] Disponível em https://itsrio.org/wp-content/uploads/2020/03/RelatorioAI.pdf

[2] Fonte: resumo detalhado dos Planos estratégicos de desenvolvimento de Inteligência Artificial – 2020 (ITS-2020). Disponível em https://itsrio.org/wp-content/uploads/2020/03/RelatorioAI.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

País	Objetivo estratégico principal
	governo se concentrar, incluindo: i) desenvolver e adotar IA para transformar as empresas australianas; ii) criar um ambiente para crescer e atrair os melhores talentos de IA do mundo; iii) utilizar tecnologias de IA de ponta para resolver os desafios nacionais da Austrália; e iv) tornar a Austrália uma líder global em IA responsável e inclusiva. Entre os benefícios esperados na implementação do plano, estão proteger o meio ambiente, melhorar os resultados da saúde, promover cidades inteligentes e impulsionar a economia. Disponível em: https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-action-plan
China	O Plano estratégico chinês é bastante compreensivo e detalhado. Tem sua implementação prevista em três fases, com o objetivo de tornar a China líder mundial em IA até 2030. O primeiro passo é alinhar a indústria de IA da China com os concorrentes até 2020; o segundo, alcançar a "liderança mundial" em alguns campos de IA até 2025; e terceiro, tornar-se o centro "primário" da inovação em IA até 2030.
Estados Unidos da América	A estratégia americana não possui, efetivamente, um plano traçado para coordenar suas atividades. Os estudos acerca do tema vêm sendo feitos desde o governo Obama, no ano de 2016, quando foram elaborados três relatórios. Com o governo Trump, assumiu-se uma abordagem diferente, orientada para o livre mercado de Inteligência Artificial. Em março de 2019, o governo federal dos EUA lançou o AI.gov para facilitar o acesso a todas as iniciativas governamentais de IA atualmente em andamento. O site consolida todas as iniciativas na área. Disponível em: http://ai.gov
França	A estratégia francesa visa o fortalecimento do ecossistema de IA no país para garantir sua atuação no contexto global. O foco é permanecer independente e criar uma frente de desenvolvimento. O governo se comprometeu a: i) a fomentar e apoiar o ecossistema de IA francesa; ii) melhor explorar seus bancos de dados centralizados para o desenvolvimento da IA; e iii) estabelecer uma estrutura ética abordando as questões levantadas pela IA. Disponível em: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ ENG-VF.pdf.
Reino Unido	O governo britânico lançou o Acordo Setorial para IA em 2018. Esse documento faz parte da estratégia industrial mais ampla do governo e tem como objetivo posicionar o Reino Unido como líder global em IA. Os esforços apontam políticas para melhorar pesquisas nos campos público e privado, investimento em educação (Ciência, Tecnologia, Engenharia e Matemática), melhorar a infraestrutura digital e, por fim, liderar o diálogo global no campo ético. Disponível em: https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sectordeal
México	O governo mexicano validou como estratégia o documento "*Towards an AI Strategy in Mexico: Harnessing the AI Revolution*" feito pelo Oxford Insights, C-Minds e encomendado pela Embaixada Britânica no México. De acordo com esse documento, o México busca ser um líder global em IA e digitalização como forma de promover desenvolvimento, tanto econômico como social, para seus cidadãos. O documento aponta uma série de recomendações organizadas nas categorias: i) governo e serviços públicos; ii) dados e infraestrutura digital; iii) pesquisa e desenvolvimento; iv) capacitação, habilidade e educação; e v) ética. Disponível em: https://www.gob.mx/mexicodigital/articulos/estrategiade-inteligencia-artificial-mx-2018

100. Em decorrência da relevância que tem tido a IA no contexto global, verifica-se que, além das estratégias individuais de cada nação, também há esforço para a construção de iniciativas regionais para o alinhamento em torno do tema. Nesse contexto, destacam-se as seguintes ações:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

a) **União Europeia**: há uma Aliança Europeia em IA que estabeleceu uma abordagem abrangente e um acordo de cooperação entre os países europeus. Vinte e cinco países europeus assinaram uma "Declaração de Cooperação" em Inteligência Artificial em abril de 2018. Embora alguns desses países também tenham iniciativas nacionais de IA, eles enfatizaram a importância de atuarem juntos para aprimorar a pesquisa e a implantação, enquanto lidam coletivamente com questões sociais, econômicas, éticas e legais;

b) **Região Nórdico-báltica**: os Ministros da Dinamarca, Estônia, Finlândia, Ilhas Faroe, Islândia, Letônia, Lituânia, Noruega, Suécia e Ilhas Åland emitiram uma declaração de colaboração sobre IA em maio de 2018, "IA na região nórdico-báltica". A iniciativa tem o objetivo de desenvolver e promover o uso da inteligência artificial para servir melhor os seres humanos. Os países pretendem colaborar no desenvolvimento de competências, acesso a dados, padrões e princípios, garantindo o papel da IA no Mercado Único Digital Europeu, evitando regulamentação desnecessária e utilizando o Conselho Nórdico de Ministros para facilitar a colaboração;

c) **Membros do G7:** os países líderes do G7 - Canadá, França, Alemanha, Itália, Japão, Reino Unido e Estados Unidos - se comprometeram com a "Visão Comum de Charlevoix para o Futuro da Inteligência Artificial". Essa visão compartilhada destaca o potencial de crescimento econômico ao lado de desafios sociais urgentes; afirma que "a IA que promove o crescimento econômico, a confiança da sociedade, a igualdade de gênero e a inclusão depende de um ambiente político previsível e estável que promova a inovação".

101. É nesse contexto mundial que se avalia na próxima seção a Estratégia Brasileira de Inteligência Artificial.

4.2 Estratégia Brasileira de Inteligência Artificial

4.2.1 Descrição geral

102. Instituída pela Portaria MCTI 4.617/2021, alterada pela Portaria MCTI 4.979/2021, a Estratégia Brasileira de Inteligência Artificial (EBIA) assume o papel de nortear as ações do Estado brasileiro em prol do desenvolvimento das ações, em suas várias vertentes, que estimulem a pesquisa, inovação e desenvolvimento de soluções em Inteligência Artificial, bem como, seu uso consciente, ético e em prol de um futuro melhor.

103. Alinhada às diretrizes da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) endossadas pelo Brasil, a EBIA fundamenta-se nos cinco princípios definidos pela Organização para uma gestão responsável dos sistemas de IA, quais sejam: (i) crescimento inclusivo, o desenvolvimento sustentável e o bem-estar; (ii) valores centrados no ser humano e na equidade; (iii) transparência e explicabilidade; (iv) robustez, segurança e proteção e; (v) a responsabilização ou a prestação de contas (*accountability*).

104. A EBIA tem seis principais objetivos: (i) contribuir para a elaboração de princípios éticos para o desenvolvimento e uso de IA responsáveis; (ii) promover investimentos sustentados em pesquisa e desenvolvimento em IA; (iii) remover barreiras à inovação em IA; (iv) capacitar e formar profissionais para o ecossistema da IA; (v) estimular a inovação e o desenvolvimento da IA brasileira em ambiente internacional; e (vi) promover ambiente de cooperação entre os entes públicos e privados, a indústria e os centros de pesquisas para o desenvolvimento da Inteligência Artificial.

105. Para tanto, a Estratégia estabelece nove eixos temáticos, caracterizados como os pilares do documento, apresenta um diagnóstico da situação atual da IA no mundo e no Brasil, destaca os desafios a serem enfrentados, oferece uma visão de futuro, e apresenta um conjunto de ações estratégicas que a aproxima dessa visão:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

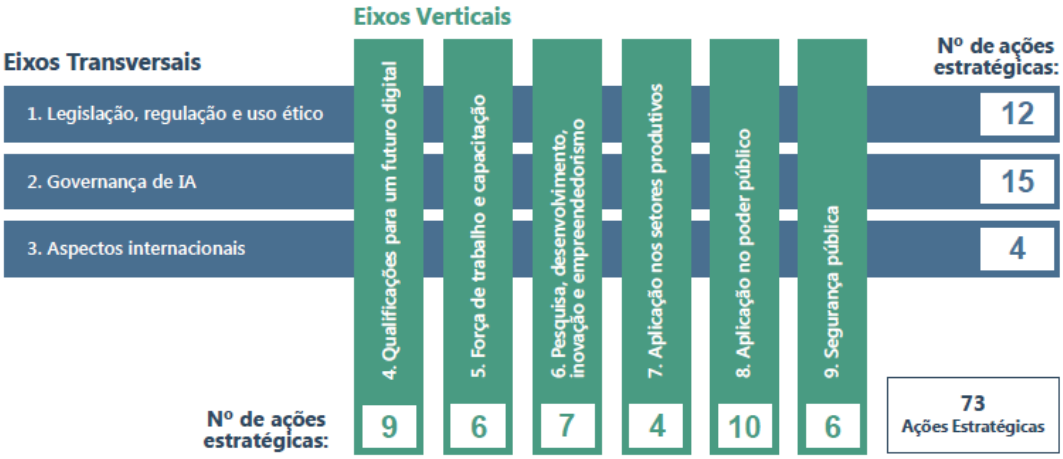


Figura 19 – Eixos temáticos da EBIA

106. O processo de elaboração da EBIA foi construído por meio do desenvolvimento das seguintes etapas:

106.1. contratação de consultoria especializada em IA para realizar estudo a respeito dos potenciais impactos sociais e econômicos das tecnologias que embasam a IA, com vistas a analisar propostas de políticas que mitigassem seus efeitos negativos ao mesmo tempo em que maximizassem seus efeitos positivos;

106.2. *benchmarking* nacional e internacional;

106.3. processo de consulta pública à sociedade, realizada por meio de plataforma eletrônica do Governo Federal, entre 12 de dezembro de 2019 e 3 de março de 2020, a qual recebeu mais de mil contribuições.

107. Ressalta-se que o texto da Portaria MCTI 4.979/2021 reconhece que a EBIA é considerada uma política pública[1] e, nesse contexto, deve ser entendida como um conjunto de intervenções e diretrizes emanadas de atores governamentais, que visam tratar, ou não, problemas públicos e que requerem, utilizam ou afetam recursos públicos.

4.2.2 Metodologia utilizada para avaliação da EBIA

108. Uma vez que o Relator determinou que fosse incluído no escopo do presente levantamento a identificação dos principais riscos e oportunidades associados à implementação da EBIA, foi necessário estabelecer primeiramente quais seriam os critérios mais adequados para atender a essa demanda.

109. Ressalta-se que o "Referencial para Avaliação de Governança em Políticas Públicas", publicado pelo TCU em 2014, assim define o termo "políticas públicas" (BRASIL, TCU, 2014b, p. 21):

> Políticas públicas são um conjunto articulado e estruturado de ações e incentivos que buscam alterar uma realidade em resposta a demandas e interesses dos atores envolvidos. Dispõem usualmente sobre "o que fazer" (ações), "aonde chegar" (objetivos relacionados ao estado de coisas que se pretende alterar) e "como fazer" (princípios e macroestratégias de ação). Política Pública diz respeito, portanto, "à mobilização político-administrativa para articular e alocar recursos e esforços para tentar solucionar dado problema coletivo".

[1] De acordo com a própria Portaria MCTI 4.617/2021, a EBIA a deve ser uma política pública constantemente acompanhada, avaliada e ajustada, pois o ritmo da evolução tecnológica da Inteligência Artificial tende a se acelerar.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

110. O objetivo do Referencial de Controle de Políticas Públicas é estabelecer uma estrutura de critérios comuns que possam compor um modelo para avaliar o nível de maturidade de políticas públicas, orientando e sistematizando ações de controle, de modo a contribuir para a melhoria do desempenho e dos resultados das políticas públicas, por meio do aprimoramento de seu processo de formulação, implementação e avaliação.

111. De acordo com esse referencial, para que os estágios de políticas sejam realizados de maneira adequada, algumas fases devem ser cumpridas com a adoção de boas práticas nacionais e internacionais para políticas públicas de qualidade. Essas fases estão organizadas com base em sete blocos de controle: (i) diagnóstico do problema e formação da agenda; (ii) análise de alternativas e tomada de decisão; (iii) desenho e institucionalização da política; (iv) estruturação da governança e gestão; (v) alocação e gestão de recursos orçamentários e financeiros; (vi) operação e monitoramento; e (vii) avaliação. A figura a seguir ilustra a interpelação entre estes componentes:

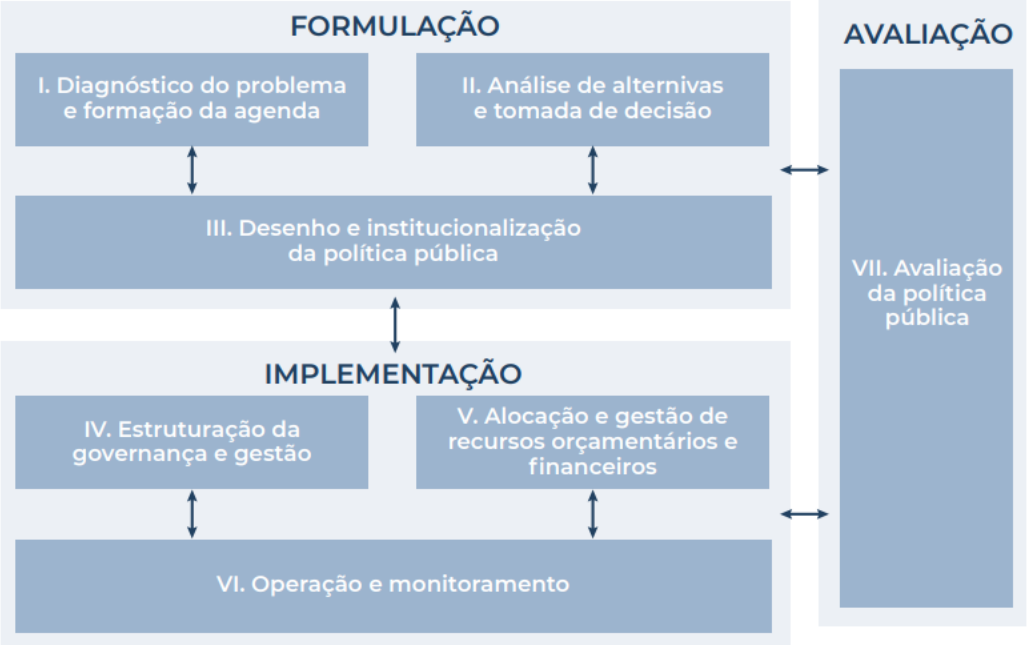


Figura 20 – Blocos de controle do Referencial de Controle de Políticas Públicas.

112. Como a EBIA foi publicada durante a execução do presente levantamento e, portanto, ainda não havia produzido completamente seus efeitos, optou-se por selecionar apenas os quatro primeiros blocos de controle para a sua avaliação, os quais são conceituados a seguir:

112.1. **Diagnóstico do problema** refere-se ao conjunto de práticas voltadas para a identificação do problema, sua delimitação e caracterização, identificação de público-alvo e de possíveis causas, efeitos e formas de tratamento. **Formação da agenda pública** envolve a escolha do conjunto de problemas ou temas considerados merecedores de intervenção pública.

112.2. **Análise de alternativas** se desenvolve por meio de escrutínios formais ou informais das consequências do problema e confronta aspectos como custo-benefício e custo-efetividade de cada alternativa disponível para apoiar o processo de tomada de decisão. **Tomada de decisão** representa o momento em que os interesses dos atores são equacionados e as intenções (objetivos e métodos) de enfrentamento de um problema são explicitadas.

112.3. **Desenho da política pública** contempla a caracterização da política com base em modelo lógico que explicita, entre outras coisas, seus objetivos, produtos, atividades, resultados e impactos. A **Institucionalização** envolve a publicação da política pública por meio de norma legal (lei, decreto, resolução etc.) apropriada, emitida por órgão dotado de legitimidade e competência para fazê-lo.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

112.4. **Estruturação da governança e gestão** envolve a definição das estruturas de governança, de gestão de riscos e controles internos, de monitoramento e de avaliação da política pública, bem como do plano de implementação da política, dos processos e operações necessárias ao seu funcionamento.

113. O objetivo da aplicação de um questionário de autoavaliação de controles (CSA) é a avaliação do nível de maturidade, o que significa verificar a qualidade do processo de formulação, implementação e avaliação da EBIA frente a um conjunto de boas práticas que, reconhecidamente, contribuem para a melhoria de seu desempenho e resultados.

114. Especificamente, buscou-se avaliar os seguintes aspectos baseados no referencial citado:

a) se o planejamento identificou os resultados esperados na transformação digital;

b) se a elaboração das estratégias contou com a participação das partes interessadas;

c) se a EBIA possui uma visão estratégica de longo prazo;

d) se foi explicitado o estágio de referência inicial da política (linha de base) que servirá de subsídio para avaliação do seu resultado;

e) se a definição dos objetivos é precisa para orientar a execução das ações;

f) se os esforços e os resultados são passíveis de mensuração, considerando aspectos como viabilidade física, financeira e temporal;

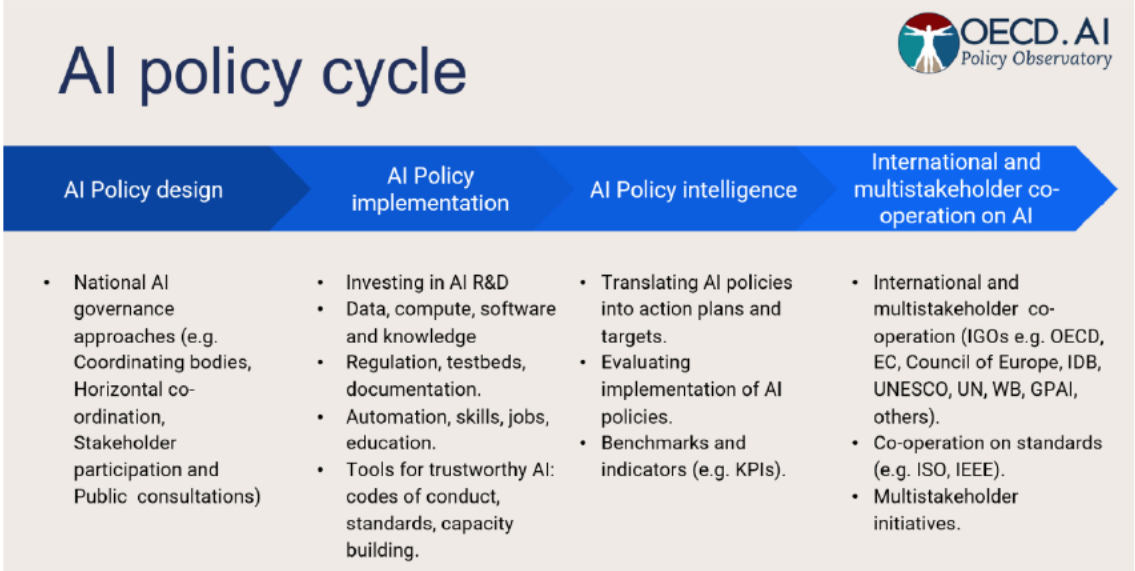
g) se os objetivos e as metas são coerentes com a visão estratégica da EBIA;

h) se a ordem de precedência para a realização de objetivos e metas dependentes entre si está identificada;

i) se foram definidos prazos para a realização dos objetivos intermediários e finais;

j) se o sistema de monitoramento e avaliação da transformação digital dispõe de processos, procedimentos e meios suficientes (recursos financeiros, pessoas, estrutura etc.) para assegurar o monitoramento e a avaliação que forneça informações confiáveis, tempestivas e necessárias para a tomada de decisão.

115. Ressalta-se que a metodologia utilizada está alinhada com a visão da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) sobre o ciclo de uma política de Inteligência Artificial. Entende-se que os critérios selecionados para avaliação da EBIA cobrem as três primeiras fases do ciclo proposto pela OCDE, o qual é constituído pelas seguintes etapas:



Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Figura 21 – Ciclo de uma Política de Inteligência Artificial. Fonte: OCDE (adaptado)

4.2.3 Achados de auditoria

***Achado I** – Os objetivos da EBIA não são específicos, mensuráveis, realistas (alcançáveis) e delimitados em um recorte temporal*

Situação encontrada

116. A operacionalização de uma política pública é delineada em função dos objetivos propostos. De acordo com a Lei 13.971/2019, que instituiu o Plano Plurianual da União para o período 2020 a 2023, objetivo é uma declaração de resultado a ser alcançado que expressa, em seu conteúdo, o que deve ser feito para a transformação de determinada realidade.

117. Feitas essas considerações, importa dizer que a EBIA não define precisamente os objetivos para enfrentamento dos problemas públicos ou aproveitamento das oportunidades no que diz respeito ao uso da Inteligência Artificial.

118. Na introdução da EBIA são apresentados seis objetivos estratégicos, conforme ilustrado na figura abaixo[1]:

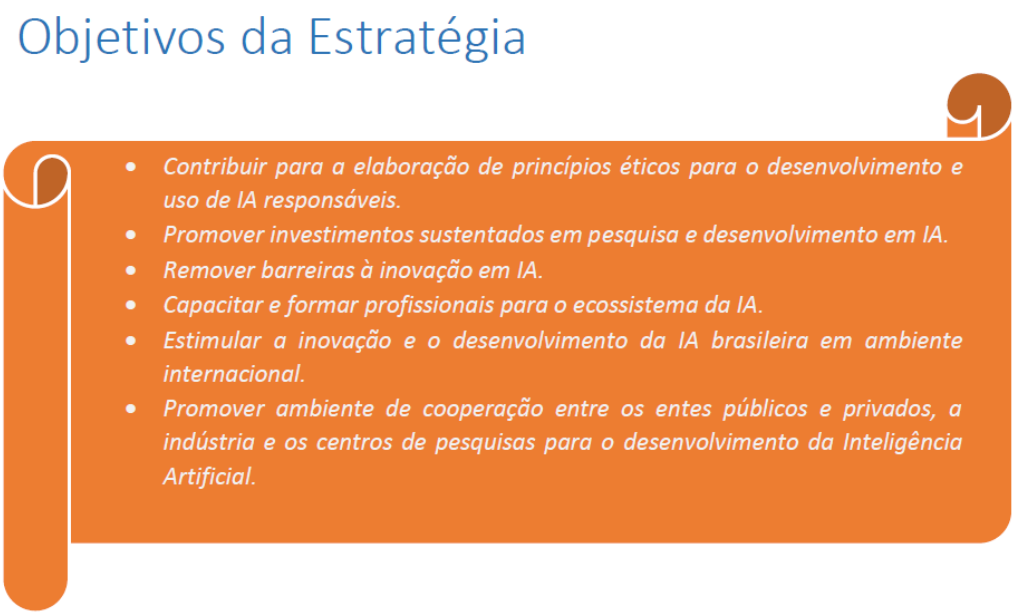


Figura 22 – Objetivos Estratégicos da EBIA

119. O TCU ("Referencial para Avaliação de Governança em Políticas Públicas do TCU"[2] e Acórdão 2.359/2018-TCU-Plenário) e o Instituto de Pesquisa Econômica Aplicada - IPEA (Guia Prático de Análise *Ex Ante*), alinhados às melhores práticas internacionais, enfatizam a necessidade de definir os objetivos e os respectivos indicadores de acordo com o modelo SMART, cujas propriedades desejáveis são as de serem:

a) específicos (*specific*): indicadores devem refletir informações simples e facilmente comunicáveis, objetivando um aspecto específico a ser submetido a mudanças pela intervenção;

b) mensuráveis (*measurable*): indicadores qualitativos e quantitativos devem ser mensuráveis, possibilitando aferir se os resultados propostos foram alcançados ou não;

[1] Fonte: Estratégia Brasileira de Inteligência Artificial

[2] Disponível em: https://portal.tcu.gov.br/governanca/governancapublica/politicas-publicas/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

c) alcançáveis (*achievable*): a meta estabelecida por um indicador deve ser alcançável e realista em relação ao contexto em que se insere a intervenção;

d) relevantes *(relevant*): indicadores devem refletir informações relacionadas aos componentes que medem;

e) temporalmente regulares (*time*): indicadores devem explicitar a periodicidade com que precisarão ser medidos para que possam ser úteis à medição de resultados.

120. Porém, os objetivos da EBIA não satisfazem totalmente a propriedade SMART (sigla em inglês para *Specific, Measurable, Attainable, Realistic e Time-bound*), uma vez que não são específicos, mensuráveis, realistas (alcançáveis) e delimitados em um recorte temporal. É percebido apenas a propriedade de serem apropriados (relevantes).

121. Em razão dessas falhas, na prática, os objetivos apresentados na EBIA correspondem mais a diretrizes estratégicas que orientam as ações derivadas dela. Não se sabe em que de concreto se traduzem os objetivos definidos.

122. Outro problema encontrado é que as ações estratégicas da EBIA não definem indicadores claros para a ação governamental, conforme requer a propriedade SMART. Além disso, embora seja perceptível que as ações estratégicas estão relacionadas com seus eixos temáticos, a EBIA não diz explicitamente a quais objetivos as ações estratégicas estão associadas, dificultando a avaliação do nível de coerência entre eles.

123. Diante dessas inconsistências, cumpre registrar que, para evitar o comprometimento de suas ações, o TCU tem o entendimento de que não é viável ou aceitável que os objetivos definidos em uma política pública não sejam expressos em termos de resultados.

124. Assim, o MCTI deveria estabelecer objetivos (ou indicadores) claros para balizar a ação governamental, à exemplo do que fez o Ministério da Educação por meio da Lei 13.005/2014, que instituiu o Plano Nacional de Educação (PNE). A figura abaixo exemplifica um dos vinte objetivos definidos pelo PNE, os quais possuem as características de específico, mensurável, apropriados, realista e com prazo delimitado:

META 9 Elevar a taxa de alfabetização da população com 15 (quinze) anos ou mais para 93,5% (noventa e três inteiros e cinco décimos por cento) até 2015 e, até o final da vigência deste PNE, erradicar o analfabetismo absoluto e reduzir em 50% (cinquenta por cento) a taxa de analfabetismo funcional.

Figura 23 – Objetivo 9 do Plano Nacional de Educação

125. Outro exemplo do Governo Federal que pode ser citado é o Decreto 10.332/2020, que institui a Estratégia de Governo Digital para o período de 2020 a 2022, o qual na Iniciativa 8.2 estabelece que devem ser implementados recursos de inteligência artificial em, no mínimo, 12 serviços públicos federais, até 2022.

126. Portanto, uma vez que a definição genérica de objetivos não é suficiente para orientar a execução das ações, entende-se que o MCTI deveria ter elaborado os objetivos ou as ações estratégicas conforme o modelo SMART.

Critérios

a) Acórdão 2.359/2018-TCU-Plenário, relatoria do Min. Vital do Rêgo;

b) Avaliação de Políticas Públicas - Guia Prático de Análise *Ex Ante*, IPEA;

c) Lei 13.971/2019.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Evidências

a) Seção "Objetivos da Estratégia" da EBIA (Portaria MCTI 4.979/2021).

Causas

a) Não identificadas.

Efeitos

a) Dificuldade de avaliar o nível de coerência entre as ações estratégicas e os objetivos definidos pela EBIA;

b) Dificuldade em estabelecer riscos específicos ao alcance de resultados da estratégia;

c) Impossibilidade de avaliar se os objetivos fixados guardam nítida relação com a finalidade última da EBIA ou com os problemas/oportunidades previamente identificados;

d) Insuficiência dos objetivos para orientarem a execução de ações.

Proposta de Encaminhamento

a) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

Achado II – *Inexistência da explicitação do estágio de referência inicial ("marco zero", "situação inicial") e da definição de indicadores e metas para subsidiar a avaliação do desempenho e resultados da EBIA*

Situação encontrada

127. A elaboração de indicadores quantitativos sobre o problema permite dimensioná-lo no tempo anterior à implementação da política, traçando o seu panorama. Nota-se que, sem valores da linha de base acerca da realidade à qual a política visa promover mudanças, não se poderá desenvolver futuramente nenhuma avaliação relevante do atingimento de seus resultados.

128. Nesse sentido, cumpre dizer que a EBIA falhou em responder questões triviais como: Quais são os indicadores-chave e metas para os principais objetivos? Qual é a situação inicial observada desses indicadores? Em qual magnitude vão ser tratados, ou seja, qual é o grau das mudanças geradas pela política que serão consideradas satisfatórias?

129. Em que pese a EBIA ter uma seção específica destinada ao diagnóstico da situação atual, apresentando dados qualitativos e quantitativos sobre do uso de IA no país, convém relatar que não foram explicitamente definidos e priorizados quais indicadores a EBIA pretende medir, com o intuito de subsidiar posterior avaliação do desempenho da intervenção pública.

130. Importante salientar que é recomendada a definição de indicadores que acompanhem o fluxo de operacionalização da política, como: insumos/recursos, processos, produtos, resultados e impactos, conforme ilustrado na figura abaixo:

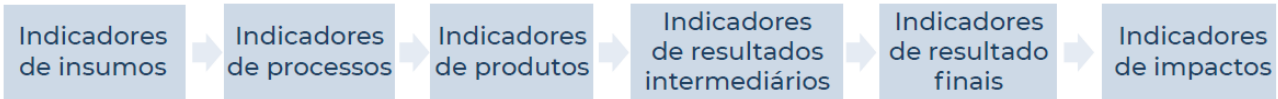


Figura 24 – Cadeia de indicadores desejável

131. Nada obstante, entende-se que, no mínimo, a EBIA deveria ter os indicadores de resultado associados aos seis objetivos estratégicos, ou às 73 ações estratégicas, conforme previsto no modelo SMART (Acórdão 2.359/2018-TCU-Plenário e Guia Prático de Análise *Ex Ante)*.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

132. Com efeito, verifica-se que os indicadores poderiam ser derivados de informações apresentadas na própria EBIA, tais como "O Brasil está situado na 66ª posição do Índice de Inovação Global de 2019" e "O número de pós-graduações em programas de tecnologias digitais no Brasil (i.e. que envolvam robótica, Inteligência Artificial / aprendizado de máquina ou big data) em 2016 era de 224, sendo 72 mestrados e 152 doutorados". Por exemplo, um indicador básico que a EBIA poderia medir seria a posição no ranking de inovação global e a quantidade de cursos de pós-graduação ofertados. Além disso, os indicadores deveriam ser associados às ações estratégicas e, consecutivamente, aos objetivos estratégicos da EBIA para acompanhar o andamento e medir o cumprimento dos objetivos de uma política.

133. Dessa forma, seria possível definir o conjunto de indicadores (associados aos eixos temáticos) a serem medidos pela EBIA com os respectivos dados sobre a situação atual, o que representaria o "estágio de referência inicial" ou "marco zero" da Estratégia, à exemplo do que foi feito pelo PNE no documento intitulado "Linha de Base" (http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkW1/document/id/493812).

134. Entretanto, isso não foi feito, o que quer dizer também que não foram definidas as metas para cada um dos indicadores. Nesse contexto, faz-se necessário esclarecer a diferença entre metas e indicadores. De acordo com a Lei 13.971/2019, indicador é instrumento gerencial que permite a mensuração de desempenho de programa em relação à meta declarada, enquanto meta é a declaração de resultado a ser alcançado, de natureza quantitativa ou qualitativa, que contribui para o alcance do objetivo.

135. Na elaboração de uma política pública, é esperado que sejam estabelecidas metas de entregas de produtos ou benefícios para um horizonte temporal, com vistas ao alcance dos objetivos pretendidos. Isso quer dizer que no caso da EBIA, deveriam ser estabelecidas metas precisas e objetivamente caracterizadas, que concorressem para a consecução dos propósitos mais gerais da intervenção pública, de modo a orientar as ações governamentais e assegurar a transparência sobre metas e resultados.

136. Uma meta necessita, pelo menos, de um quantitativo explícito e de um prazo máximo de execução para ser clara e mensurável. Dessa forma, quanto à possibilidade de mensuração dos resultados da EBIA, considerando aspectos como viabilidade física e temporal, a equipe de fiscalização verificou que, das 73 ações estratégicas, poucas apresentaram um quantitativo claro ou o ano máximo de execução (sendo raras exceções).

137. Diante dessas falhas, um exemplo do que poderia ser feito para atender ao Modelo SMART é, por exemplo, reescrever a ação estratégica "instituir programas de formação tecnológica para professores e educadores", do eixo vertical "Qualificações para um futuro digital" adicionando informações, tais como: "instituir, **na modalidade EAD**, **três** programas de formação tecnológica, **nas áreas de estatística, ciência de dados e inteligência artificial** para atender a no mínimo **2.000** professores e educadores, **até o ano de 2023**".

138. Entende-se que a definição de metas quantitativas é essencial para a avaliação da eficácia, eficiência e efetividade da política. O uso de indicadores quantitativos e/ou qualitativos apropriados deveria fazer parte do desenho da política, de modo que esses parâmetros iniciais possam ser comparados com os resultados efetivos e as metas alcançadas. Contudo, faltou clareza da EBIA em relação aos resultados finais que pretende atingir com a sua implementação.

Critérios

a) Acórdão 2.359/2018-TCU-Plenário, relatoria do Min. Vital do Rêgo;

b) Avaliação de Políticas Públicas - Guia Prático de Análise *Ex Ante*, IPEA;

c) Lei 13.971/2019, art. 3º, inciso II.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Evidências

a) Estratégia Brasileira de Inteligência Artificial (Portaria MCTI 4.979/2021).

Causas

a) Não identificadas.

Efeitos

a) Falta de clareza acerca dos produtos e efeitos esperados da política;

b) Impossibilidade de monitorar o desempenho da política em termos de alcance de objetivos e produção de resultados de curto e médio prazo;

c) Ausência de entendimento dos resultados pretendidos com a EBIA pela sociedade.

Proposta de Encaminhamento

d) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

***Achado III** – Falhas formais na apresentação do modelo lógico e sobre como a EBIA incide sobre os problemas/oportunidades, projetando seus resultados e impactos de longo prazo*

Situação encontrada

139. Uma das formas de se entender a EBIA é analisá-la como uma solução estruturada para capturar benefícios, solucionar problemas e remover barreiras do uso da IA no Brasil.

140. Nesse sentido, a literatura sobre planejamento no setor público considera como boa prática o uso de técnicas que auxiliam no desenho da política pública. Nada obstante, importa dizer que foram observadas falhas formais na elaboração da EBIA, devido à ausência de artefatos requeridos por normativos legais, relacionados à qualidade do processo de formulação de uma política pública.

141. Verificou-se que a EBIA não atendeu a dois requisitos importantes na fase de diagnóstico do problema e desenho, quais sejam: (i) apresentar de forma clara e objetiva a teoria que sustenta a política, explicitando a forma como a política incide sobre as causas e os efeitos do problema; e (ii) elaborar modelo lógico que detalha entradas, atividades, produtos, impactos e resultados.

142. Em relação ao primeiro item, vale salientar que, para uma melhor compreensão da política, o primeiro passo é delimitar qual oportunidade se quer aproveitar ou qual o problema se quer mitigar ou resolver com a intervenção da política proposta. Problema público pode ser definido como a diferença entre a situação existente (realidade) e a situação desejada.

143. O Guia Prático de Análise Ex Ante[1] do IPEA, em consonância ao Decreto 9.191/2017, art. 27, inciso I, sugere que a caracterização do problema/oportunidade pode ser decomposta da seguinte maneira:

1) Identificação do problema/oportunidade.
2) Causas potenciais dos problemas/benefícios das oportunidades.
3) Dados quantitativos acerca do problema/oportunidade:
    a) evidências na realidade brasileira; e
    b) comparação internacional.

[1] O Guia Prático de Análise *Ex Ante* aborda apenas os problemas. Como uma tecnologia emergente abre a possibilidade para um país ganhar vantagem competitiva e capturar benefícios, o modelo proposto pelo referido guia foi adaptado para também incluir as oportunidades de inovação.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Figura 25 – Decomposição dos problemas/oportunidades. Fonte: Guia Prático de Análise *Ex Ante* (adaptado)

144. Após a criação de uma lista de problemas, pode-se utilizar um método específico, como Árvore de Problema, para verificar se as possíveis causas e consequências para a ocorrência do problema foram levantadas, no momento de sua identificação e se existe consistência lógica e coerência entre o problema, as causas e os efeitos apresentados. Essa árvore contém o problema central a ser atacado; as causas potenciais e o relacionamento entre elas e as consequências. A sua elaboração deixa a análise do problema mais didática e transparente, permitindo posteriormente à sociedade verificar se a política realmente ataca as causas centrais dos problemas identificados. Entende-se também que a construção de árvore similar pode ser utilizada para identificar os efeitos positivos do aproveitamento das oportunidades de inovação com o uso de inteligência artificial.

145. Ressalta-se que a seção de diagnóstico e dos eixos temáticos da EBIA e o documento produzido pela consultoria contratada pelo MCTI para elaboração da EBIA (peça 626) abordam sobre dados quantitativos dos problemas/oportunidades discutidos. Porém, devido à ausência do uso de metodologia que identificasse explicitamente os problemas e as oportunidades, não é possível dizer se os dados quantitativos coletados são adequados e suficientes.

146. Ainda sobre o primeiro item, as melhores práticas orientam que "a teoria de como a política incide sobre as causas do problema, projetando seus resultados e impactos de longo prazo" seja descrita da seguinte forma: se (descreva o desenho do programa), então (descreva os resultados), o que então levará a (descreva os impactos).

147. A segunda falha formal identificada na elaboração da EBIA é que não foi apresentado explicitamente um modelo lógico que demonstre como os recursos e atividades geram produtos, resultados e seus respectivos impactos.

148. O modelo lógico geralmente é composto por um fluxo, exibindo a racionalidade de conexão entre as atividades propostas através da política e os objetivos que se pretende atingir, conforme demonstrado na figura abaixo:

Figura 26 – Etapas do modelo lógico de uma política pública. Fonte: Guia Prático de Análise *Ex Ante*

149. Como se pode observar, essa é uma estrutura racional (passo a passo estruturado) que demonstra como recursos e ações mudam comportamentos, geram produtos e produzem resultados e impactos. Trata-se, portanto, de um importante fluxo que estabelece conexão entre os resultados pretendidos e os recursos e ações necessárias ao alcance daqueles.

150. Ao não apresentar o referido modelo, entende-se que a EBIA não apresentou transparência adequada quanto à identificação e definição de resultados esperados da política.

<u>Critérios</u>

a) Decreto 9.191/2017, art. 27, inciso I;

b) Avaliação de Políticas Públicas - Guia Prático de Análise *Ex Ante*, IPEA;

<u>Evidências</u>

a) Estratégia Brasileira de Inteligência Artificial (Portaria MCTI 4.979/2021).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

b) Documento produzido pela Consultoria contratada no Âmbito da elaboração da Estratégia Brasileira de IA (Peça 626).

Causas

a) Não identificadas.

Efeitos

a) Dificuldade de validar os materiais que embasaram a elaboração da EBIA pela sociedade;

b) Dificuldade de avaliar se as soluções propostas são adequadas aos problemas/oportunidades relacionados ao uso de IA;

c) Dificuldade de entender como a EBIA organiza os processos organizacionais e o fluxo de insumos e processos de maneira a gerar produtos, resultados e impactos;

a) Falta de clareza acerca dos produtos e efeitos esperados da política.

Proposta de Encaminhamento

a) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

Achado IV – *As estruturas de governança e gestão necessárias à implementação da política não estão formalmente institucionalizadas na EBIA*

Situação encontrada

151. Políticas lançadas sem uma clara definição de quem faz o quê, com que grau de autonomia, com que formas de monitoramento e controle – e sem muitas vezes especificar seu passo a passo operacional e sem planejamento de seus riscos – geram incertezas, brechas e pontos de travamento ("gargalos"). Sendo assim, para mitigar tais riscos, a política pública deve vir acompanhada pela criação de instâncias de gestão e governança.

152. De acordo com o Guia Prático de Análise *Ex Ante,* a gestão diz respeito à maneira como os atores determinam e operam os insumos e processos que geram produtos, de forma a alcançar resultados positivos. Ela também se debruça sobre os aspectos que foram especificados no modelo lógico, organizando soluções técnicas e tecnológicas capazes de garantir as entregas dos produtos com qualidade e no tempo esperado.

153. Já os modelos de governança definem quem comanda, quem coordena, quem acompanha (supervisiona) e quem gera a política. A governança de uma política define as competências e as atribuições de sua execução e a responsabilidade por seus ajustes. Um modelo estabelece quem pode o quê (competências) e quem faz o quê (atribuições).

154. Importante ressaltar que parte significativa das definições sobre a gestão e a governança devem vir previamente institucionalizadas, ou seja, inscritas nos próprios normativos dos programas, contendo, no mínimo, conforme previsto no Decreto 9.203/2017, informações sobre: (i) se a atuação governamental conjunta é formalmente definida; (ii) se os papéis e responsabilidades das partes interessadas estão claramente definidos; e (iii) se os processos decisórios da atuação conjunta estão definidos.

155. Por essa perspectiva, analisando o conteúdo da EBIA, verifica-se que não está estabelecida nenhuma estrutura de governança e gestão que satisfaça os requisitos acima. Quanto a essa ausência, vale dizer que a publicação da EBIA sem a definição de arranjos institucionais que condicionam a forma pela qual as políticas são formuladas, implementadas e avaliadas prejudica o funcionamento de instrumentos como a *accountability* (responsabilização e prestação de contas) e transparência da política pública.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

156. Nada obstante, de acordo com a resposta do MCTI ao questionário de autoavaliação (peça 627), tal estrutura de governança será construída mediante a elaboração de um regimento interno, disciplinando a competência do MCTI referente ao art. 2º, inciso II da Portaria 4.617/2021, que instituiu a Estratégia Brasileira de Inteligência Artificial e seus eixos temáticos:

> Art. 2º Caberá ao Ministério de Ciência, Tecnologia e Inovações, no âmbito de suas atribuições:
> I - criar instâncias e práticas de governança para priorizar, implantar, monitorar e atualizar as ações estratégicas estabelecidas na Estratégia Brasileira de Inteligência Artificial;

157. Sendo assim, entende-se que a elaboração do regimento supramencionado deverá envolver as estruturas, os processos, os mecanismos, os princípios, as regras e as normas que influenciam a governança e a gestão de uma política pública.

Critérios

a) Decreto 9.203/2017, art. 4º, inciso IV, VIII e X;

b) Avaliação de Políticas Públicas - Guia Prático de Análise *Ex Ante*, IPEA.

Evidências

a) Estratégia Brasileira de Inteligência Artificial (Portaria MCTI 4.979/2021).

Causas

a) Não identificadas.

Efeitos

a) Ausência de estrutura atuante de coordenação das políticas federais de fomento à inovação sob uma perspectiva integrada de governo;

b) Inoperância de instâncias responsáveis pela política;

c) Os atores governamentais envolvidos com a política pública podem não atuar de maneira coerente e coordenada entre si;

d) Ausência de mecanismos de avaliação e monitoramento de resultados.

Proposta de Encaminhamento

e) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

***Achado V** – A estrutura de monitoramento e avaliação da EBIA não está formalmente definida em norma*

Situação encontrada

158. Monitoramento é uma atividade gerencial interna, realizada sistematicamente durante o período de execução e operação, para se saber como uma intervenção evolui ao longo do tempo, através de dados da gerência da política pública sobre metas iniciais, indicadores e resultados associados aos programas. O sistema de monitoramento e avaliação, por sua vez, consiste no conjunto de estruturas, processos, tecnologias e recursos desenhados e operacionalizados por atores públicos e/ou privados, os quais são responsáveis por apresentar evidências quantificáveis do problema; de sua evolução, frente as intervenções; e de seus resultados e impactos ao longo do tempo.

159. Nada obstante, necessário consignar que, embora o MCTI tenha informado que tal estrutura será construída mediante um regimento interno (peça 627), o fato é que a EBIA não definiu formalmente em norma uma estrutura de monitoramento e avaliação, contrariando o disposto no Decreto 9.203/2017, art. 4º, inciso III.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

160. Prever a estrutura de monitoramento no corpo do documento que instituí a política é necessário, pois a avaliação das políticas públicas deve assegurar o aprendizado e aperfeiçoamento contínuos, e, também, criar condições para que haja *accountability*, envolvendo as dimensões de transparência da ação pública e responsabilização perante a sociedade.

161. Considera-se boa prática que complementa o sistema de monitoramento: (i) o estabelecimento de formas de revisão, fixando a periodicidade da avaliação de desempenho da política; (ii) a definição de uma cadeia de indicadores de: insumos/recursos, processos, produtos, resultados e impactos para acompanhar o desempenho da política; e (iii) o estabelecimento de um cronograma, ou marcos temporais, de verificação do desempenho da política.

162. Além disso, é importante verificar se o sistema de monitoramento prevê a elaboração de relatórios de acompanhamento de implementação e progresso, a fim de se constatar se foi efetivamente implementado um plano de monitoramento e avaliação para acompanhar o progresso e o alcance das metas e objetivos estabelecidos para a política pública. Nesse contexto, merece registro a iniciativa do Governo Alemão, que após dois anos da publicação de sua estratégia nacional de IA em 2018, publicou um documento de acompanhamento e atualização da estratégia, em 2020[1].

163. Por fim, pressupõe-se que o monitoramento da política deve ocorrer até o seu término. Sobre este assunto, importa dizer que a EBIA não estabeleceu um prazo final para implementação de suas ações. Dessa maneira, entende-se que a ausência de uma data final para cumprimento das ações e iniciativas previstas, somada à ausência de estrutura de monitoramento implica na dificuldade em compreender os resultados pretendidos com a implementação da EBIA e a identificação de interdependências e obstáculos, o que deveria ter sido evitado pelo MCTI.

Critérios

a) Decreto 9.203/2017, art. 4º, inciso III.

Evidências

a) Estratégia Brasileira de Inteligência Artificial (Portaria MCTI 4.979/2021).

Causas

a) Não identificadas.

Efeitos

a) Impossibilidade de verificar se a política pública opera de forma consistente, em conformidade com o que foi definido durante sua estruturação;

b) Impossibilidade de verificar se o desempenho é satisfatório, em termos de eficiência e alcance de objetivos;

c) Impossibilidade de verificar se há problemas de formulação e/ou implementação que precisam ser tempestivamente tratados.

Proposta de Encaminhamento

d) Não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020.

5 Inteligência Artificial no Controle

5.1 Utilização de IA na atividade de controle

[1] Disponível em https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

164. As tecnologias de IA, em especial as referentes à aprendizagem de máquina, já vêm sendo utilizadas e proporcionando maior produtividade e precisão às atividades de controle. Historicamente, a auditoria dependia de amostragem para identificar os riscos. Essa abordagem não leva em consideração a maioria das informações disponíveis e depende do processamento manual. Com o aprendizado de máquina, os auditores podem tratar e analisar grandes volumes de dados para identificar anomalias e transações de alto risco que podem ser posteriormente analisadas mais detalhadamente.

165. Conforme argumenta a Intosai, a IA pode facilitar o processo de auditoria pública, pois ajuda a realizar muitas tarefas demoradas com mais eficiência. Nesse contexto, a IA pode ser aplicada em áreas como[1]:

a) Automação de processos robóticos (RPA): muitas tarefas de auditoria iterativa podem ser executadas com mais eficiência usando a RPA. Depois que os dados das entidades auditadas são carregados no sistema do órgão fiscalizador, a RPA pode identificar inconsistências e *outliers* que os auditores humanos podem então resolver;

b) Ferramentas de otimização de pesquisa: os algoritmos de IA podem ser usados para classificar e agrupar entidades, enquanto um mecanismo de risco projetado por IA pode calcular uma pontuação na natureza crescente do risco;

c) Rede Neural Artificial: as redes neurais artificiais reconhecem e memorizam dados ou padrões de transação, de forma que com base em problemas ou observações identificadas em auditorias anteriores, podem-se detectar casos semelhantes;

d) Extração de informações e mineração de dados: a IA pode ser usada para coletar, agrupar, extrair e analisar receitas e despesas do governo armazenadas em vários sistemas. Os algoritmos também podem confrontar dados de importação e exportação de sistemas integrados para identificar anomalias rapidamente;

e) Processamento de linguagem natural: esses algoritmos podem automatizar as classificações de observação de auditoria com base em dados históricos para aprender critérios e aplicar regras semelhantes para automação de tarefas.

166. Dessa forma, o uso de tecnologias de IA permite que o auditor se concentre nas áreas de maior risco na auditoria, reduzindo os custos envolvidos na atividade e indo ao encontro do momento atual por que passam os órgãos de controle da Administração Pública no tocante à diminuição de seus quadros de colaboradores.

167. Nesse contexto, o TCU vem desenvolvendo soluções que visam apoiar as diversas atividades de auditoria como os sistemas Alice (Análise de Licitações e Editais), Monica (Monitoramento Integrado para o Controle de Aquisições), Adele (Análise de Disputa em Licitações Eletrônicas), Sofia (Sistema de Orientação sobre Fatos e Indícios para o Auditor), Carina (Crawler e Analisador de Registros da Imprensa Nacional), e Ágata (Aplicação Geradora de Análise Textual com Aprendizado).

168. O Apêndice XI apresenta o detalhamento de algumas dessas soluções, bem como iniciativas de outros órgãos de controle na utilização de IA no suporte às suas atividades fim.

5.2 Soluções de IA para a atividade de controle

169. Se por um lado a inteligência artificial tem o potencial de aumentar os ganhos em eficiência, melhorar serviços e reduzir custos, por outro, ela também apresenta novos desafios e riscos, conforme apresentado na seção 2.8 – Considerações sobre o uso de IA. Tais desafios envolvem questões como a

[1] Conteúdo disponível em http://intosaijournal.org/artificial-intelligence-preparing-for-the-future-of-audit/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

segurança de dados, a possibilidade de tratamento automatizado desigual e a produção em massa de decisões institucionais incorretas e/ou discriminatórias.

170. Nesse contexto, é necessário o uso de controles pelos gestores públicos para garantir a confiabilidade dos sistemas de IA, assim como é necessário que os órgãos de controle auditem as aplicações baseadas em algoritmos de IA. Dessa forma, várias entidades que atuam no segmento de controle já desenvolveram propostas e frameworks que visam endereçar as questões relativas à asseguração da utilização ética e segura da IA.

171. Nesse sentido, dado o potencial de desenvolvimento e implementação de soluções de IA pela APF, é importante que o corpo técnico do TCU esteja devidamente capacitado para o adequado acompanhamento e fiscalização das referidas ações. Assim, entende-se que seja necessário a elaboração de um programa de capacitação específico sobre o tema, o qual poderá ter como base os diversos referenciais já existentes, a seguir descritos.

5.2.1 GAO - An Accountability Framework for Federal Agencies and Other Entities

172. A Instituição Superior de Auditoria dos Estados Unidos (*U.S. Government Accountability Office* - GAO) publicou um framework para governança em Inteligência Artificial intitulado "Artificial Intelligence: *An Accountability Framework for Federal Agencies and Other Entities*"[1]. O documento apresenta uma estrutura de responsabilidade de IA projetada para garantir a prestação de contas e o uso responsável da IA em processos de negócio do governo.

173. A publicação define uma estrutura que (i) serve como um guia prático para avaliar e auditar sistemas de IA; (ii) é flexível o suficiente para se adaptar às tecnologias em evolução; (iii) permite que não especialistas auditem sistemas de IA; e (iv) estabelece um *benchmark* para avaliar a segurança, justiça e eficácia em sistemas de IA implantados pelo governo.

174. A estrutura é organizada em torno de quatro princípios complementares que tratam sobre governança, dados, desempenho e monitoramento. Cada um destes componentes inclui itens de verificação, procedimentos de auditoria e os tipos de evidências a serem coletadas. Também são descritas as principais práticas-chaves que as entidades de governo devem assegurar na implementação e utilização dos sistemas de IA. Cada prática inclui um conjunto de questões e controles que devem ser verificados pelos auditores.

[1] Disponível em https://www.gao.gov/assets/gao-21-519sp.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

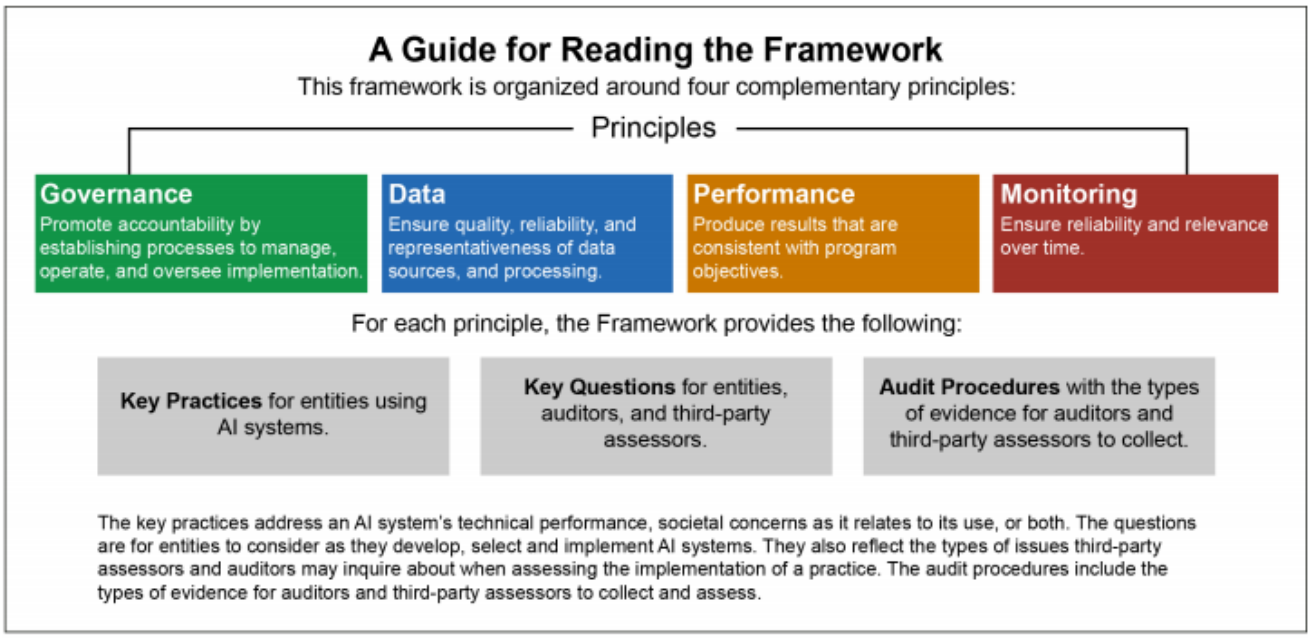


Figura 27 – Detalhamento da estrutura do Framework do GAO

5.2.2 ISSAI - Auditing machine learning algorithms

175. Foi publicado no âmbito do *The International Standards of Supreme Audit Institutions* (ISSAI) um artigo[1] que propõe uma abordagem metodológica à realização de auditoria pelas Instituições Superiores de Auditoria (SAIs) sobre algoritmos de aprendizado de máquina desenvolvidos pelas entidades governamentais. O objetivo é ajudar as SAIs a auditarem aplicações de IA, abrangendo os riscos relacionados ao uso de modelos de aprendizagem de máquina, além de sugerir possíveis testes para obter evidências.

176. Ressalta-se que o modelo proposto requer conhecimento específico dos auditores em ciência de dados e métodos quantitativos. Além disso, é desejável que o auditor tenha entendimento básico sobre linguagem de programação e infraestrutura de TI. Essa limitação ao público-alvo é em razão do artigo focar em questões que são específicas para aplicações de IA, como a transparência, a autonomia, a igualdade e a justiça das decisões de um algoritmo de ML.

177. O *framework* é estruturado em cinco áreas de auditoria e baseado no modelo de processo de mineração de dados amplamente conhecido como *Cross Industry Standard Process for Data Mining* (CRISP-DM)[2], que inclui todas as fases do ciclo de vida de uma aplicação de IA - desde a compreensão do negócio até a implantação e operação contínua.

178. O modelo vem acompanhado ainda de uma lista de verificação de auditoria , a qual resume os requisitos mínimos que uma organização auditada deve atender. A lista inclui controles e testes de auditoria sugeridos para mitigar possíveis riscos. Registra-se que o modelo pode ser utilizado de forma flexível, uma vez que os auditores, com base em sua avaliação de risco, podem selecionar a(s) etapa(s) que querem auditar ao longo do ciclo CRISP-DM.

5.2.3 ICO - Guidance on the AI auditing framework

179. O *Information Commissioner's Office* (ICO), autoridade independente do Reino Unido criada para proteger os dados e informações no que tange ao interesse público, criou um *framework* de

[1] Disponível em https://www.auditingalgorithms.net/

[2] CRISP-DM: https://pt.wikipedia.org/wiki/Cross_Industry_Standard_Process_for_Data_Mining

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

auditoria para sistemas de IA. A premissa é que, embora a IA traga diversos benefícios para organizações e indivíduos, ela também apresenta diversos riscos que podem ameaçar o direito à privacidade dos cidadãos.

180. O *framework* tem como principal objetivo garantir que a solução de IA (própria ou de terceiros) processe os dados pessoais de maneira lícita, justa e transparente e é destinado tanto aos especialistas em TI (especialistas em aprendizado de máquina, cientistas de dados, desenvolvedores e engenheiros de software) como aos gestores que lidam com a conformidade dos dados.

181. O *framework* tem como premissa que um gestor deve: (i) ser responsável pela conformidade do sistema; (ii) avaliar o impacto de proteção aos dados e mitigar os riscos; e (ii) documentar e demonstrar as escolhas que foram feitas no projeto.

182. O modelo proposto pelo ICO subdivide-se em duas grandes áreas: i) governança e accountability; e ii) áreas de riscos específicas de IA. Os riscos de IA são tratados em oito áreas específicas:

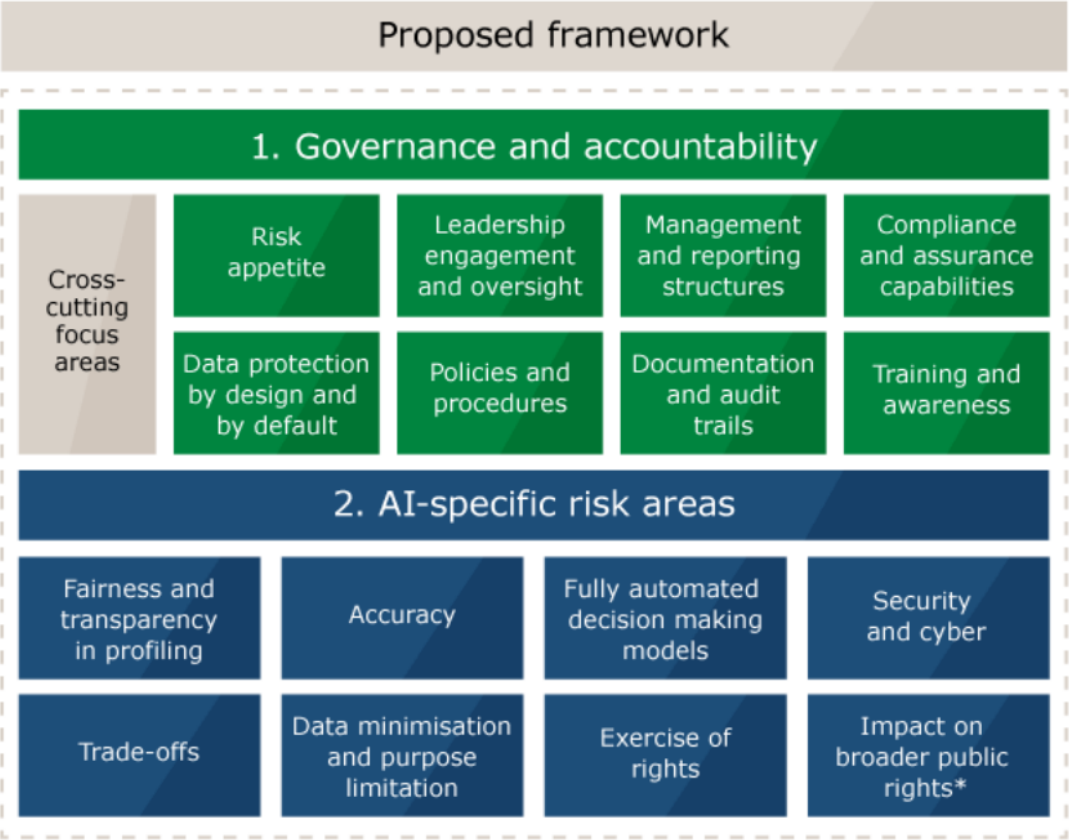


Figura 28 - Detalhamento da estrutura do Framework do ICO

183. Para cada área de risco, é descrito: (i) os riscos envolvidos; (ii) como a IA pode aumentar sua probabilidade e / ou impacto; e (iii) algumas medidas possíveis que a organização pode usar para identificar, avaliar, minimizar, monitorar e controlar esses riscos.

184. É ainda disponibilizado um *toolkit* projetado para fornecer suporte prático para auditoria, que consiste na avaliação de 61 controles relacionados às oito áreas de riscos e às fases de desenvolvimento de soluções de IA (https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/about-this-guidance).

5.2.4 Comissão Europeia - Assessment List for Trustworthy Artificial Intelligence

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

185. Na mesma linha do ICO, a Comissão Europeia, por meio do *High-Level Expert Group on Artificial Intelligence* (AI HLEG) publicou o documento "*Assessment List for Trustworthy Artificial Intelligence – (ALTAI)*"[1] com o propósito de auxiliar as organizações a avaliarem se o sistema de IA que está sendo desenvolvido, implantado, adquirido ou usado cumpre os sete requisitos de Inteligência Artificial Confiável, conforme especificado pelas diretrizes de ética da Comissão Europeia.

186. O conceito de IA confiável foi introduzido nas Diretrizes de ética para Inteligência Artificial Confiável (IA) e é baseado em sete requisitos principais: i) Agência Humana e Supervisão; ii) Robustez técnica e segurança; iii) Privacidade e governança de dados; iv) Transparência; v) Diversidade, não discriminação e justiça; vi) Bem-estar ambiental e social; e vii) Prestação de contas.

187. Os princípios de IA são traduzidos em uma lista de verificação dinâmica por meio de um questionário de autoavaliação, que orienta, na prática, aos desenvolvedores de IA a cumprirem os princípios contidos nos sete requisitos básicos.

188. Cumpre observar que a lista de avaliação para IA confiável pode ser utilizada de forma flexível, uma vez que as organizações podem escolher os elementos relevantes para o sistema de IA específico ou adicionar elementos que acharem relevantes.

189. O guia de autoavaliação é destinado a equipes multidisciplinares, envolvendo especialista em IA, cientistas de dados, usuários do sistema, área de negócio e controle interno. Assim, a ALTAI auxilia as organizações a entenderem o que é IA confiável, em particular quais riscos um sistema de IA pode gerar e como minimizar esses riscos e, ao mesmo tempo, maximizar os benefícios da IA.

5.2.5 Transparência Brasil - Estrutura de avaliação de riscos a direitos e de transparência

190. A organização Transparência Brasil elaborou uma "Estrutura de avaliação de riscos a direitos e de transparência"[2] para o uso de IA pelo poder público, o qual permite entender a transparência necessária em processos de desenvolvimento, aquisição e implementação de IA pelo estado, de forma a garantir o devido controle social democrático.

191. O modelo funciona como um guia para encontrar pontos críticos e, a partir disso, elaborar recomendações ao governo, exigir mais transparência pública, correções ou testes para garantir não-discriminação e erros algorítmicos, ou até a eventual descontinuidade de alguma ferramenta cujo risco seja incontornável.

192. A estrutura de avaliação é dividida da seguinte forma: (i) riscos a direitos pela natureza da ferramenta; (ii) riscos a direitos por discriminação algorítmica; (iii) riscos ao direito à privacidade; (iv) potencial abuso autoritário do espaço cívico; e (v) transparência no uso de sistemas de inteligência artificial. Cada uma dessas áreas tem itens de verificação associados. Assim, avalia-se o nível de transparência existente como instrumento necessário para acompanhar todo o processo de utilização pública de sistemas de IA, tendo em vista os riscos a direitos elencados:

Tabela 6 - Avaliações a serem realizadas

I - No fluxo de utilização da ferramenta, há supervisão humana em todas as decisões sugeridas ou tomadas pelo algoritmo?	VIII - Existem evidências de que este algoritmo funcione no ambiente em que ele está sendo utilizado? As evidências são baseadas em experimentos científicos relevantes?

[1] Disponível em https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

[2] Disponível em https://www.transparencia.org.br/downloads/publicacoes/Estrutura_Avaliacao_Risco.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

II - Em caso de erro do algoritmo corrigido por humano, essa informação é usada para aprimoramento do algoritmo?	IX - Existe regulamentação específica sobre o uso deste algoritmo na área em que ele está sendo aplicado? Quais são? Se não, existe apoio de uma equipe jurídica especializada para garantir que haja respaldo jurídico?
III - A ferramenta, por sua natureza, pode impactar direitos fundamentais, seja por erro ou por design do seu algoritmo, seja direta ou indiretamente? Se sim, quais?	X - Existe uma equipe técnica que acompanha e monitora a implementação deste algoritmo? Esta equipe contém funcionários do órgão capazes de analisar criticamente os caminhos tomados?
IV - Quais grupos ou populações serão afetadas por esse algoritmo? Esses segmentos foram considerados no processo de treinamento da ferramenta?	XI - A equipe responsável pelo desenvolvimento do algoritmo inclui especialistas da área na qual o algoritmo será aplicado?
V - Existem órgãos ou pessoas dentro da entidade que utilizam o algoritmo que podem prestar informações sobre seu uso às autoridades competentes?	XII - Um comitê de ética acompanha/acompanhou o desenvolvimento do algoritmo e os ritos envolvidos na coleta e uso de dados?
VI - O impacto negativo é criado ou acentuado a partir do algoritmo?	
VII - Esse algoritmo é imprescindível para atingir o objetivo apontado? Se ele tem o potencial de afetar o exercício de direitos fundamentais ou de se colocar como intermediário para acesso a eles, existem formas alternativas para exercício de tal direito? Se sim, quais?	

5.2.6 **LIAA-3R - Diretrizes de auditabilidade e conformidade no desenvolvimento e testes de soluções de IA**

193. O Laboratório de Inteligência Artificial Aplicada do Tribunal Regional Federal da 3ª Região (LIAA-3R), com o objetivo de incentivar a pesquisa e contribuir para o aprimoramento dos serviços judiciais e administrativos, elaborou o documento[1] que consolida as diretrizes de auditabilidade e conformidade, para orientar o desenvolvimento de modelos de inteligência artificial.

194. O documento consolida os critérios legais que devem ser seguidos no âmbito do LIAA-3R e é organizado em três partes. Na primeira delas, intitulada "Diretrizes Gerais", procurou-se seguir a divisão de matérias e os preceitos da Resolução CNJ 332/2020, da Portaria do CNJ 271/2020, assim como do documento intitulado *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*.

195. Na segunda parte, intitulada "Diretrizes Específicas", abordam-se algumas das regras aplicáveis especificamente à Justiça Federal da 3ª Região e a seus laboratórios de inovação. A terceira e última parte cuida de algumas diretrizes relacionadas à observância da LGPD.

196. Além disso, o documento apresenta duas listas de verificação, na forma de questões técnicas, destinadas à equipe de desenvolvimento e ao Grupo de Validação Ética e Jurídica do LIAA-3R (GVE), com a finalidade de avaliar se requisitos específicos foram observados, os quais envolvem as seguintes áreas:

[1] Disponível em https://www.trf3.jus.br/documentos/adeg/Inova/LIAA-3R/Aquivos/Manual_de_Conformidade_-_final_13.abril.pdf

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Tabela 7 – Diretrizes LIAA-3R – Questões Técnicas

Equipe de desenvolvimento	GVE
1. Questões relacionadas ao escopo e à finalidade do projeto	1. Ação e supervisão humanas
2. Questões relacionadas aos usuários e ao contexto de uso:	2. Solidez técnica e segurança
3. Questões relacionadas aos *datasets* utilizados	3. Privacidade e gestão de dados
4. Questões relacionadas à arquitetura e tecnologias adotadas	4. Transparência
5. Questões relacionadas aos testes da solução	5. Diversidade, não discriminação e equidade (*fairness*)
	6. Bem-estar social e ambiental
	7. Prestação de Contas (*Accountability*)
	8. Questões relacionadas às aprovações obtidas
	9. Questões relacionadas a possíveis conflitos de interesses
	10. Questões relacionadas à equipe do projeto

6 Conclusão

197. O presente levantamento foi proposto com o objetivo geral de conhecer a atual utilização das tecnologias de Inteligência Artificial na Administração Pública Brasileira, avaliar os riscos e oportunidades da Estratégia Brasileira de Inteligência Artificial e verificar os principais aspectos referentes ao uso de tecnologias de IA nas atividades de controle.

198. Nesse sentido, o trabalho buscou inicialmente descrever na Seção 2 – Visão Geral do Objeto os principais conceitos, tecnologias envolvidas, domínios de aplicação da tecnologia bem como considerações que devem ser referentes à ética, segurança e privacidade dos dados quando do desenvolvimento e implementação de soluções de IA.

199. Com relação à utilização de IA na Administração Pública Federal, a Seção 3 abrangeu os principais benefícios na utilização de soluções de IA no Setor Público, entre eles a redução de gastos em decorrência da maior eficiência operacional, o aprimoramento dos serviços disponíveis ao cidadão e a disponibilização de novos serviços, além das vantagens da utilização dos dados para tomadas de decisão pelos entes governamentais.

200. Nesse contexto, verificou-se em estudo que no Brasil, apesar do país apresentar diferenciais competitivos que incentivam a maior implementação e soluções de IA, como a grande disponibilidade e a representatividade dos dados, o Setor Público brasileiro ainda carece de melhores políticas e investimentos no sentido de aprimorar sua infraestrutura, capital humano e capacidade de inovação.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

201. Com o objetivo de se obter um panorama abrangente do estado atual de utilização da IA nas organizações da APF, realizou-se pesquisa por meio de questionário eletrônico que foi respondido por 263 organizações. Entre as principais constatações a partir da análise das respostas, tem-se que grande parte das instituições ainda não planejam utilizar soluções de IA em seus processos produtivos ou na disponibilização de serviços aos cidadãos. Entre os principais dificultadores informados para a adoção da tecnologia, foi reportado o baixo número de colaboradores disponíveis com capacitação suficiente para a implementação.

202. A pesquisa também evidenciou que a maior parte das soluções são desenvolvidas por equipe própria dos órgãos, sendo que 48% das organizações pesquisadas não realizaram nenhuma capacitação interna em relação ao desenvolvimento e implementação de tecnologias de IA.

203. Verificou-se também que o Poder Judiciário está à frente em relação à utilização de soluções de IA, que tem beneficiado e contribuído principalmente na diminuição do estoque processual daquele Poder. Também estão sendo implementadas pelo Poder Judiciário estratégias e infraestruturas que permitem o compartilhamento de dados e o intercâmbio de soluções entre os diversos órgãos que o compõem.

204. Entre os principais benefícios apontados com a utilização de IA, as organizações pesquisadas informaram a automação de processos repetitivos com aumento de produtividade, a redução do tempo e custos na execução dos processos, a otimização na alocação de recursos humanos, o aumento da confiabilidade e transparência dos processos e maior tempestividade na resposta aos clientes.

205. Em relação à Estratégia Brasileira de Inteligência Artificial (EBIA), apresentou-se como algumas nações têm tratado o tema, tanto em nível nacional como em nível regional, com o objetivo de se tornarem líderes no segmento.

206. Nesse sentido, a avaliação realizada da EBIA, realizada tendo como base o Referencial para Controle de Políticas Públicas elaborado pelo TCU, evidenciou falhas em sua elaboração que podem comprometer a sua adequada implementação, monitoramento e atingimento dos objetivos. Foram identificados os seguintes achados em relação à EBIA:

a) os objetivos da EBIA não são específicos, mensuráveis, realistas (alcançáveis) e delimitados em um recorte temporal;

b) inexistência da explicitação do estágio de referência inicial ("marco zero", "situação inicial") e da definição de indicadores e metas para subsidiar a avaliação do desempenho e resultados da EBIA;

c) falhas formais na apresentação do modelo lógico e sobre como a EBIA incide sobre os problemas/oportunidades, projetando seus resultados e impactos de longo prazo;

d) as estruturas de governança e gestão necessárias à implementação da política não estão formalmente institucionalizadas na EBIA;

e) a estrutura de monitoramento e avaliação da EBIA não está formalmente definida em norma.

207. Cabe ressaltar que, apesar do disposto nos achados, não serão feitas recomendações em função da restrição imposta pelo item 21.3 da Portaria-Segecex 5/2021 (Roteiro de Levantamento) c/c art. 13 da Resolução TCU 315/2020. Entretanto, com o objetivo de contribuir para o aperfeiçoamento da EBIA, as informações contidas nos achados devem ser encaminhadas ao MCTI, órgão gestor da referida política pública.

208. O Levantamento abordou a utilização a utilização de IA na atividade de controle, identificando casos de uso da solução nas organizações da Administração Pública.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

209. O trabalho abordou ainda o controle de soluções de IA pelo Poder Público, abordando os principais referenciais já produzidos a respeito. Em decorrência das várias características específicas que as soluções de IA apresentam em relação aos demais sistemas computacionais, entende-se que deve ser elaborado plano de capacitação específico sobre o tema para o corpo técnico do TCU, em especial da Secretaria de Fiscalização de Tecnologia da Informação.

210. Por fim, de forma a contribuir para o aprimoramento da IA no Brasil, propõe-se encaminhar o presente relatório ao MCTI, para que este, caso julgue necessário, adote as providências pertinentes.

7 Proposta de Encaminhamento

211. Ante todo o exposto, propõe-se o encaminhamento dos autos ao gabinete do Exmo. Relator Aroldo Cedraz com as seguintes propostas:

a) levantar o sigilo dos autos, por conter informações relevantes à Administração Pública;

b) autorizar a Secretaria de Fiscalização de Tecnologia da Informação (Sefti) a divulgar as informações consolidadas constantes deste levantamento, preferencialmente por intermédio de fichas síntese, sumário executivo e infográfico;

c) encaminhar ao Ministério da Ciência, Tecnologia e Inovações (MCTI), à Comissão de Ciência e Tecnologia, Comunicação e Informática da Câmara dos Deputados, à Comissão de Ciência, Tecnologia, Inovação, Comunicação, Inovação e Informática do Senado Federal, bem como aos órgãos discriminados no Apêndice IX do presente relatório, cópia do acórdão que vier a ser proferido, bem como cópia do relatório da unidade técnica e seus respectivos apêndices, destacando que o relatório e o voto que fundamentaram a deliberação ora encaminhada podem ser acessados por meio do endereço eletrônico www.tcu.gov.br/acordaos;

d) arquivar o presente processo, com fulcro no art. 169, inciso V, do Regimento Interno do TCU.

Sefti, 30 de setembro de 2021.

(Assinado eletronicamente)
Diego Oliveira Farias
AUFC – Matr. 10642-9

(Assinado eletronicamente)
Eric Hans Messias
AUFC – Matr. 9652-0

(Assinado eletronicamente)
Erick Muzart Fonseca dos Santos
AUFC – Matr. 7.666-0

(Assinado eletronicamente)
Monique Louise de Barros Monteiro
AUFC – Matr. 8677-0

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

(Assinado eletronicamente)
Tibério César Jocundo Loureiro
AUFC – Matr. 6520-0

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE I – Principais algoritmos de Aprendizagem de Máquina

Aprendizagem Supervisionada

Regressão Linear	Os algoritmos de regressão linear mostram ou preveem a relação entre duas variáveis ou dois fatores ajustando uma linha reta contínua aos dados. A linha geralmente é calculada com a função Custo de Erro Quadrado. A regressão linear é um dos tipos mais populares de análise de regressão.
Regressão Logística	Os algoritmos de regressão logística ajustam uma curva contínua em forma de S aos dados. A regressão logística é outro tipo popular de análise de regressão, embora seja utilizada para tarefas de classificação.
K-Vizinhos mais próximos (K-Nearest Neighbors)	Os algoritmos de K-vizinho mais próximos armazenam todos os pontos de dados disponíveis e classificam cada novo ponto de dados com base nos pontos de dados mais próximos a eles, conforme medido por uma função de distância.
Máquinas de Vetores de Suporte (SVMs)	Os algoritmos de máquinas de vetores de suporte desenham um hiperplano entre os dois pontos de dados mais próximos. Isso diferencia as classes e maximiza as distâncias entre elas para diferenciá-las mais claramente.
Árvores de Decisão (Decision Trees)	Os algoritmos de árvore de decisão dividem os dados em dois ou mais conjuntos homogêneos. Eles usam as regras se–então para separar os dados com base no diferenciador mais significativo entre os pontos de dados.
Florestas Aleatórias (Random Forests)	Os algoritmos de floresta aleatória se baseiam nas árvores de decisão, mas em vez de criar uma árvore, eles criam uma floresta de árvores. Em seguida, agregam votos de diferentes formações aleatórias das árvores de decisão para determinar a classe final do objeto de teste.
Redes Neurais	

Aprendizagem Não Supervisionada

K-Means	Os algoritmos K-means classificam os dados em grupos (clusters) – em que K é igual ao número de clusters. Os pontos de dados dentro de cada cluster são homogêneos e são heterogêneos para pontos de dados em outros clusters.
DBSCAN	Com base em um conjunto de pontos, o DBSCAN agrupa pontos que estão próximos uns dos outros com base em uma medição de distância (geralmente distância euclidiana) e um número mínimo de pontos. Ele também marca como outliers os pontos que estão em regiões de baixa densidade.
Análise de Agrupamentos Hierárquicos (HCA)	Os algoritmos HCA usam uma estratégia que busca construir uma hierarquia de clusters que tenha uma ordenação estabelecida de cima para baixo.
Análise de Componente Principal (PCA)	Os algoritmos de análise de componente principal ou PCA são utilizados para redução de dimensionalidade (campos) de grandes arquivos de dados. Reduzir o número de campos torna mais simples explorar e visualizar grandes arquivos de dados.

Aprendizagem por Reforço

REINFORCE	Algoritmo clássico de 1992 que deu visibilidade à área de aprendizado por forço, REINFORCE é uma variante de Monte Carlo de um algoritmo de gradiente de política (policy) no aprendizado por reforço. O agente coleta amostras de um episódio usando sua política (policy) atual e as usa para atualizar o parâmetro θ da política. Uma vez que uma trajetória completa deve

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

	ser concluída para construir um espaço de amostra, ela é atualizada como um algoritmo off-policy.
SARSA	Sarsa usa diferença temporal para aprendizagem (*TD-learning*), que combina métodos de Monte Carlo, para aprender direto da experiência obtida pelo agente sem um modelo representativo das dinâmicas do ambiente, com o uso de programação dinâmica.
Deep Q-Learning (DQL)	O objetivo do Q-Learning é aprender uma política que diz a um agente que ação tomar sob determinadas circunstâncias (estados). O DQL estende e adapta a ideia por trás do Q-Learning para uso com redes neurais como aproximadores universais para representar problemas de alta dimensionalidade.
Rainbow	Várias extensões foram feitas aos algoritmos de DQL e o Rainbow combina 6 das técnicas mais proeminentes em um único algoritmo com ótimos resultados.
Deep Deterministic Policy Gradient (DDPG)	DDPG é um algoritmo off-policy, actor-critic, em que são combinadas duas redes neurais especializadas. Uma das redes neurais é o ator que prevê a ação a ser realizada e a outra rede, o crítico, avalia a performance do ator para indicar qual o melhor caminha de maximizar a recompensa.
Proximal Policy Optimization Algorithms (PPO)	PPO é um algoritmo on-policy de otimização de primeira ordem que busca políticas (policies) mais eficientes e compara com a corrente para troca por vários ciclos.
Asynchronous Advantage Actor-Critic (A3C)	O A3C é um algoritmo que usa múltiplos atores experimentando o ambiente em paralelo e, com isso, torna os dados em um processo mais estacionário e também torna a exploração dos estados mais diversa.
Twin Delayed Deep Deterministic policy gradient (TD3)	O TD3 baseia-se no algoritmo DDPG, com algumas modificações destinadas a lidar com o viés de superestimação com a função de valor. Em particular, ele utiliza clipped double Q-learning, atualização menos frequente de redes alvo (target) e política (policy) e suavização de política alvo (target policy) (semelhante a uma atualização baseada em SARSA. Uma atualização mais estável).
AlphaZero	AlphaZero é um algoritmo de aprendizado por reforço para jogos de tabuleiro como Go, xadrez e shogi. Ele ganhou grande visibilidade em 2017 ao atingir níveis sobre-humanos nós três jogos citados. Ele é um marco para o Go, pois é o primeiro algoritmo a conseguir vencer campeões mundiais do jogo. O AlphaZero trouxe enorme interesse da comunidade de IA para o aprendizado por reforço por usar IA para resolver problemas que seriam computacionalmente intratáveis de maneira exata com algoritmos tradicionais.
MuZero	O MuZero combina a busca utilizada no AlphaZero (Monte Carlo Tree Search – MCTS) com um modelo aprendido das dinâmicas do ambiente (environment). Ele atinge o nível sobre-humano em 51 dos 57 jogos do videogame Atari, que é muito usado para *benchmarking* por envolver tarefas variadas e com vários níveis de complexidade e fácil obtenção de recompensas (ex.: pontuação nos jogos).
Agent57	O Agent57 é um algoritmo bem recente, 2020, e é o primeiro algoritmo de aprendizado por reforço a atingir nível sobre-humano em todos os 57 jogos. Ele é *model-free*, ao contrário do MuZero, e possui um meta-controlador que possui um mecanismo adaptativo para escolher qual política (policy) deve fazer a exploração do espaço e qual deve fazer o *exploiting* além do desmembramento da recompensa entre extrínseca e intrínseca para melhor estabilidade do treinamento.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE II – Redes Neurais Artificiais[1]

0. Algoritmos tradicionais de aprendizagem de máquina dependem fortemente da representação dos dados para serem capazes de criar relacionamentos entre tais dados e as predições a que eles podem conduzir. Note-se, por exemplo, a diferença entre um sistema de diagnóstico que depende de informações de um paciente fornecidas pelo médico (ex.: IMC, tipo sanguíneo, glicemia) para ser capaz de propor um diagnóstico, de um sistema capaz de identificar tumores a partir de uma imagem radiográfica. Enquanto algoritmos tradicionais são capazes de extrair correlações entre o primeiro grupo de informações informadas pelo médico – aqui referidas como *features* ou atributos - e um possível diagnóstico, no segundo exemplo tais sistemas possuem limitações na análise de dados não estruturados como imagens, pois não são capazes de extrair significado a partir de tão somente um conjunto de pixels.

1. Uma solução para esse tipo de problema é o uso de técnicas que sejam capazes não somente de aprenderem relacionamentos entre os atributos e a resposta ou saída (predição), mas também de aprenderem a melhor forma de representar os dados de entrada. É nesse contexto que se insere a técnica de *deep learning*, que permite expressar representações complexas em termos de outras mais simples.

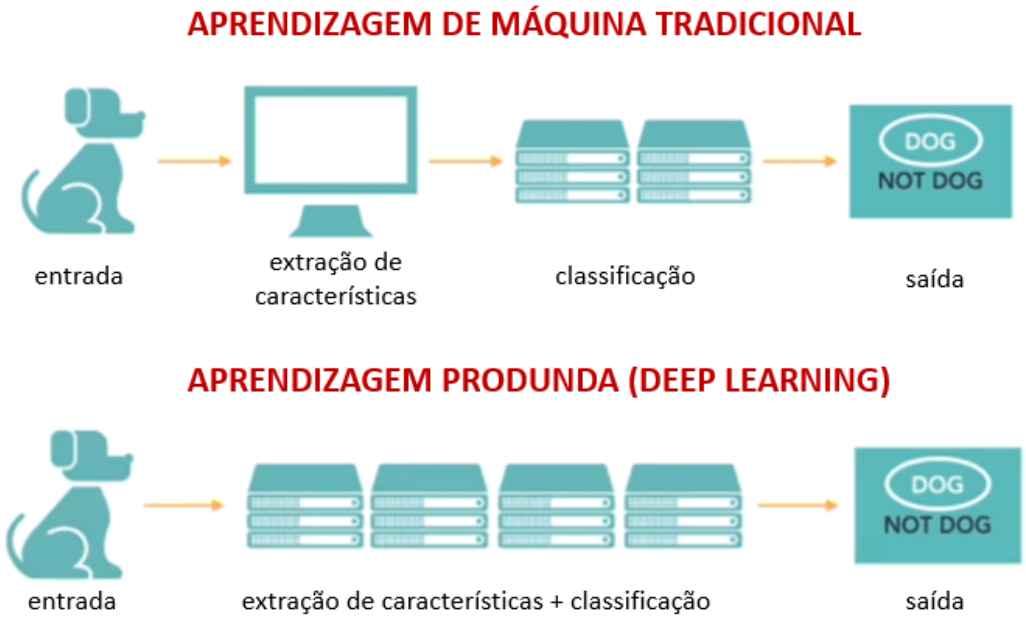


Figura 29 – Aprendizagem Profunda

2. *Deep Learning*, ou aprendizagem profunda, refere-se a uma classe de algoritmos que, baseada em modelos simplificados do funcionamento do cérebro, objetiva representar conceitos complexos em termos de conceitos mais simples, de forma hierárquica ou "multicamada". Cada "camada" de representação agrega conhecimento adquirido por camadas inferiores, que se especializam por sua vez em identificar os conceitos primitivos. Podemos imaginar, portanto, que quanto maior o número de camadas, ou melhor, quanto mais longa ou profunda for a sequência de camadas, maior a expressividade do modelo – daí a origem do termo *deep learning*.

3. Um exemplo desse tipo de entendimento hierárquico é a interpretação de uma imagem por um computador: as primeiras camadas de representação tendem a reconhecer padrões de pixels adjacentes – por exemplo, em termos de cores, para identificação de contornos. As camadas subsequentes tendem a especializar-se em reconhecer conceitos mais elaborados (ex.: vértices como

[1] O resumo apresentado nesta seção é baseado parcialmente em tradução livre do livro (Goodfellow, Bengio, & Courville, Deep Learning, 2016).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

conjuntos de contornos) e posteriormente formas geométricas (constituídas por sua vez por padrões específicos de contornos e vértices). Já as próximas camadas tendem a reconhecer conceitos mais complexos formados a partir de determinados padrões de formas geométricas (ex.: diferenciar um pneu de um prato, uma vez que ambos têm formato circular), aprendendo a reconhecer padrões de textura, por exemplo. E assim, sucessivamente, por meio de múltiplas camadas de reconhecimento de padrões, o algoritmo chega ao nível de identificar objetos em cenas ou classificar imagens.

4. Para dados textuais, é possível pensar em hierarquias como sequências de caracteres que formam palavras de um vocabulário. Tais palavras por sua vez formam sentenças, que em seguida formam textos com determinados padrões semânticos que permitem classificar o assunto ao qual o texto se refere, ou então identificar, por exemplo, se um determinado texto é uma crítica ou um elogio, se contém ironia etc.

5. O exemplo mais simples de um modelo de *deep learning* é a rede neural *multilayer perceptron* (MLP), uma função matemática composta por inúmeras funções mais simples, que processam sequencialmente os dados de entrada – sejam eles um arquivo de imagem ou um documento textual - de modo que o valor computado por uma função é utilizado como valor de entrada pela função subsequente, e assim sucessivamente. Cada uma dessas funções é denominada de camada de representação. A figura abaixo apresenta um diagrama de alto nível que explica as camadas de entrada, ocultas e de saída em uma MLP:

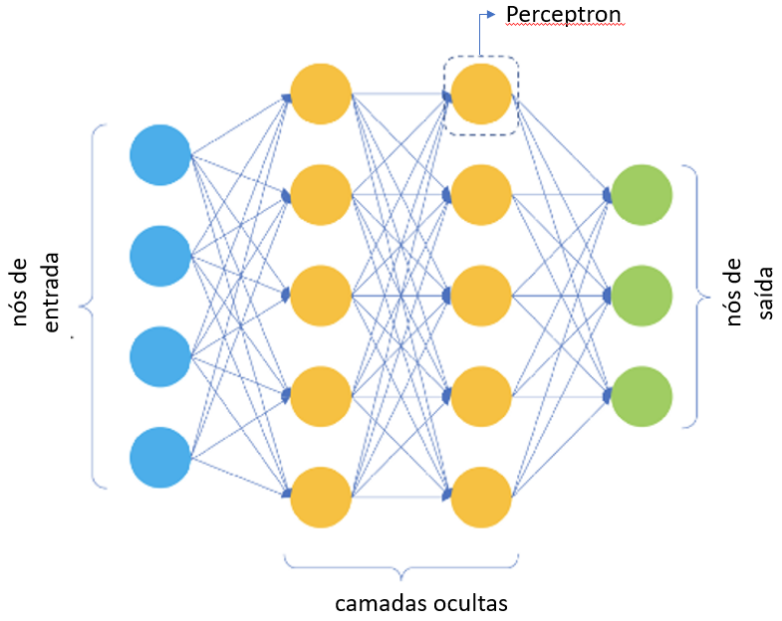


Figura 30 – Representação de uma rede neural *multilayer perceptron*

6. Por fim, é importante ressaltar que as redes neurais artificiais têm um histórico que remonta desde a década de 1940, com os primeiros modelos conceituais (McCulloch & Pitts, 1943), passando pelo algoritmo de retro propagação ou *backpropagation* (Rumelhart, Hinton, & Williams, 1986) - utilizado até os dias atuais – até chegar a meados do ano de 2006. Naquele ano, a terceira onda de pesquisa e uso de redes neurais foi impulsionada por inovações em termos de algoritmos e estratégias de treinamento, adaptáveis para diferentes arquiteturas de redes neurais e com maior capacidade não apenas de otimização, mas também de generalização.

7. O termo *deep learning* começa então a ser adotado, dada a importância teórica do maior número de camadas de unidades computacionais em redes neurais. Além disso, o maior poder computacional disponível na época veio a tornar viável o treinamento de redes neurais profundas. Um

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

exemplo é o advento do uso de unidades de processamento gráfico ou GPUs (*graphical processing units*), responsável por maior grau de paralelismo no processamento dos algoritmos de treinamento.

8. A inspiração biológica original – neurônios humanos – passou a ser cada vez mais substituída por inspiração advinda de outros campos, como álgebra linear, probabilidade, estatística, teoria da informação e otimização numérica. Além disso, o advento do Big Data, com um número cada vez maior de dispositivos computacionais produzindo grandes volumes de dados, ajudou a superar o maior obstáculo para a estatística estimativa: ser capaz de generalizar para novos dados depois de observar uma amostra. Assim, parte-se da premissa de que, aumentando a amostra, tem-se cada vez mais capacidade de aprendizado e generalização.

Tendências recentes em redes neurais

Transferência de Aprendizagem[1]

9. Transferência de aprendizagem (do inglês *transfer learning*) é uma técnica cada vez mais empregada, em especial nas áreas de visão computacional e processamento de linguagem natural, na qual o conhecimento adquirido por um modelo pré-treinado em um determinado domínio/conjunto de tarefas é "transferido" para outro domínio/conjunto de tarefas. Permite a "democratização" do uso de modelos de Inteligência Artificial, uma vez que novos modelos podem ser treinados com apenas uma fração da quantidade de dados e recursos computacionais que seriam utilizados caso um modelo tivesse que ser treinado "do zero".

10. A transferência de aprendizagem é ligeiramente inspirada na forma pela qual humanos aprendem, uma vez que dificilmente aprendemos algo do zero, mas comumente aprendemos por meio de analogia, incorporando a experiência adquirida anteriormente em problemas e situações similares a novos contextos.

Transferência de aprendizagem em processamento de linguagem natural (PLN)

11. Não restam dúvidas de que as arquiteturas e estratégias de treinamento de redes neurais adotadas nos últimos levaram a consideráveis avanços em tarefas como tradução de textos, respostas a perguntas (*question answering*) e *chatbots*, mesmo em tarefas treinadas do zero. Entretanto, mudanças significativas na distribuição amostral dos dados levavam à degradação de desempenho, indicando que os modelos haviam se especializado em desempenhar bem apenas com entradas de um determinado tipo (ex.: idiomas ou tipos de texto específicos).

12. Restavam desafios a serem superados para idiomas menos populares em comparação com o inglês, ou mesmo tarefas mais específicas ou ainda não exploradas. No caso concreto de idiomas, atente-se para o problema de línguas pouco faladas que, como consequência, contam com reduzida disponibilidade de corpus rotulado para treinamento de modelos de PLN.

13. Na década de 1960, um primeiro passo rumo à transferência de aprendizagem consistiu no uso de espaços vetoriais para representar palavras como vetores de números.

14. A metade da década de 2010 assistiu ao advento de modelos como word2vec (Mikolov, Chen, Corrado, & Dean, 2013), (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013), sent2vec (Pagliardini, Gupta, & Jaggi, 2018) (Gupta, Pagliardini, & Jaggi, 2019) e doc2vec (Le & Mikolov, 2014). Tais modelos eram treinados para representar, respectivamente, palavras, sentenças e documentos em espaços vetoriais de forma que a distância entre os vetores estivesse relacionada à diferença de significado entre as entidades correspondentes. O treinamento para que essa vetorização fosse obtida

[1] O resumo apresentado nesta seção é parcialmente baseado em (Azunre, 2021)

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

utilizava como premissa atrelar o significado de uma palavra ao seu contexto, ou seja, às palavras adjacentes no texto. Tem-se aqui portanto um exemplo de treinamento não supervisionado.

15. Uma vez que as palavras, sentenças ou parágrafos estejam representados como vetores, é possível utilizar, por exemplo, algoritmos de classificação ou *clustering*, para os quais os dados de entrada são representados tão somente como pontos em um espaço vetorial. No exemplo concreto de classificação, tem-se, portanto, uma abordagem semisupervisionada, uma vez que, apesar da tarefa de classificação ser supervisionada, a representação dos dados de entrada foi obtida de forma não supervisionada, sendo ainda assim capaz de embutir a semântica textual.

16. Posteriormente, para lidar com palavras não previstas no vocabulário inicial (ex.: palavras novas, gírias, *emojis*, estrangeirismos ou nomes de pessoas), passou-se a utilizar vetorização a nível de caracteres.

17. A descrição acima pode ser entendida como uma forma embrionária de transferência de aprendizagem, uma vez que o modelo pré-treinado de vetorização já é capaz de embutir um certo nível de semântica ou significado a palavras, sentenças etc.

18. O ano de 2018 assistiu a uma verdadeira revolução na área de PLN, quando pesquisadores começaram a empregar transferência de aprendizagem em um nível mais abstrato, disponibilizando não mais modelos de vetorização pré-treinados, porém redes neurais inteiras pré-treinadas em tarefas genéricas, não supervisionadas, e ainda assim de mais alto nível. Como exemplo, temos redes neurais que implementam modelos de linguagem, que nada mais são do que modelos estatísticos treinados para prever a próxima palavra ou conjunto de palavras dados os termos anteriores. Em um processo conhecido como fine-tuning, é possível, obtendo um desses modelos pré-treinados, apenas fazer breves treinamentos adicionais, voltados para otimizar o modelo para a tarefa específica que se deseja treinar, ajustando os pesos da rede. A esse movimento dá-se inclusive o nome do "momento ImageNet", em referência ao uso até então bastante difundido de redes neurais pré-treinadas na base de imagens ImageNet (Stanford Vision Lab, Stanford University, Princeton University, 2021) para aplicações das mais diversas em visão computacional.

19. A tabela a seguir mostra um resumo das principais inovações consideradas pioneiras na área de transferência de aprendizagem para PLN, e que são em sua maioria utilizadas atualmente. Como o objetivo aqui é tão somente motivar e fornecer uma visão geral do estado da arte na área, o detalhamento da arquitetura, princípios matemáticos e algoritmos de treinamento de cada um desses modelos está fora do escopo deste documento e pode ser consultado nas referências:

Embeddings from Language Models (ELMo) (Peters, et al., 2018)	Um problema da formulação original do word2vec era a desambiguação: um único vetor representava cada palavra de um vocabulário, mesmo que tal palavra pudesse ter diferentes significados a depender do contexto. O ELMo resolveu esse problema gerando vetores contextualizados, a partir do treinamento de um modelo de linguagem. É possível utilizar grandes bases de texto não supervisionadas como por exemplo Wikipedia para o treinamento do modelo de linguagem.
Universal Language Model Fine-tunning (ULMFit) (Howard & Ruder, 2018)	Modelo voltado para adaptar qualquer modelo de linguagem baseado em rede neural para qualquer tarefa específica, tendo sido incialmente demonstrado em tarefas de classificação textual. Utiliza como estratégia para *fine-tuning* o uso de

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

	diferentes taxas de aprendizagem para as diferentes camadas da rede.
OpenAI Generative Pretrained Transforer (GPT) (Radford & Narasimhan, Improving Language Understanding by Generative Pre-Training, 2018)	Baseado no modelo de rede neural denominado *Transformer* (Vaswani, et al., 2017), grosso modo uma arquitetura de rede que permite maior paralelismo e desempenho em relação a arquiteturas anteriores, que além de não suportarem o mesmo grau de paralelismo, apresentavam dificuldade para lidar com textos longos. Em sua formulação mais recente – o GPT-3 (Brown, et al., 2020) – é capaz de gerar automaticamente textos realistas, semelhantes aos que seriam escritos por humanos, além de suportar recursos como aprendizado multitarefa, *one-shot learning*[1] e *zero-shot learning*[2].
Bidirecional Endoder Representations from Transformers (BERT) (Devlin, Chang, Lee, & Toutanova, 2018)	Baseado no mesmo modelo de rede neural do GPT, porém com variações. Utiliza mascaramento de palavras a serem inferidas, adotando a acurácia dessa predição como métrica para o treinamento.

20. Para todos os modelos acima, foi demonstrado que era possível adaptá-los para tarefas mais especializadas, porém com reduzida quantidade de dados rotulados, não requerendo, portanto, alto custo em termos de base de textos para treinamento. Atualmente, muitos modelos de linguagem pré-treinados em vários idiomas estão disponíveis em repositórios, de forma aberta, para que possam ser livremente retreinados em tarefas específicas. Um exemplo desse tipo de repositório é o disponibilizado pela biblioteca HuggingFace (Huging Face, 2021), que disponibiliza milhares de modelos inclusive para tarefas especializadas como sumarização, classificação, geração de textos, tradução etc.

Generative Adversarial Networks

21. *Generative Adversarial Networks*, também conhecidas pelo acrônimo GANs ou, em tradução livre, redes generativas adversariais, referem-se a uma classe de arquitetura de redes neurais introduzida por um artigo de Ian Goodfellow e demais pesquisadores em 2014 (Goodfellow, et al., 2014).

22. Como o próprio nome dá a entender, modelos generativos especializam-se na geração de dados, sejam eles imagens, vídeos, textos, músicas ou mesmo discursos falados, de forma a assemelhar-se a dados criados por humanos, diferentemente de modelos discriminativos, que se especializam em classificar ou "compreender" informações (ex.: classificação de imagens ou detecção de objetos em aplicações de visão computacional, compreensão de textos em processamento de linguagem natural etc.).

23. Exemplos de aplicações de redes GANs têm sido noticiados pela mídia nos últimos anos, como por exemplo algoritmos capazes de gerar fotos de pessoas que não existem, ou fotos de ambientes (paisagens, cenários internos etc.) igualmente inexistentes (Karras, et al., 2020). Um exemplo de uso desvirtuado desse tipo de tecnologia inclui deep fake – uma prática que pode resultar, por exemplo, na criação de fotos para perfis falsos de redes sociais ("robôs"), montagens e manipulação de imagens fotográficas ou audiovisuais, geração de falsos discursos atribuídos a pessoas públicas etc. Por outro lado, bons usos de redes GANs incluem, por exemplo, a geração de obras de arte por artistas-programadores, e até mesmo usos na área de medicina, como a geração de imagens sintéticas para

[1] Técnica de treinamento na qual apenas um ponto de dado rotulado é necessário. Entradas – descritas em linguagem natural - são incluídas para instruir o modelo sobre qual tarefa deverá ser desempenhada.

[2] Técnica de treinamento na qual nenhum dado rotulado é necessário.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

aumentar bases de treinamento para algoritmos de diagnóstico a partir de imagens (Frid-Adar, Klang, Amitai, Goldberger, & Greenspan, 2018).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE III – Principais domínios de aplicação de IA

Predição

0. Um Modelo Preditivo é, de forma simplificada, uma função matemática que pode ser aplicada a uma grande quantidade de dados soltos, para que seja possível a identificação de padrões que possam mostrar tendências futuras. É como se fosse possível prever com eficiência o futuro, de forma matemática, com probabilidade, estatística etc.

1. O Modelo Preditivo usa dados, algoritmos estatísticos e técnicas de Aprendizagem de Máquina para identificar a probabilidade de resultados futuros, a partir de dados armazenados em um determinado histórico.

2. Ter um Modelo Preditivo bem-feito e calibrado faz com que riscos e oportunidades sejam identificados com antecedência suficiente para a tomada de decisão mais adequada. Os exemplos de uso são muitos, e vão desde a prevenção contra fraudes e otimização de campanhas de marketing até melhorias em processos operacionais.

3. O Modelo Preditivo, em última análise, serve para embasar decisões, que se tornam mais eficientes por serem realizadas de acordo com um cenário de necessidades específicas, já que o modelo é moldado às necessidades de quem o criou e o alimenta.

4. Para criar um modelo preditivo, algumas etapas precisam ser cumpridas. O Apêndice IV – Ciclo de desenvolvimento de soluções de IA aborda esse tema.

Visão Computacional

5. Visão computacional é um domínio de aplicação e área de estudos que precede o aprendizado de máquina e envolve como os computadores enxergam, entendem e extraem informações relevantes sobre imagens e vídeos. A visão computacional tenta mimetizar a habilidade biológica humana de enxergar e compreender o ambiente ao seu redor para que possa tomar decisões baseadas nesses dados.

6. A visão computacional tem aplicação útil na Administração Pública em áreas que necessitem de análises visuais como: acompanhamento de obras por imagens de satélite como rodovias, usinas e qualquer outro tipo de objeto que possa ser capturado por satélites com resoluções cada vez mais altas; reconhecimento de áreas e busca de sobreviventes com drones por meio do reconhecimento de pessoas pelos vídeos capturados; reconhecimento facial de pessoas procuradas; reconhecimento de placas de veículos, detecção precoce de doenças como diabetes, alguns tipos de câncer, Alzheimer, entre vários outros casos na área médica a partir de análise de imagens.

7. A área de visão em seu período anterior ao uso de aprendizado de máquina se baseava em algoritmos determinísticos com filtros para detecção de bordas, limites, curvas, cantos, entre outros para que fosse possível compor uma representação e entendimento de mais alto nível. Em 1989, Yann LeCun construiu uma solução para o problema baseada em uma nova tecnologia à época denominada redes convolucionais, para reconhecimento de códigos de endereçamento postal (CEPs) escritos a mão em cartas nos Estados Unidos, com o objetivo de automação desta atividade. Abaixo está representada uma rede Convolucional para reconhecimento de dígitos de 0 a 9.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

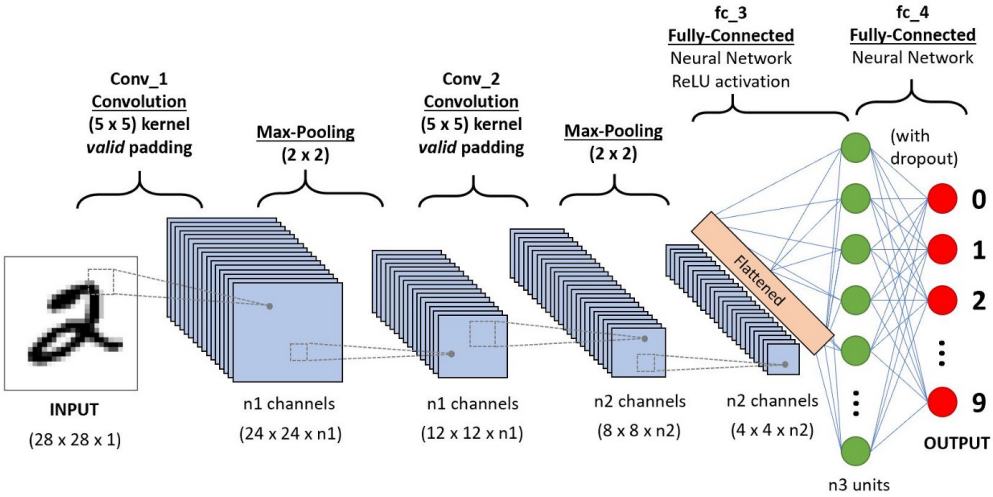


Figura 31 – Representação do reconhecimento de números.

8. A grande popularização recente da área de visão computacional aconteceu a partir de 2012, com o advento de competições como a *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) (https://www.image-net.org/challenges/LSVRC/), que buscam a avaliação de soluções nesta tecnologia.

9. Nos anos que se seguiram, a competição foi dominada por modelos de aprendizado profundo, baseando-se principalmente em arquiteturas de redes convolucionais cada vez mais profundas. As taxas de erro de classificação foram baixando a ponto de superar o nível humano médio. Com os resultados melhores e as redes cada vez mais profundas e demoradas para se treinar, a transferência de aprendizagem foi ganhando cada vez mais espaço para reaproveitar o treino prévio e apenas refinar os pesos da rede em novas aplicações sem a necessidade de treinar do zero.

10. Mesmo com os significativos avanços dos últimos anos em detecção de objetos em imagens, o problema está muito longe de ser considerado resolvido quando comparado com o que é feito por seres humanos. Exemplos disso incluem a descrição de imagens com inúmeros objetos interagindo entre si, entendimento de vídeos ou sequências de vídeos longos como filmes e séries contextualizando o que é visto com o que é de conhecimento geral de pessoas para generalizar suas observações.

11. Uma compilação de artigos publicados na área de visão computacional com o estado da arte e código fonte da implementação está disponível em https://paperswithcode.com/area/computer-vision.

Processamento de Linguagem Natural

12. Processamento de linguagem natural (PLN) é uma subárea da ciência de computação, mais especificamente, da inteligência artificial, que está associada aos problemas não apenas de processamento, mas também de compreensão da linguagem natural humana em seus vários idiomas, e de geração de linguagem. Engloba tanto a linguagem falada como a escrita. Suas aplicações envolvem desde a análise e processamento de texto escrito até processamento e geração de diálogos entre humanos e agentes não humanos.

13. O uso de PLN é uma área de forte aplicabilidade no serviço público, dada principalmente a importância da linguagem escrita (normativos, documentos digitalizados etc.), em que se vislumbra por exemplo o processamento automático de centenas ou milhares de documentos (ex.: peças de processos judiciais ou de órgãos de controle) em que a aplicação tão somente da capacidade humana tem-se mostrado inviável, o que historicamente tem contribuído para a tão conhecida lentidão na tramitação de processos e outros assuntos diretamente relacionados à vida dos cidadãos.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

14. Os primeiros usos do que hoje vem a ser conhecido como processamento de linguagem natural fazia uso quase exclusivo de reconhecimento determinístico de padrões por meio de expressões regulares, um tipo de programação baseado portanto em regras bem definidas. Posteriormente, os usos de aprendizagem de máquina foram focados no treinamento de modelos tradicionais como regressão logística, Naïve Bayes[1], máquinas de vetores de suporte, entre outros, em que os dados de entrada continham inúmeras combinações de atributos que visavam a caracterizar completamente os textos (ex.: classificação gramatical de cada palavra obtida por algoritmos específicos, classificação gramatical das *N* palavras subsequentes e/ou anteriores etc.), em um processo conhecido como engenharia de atributos (*feature engineering)*. A escolha das características ou *features* a serem utilizadas era um processo manual, de alto custo, que consumia a maior parte do tempo alocado ao desenvolvimento de modelos de aprendizagem. Nos últimos anos, redes neurais têm sido empregadas largamente não somente na etapa de aprendizado de tarefas propriamente dito, em substituição aos algoritmos citados no parágrafo anterior, como principalmente na geração da representação do texto de entrada. Em suma, o desenvolvedor não mais se preocupa em determinar as características de cada palavra do texto a ser analisado, mas delega à rede a tarefa de, dada uma sequência de palavras ou subpalavras presentes em um vocabulário, gerar a representação mais adequada à semântica do texto ou discurso falado, para que essa representação seja utilizada pela mesma rede ou até mesmo por outro algoritmo para a tarefa desejada (ex.: classificação textual).

15. Por fim, (Jurafsky & Martin, 2020) chama atenção para possíveis maus usos de modelos de inteligência artificial, a exemplo do modelo GPT-3 (Brown, et al., 2020), uma rede neural geradora de textos criada pela OpenAI. Considerado por alguns como "a inteligência artificial mais avançada já criada pelo homem" (Rodrigues, 2021), o GPT-3 tem levantado sérios questionamentos éticos devido à incrível semelhança dos textos por ela gerados em relação a textos escritos por humanos.

Principais aplicações e técnicas de processamento de linguagem natural (PLN)

1. Aplicações supervisionadas

1.1. Classificação textual e análise de sentimentos

16. Uma classe de tarefas facilmente compreensível e comum em inúmeros cenários de uso de processamento de linguagem natural certamente inclui classificação de documentos, definida em (Ruder, Text classification, 2019) como "a tarefa responsável por atribuir a uma sentença ou documento uma categoria apropriada, dentre uma lista de categorias dependente da base de dados" (tradução livre) ou da aplicação específica.

17. Modelos baseline para algoritmos de classificação textual em geral incluem Naïve Bayes, regressão logística e máquinas de vetores de suporte. Tais modelos costumam ser a primeira escolha e em muitas situações são de fato os mais adequados, não apenas pelo desempenho em termos computacionais, mas também por atingirem acurácia suficiente para o cenário pretendido de uso, além de não exigirem recursos como GPU ou computação em nuvem.

18. Em (Ruder, Text classification, 2019), é disponibilizada uma lista de bases de dados para treinamento de modelos para a língua inglesa, bem como respectivos modelos estado-da-arte. Dentre tais modelos, vemos que no momento da escrita deste relatório, é comum o uso de tecnologia de rede neural baseada em arquiteturas como *Transformers* (Yang, et al., 2019), ao lado de redes recorrentes.

19. Uma tarefa relacionada à classificação textual, que pode ser vista como uma especialização, é a chamada **análise de sentimentos**, um tipo de aplicação em geral voltado para analisar opinião de

[1] Naive Bayes classifier: https://en.wikipedia.org/wiki/Naive_Bayes_classifier

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

usuários e verificar a polaridade de opiniões ou sentimentos. Aqui, aplicam-se basicamente os mesmos algoritmos empregados em classificação textual, tanto para algoritmos de baseline como estado-da-arte.

20. Como exemplo de *dataset* para o idioma português, temos o TweetSentBR Dataset (Brum & das Graças Volpe Nunes, 2018), além de um corpus de avaliação de hotéis (Freitas & Vieira, 2015).

1.2. *Part of speech recognition*

21. Reconhecimento de partes de um discurso (ou **part of speech recognition**) refere-se à tarefa de classificar cada termo ou palavra em um texto em uma determinada categoria gramatical (ex.: adjetivo, advérbio, substantivo comum/próprio, preposição, interjeição, verbo auxiliar, numeral, pronome, sinal de pontuação etc.). Segundo (Jurafsky & Martin, 2020), ao lado do reconhecimento de entidades mencionadas, é uma tarefa capaz de fornecer pistas úteis para indicar a estrutura e significado de sentenças.

22. O reconhecimento de partes de um discurso – referenciado deste ponto em diante pelo acrônimo inglês *POS tagging* – pertence a uma classe de tarefas denominada rotulagem de sequências. São tarefas que atribuem a cada palavra x_i em uma sequência de entrada X um rótulo y_i (ex.: verbo ou substantivo etc.), gerando, portanto, uma nova sequência Y com o mesmo tamanho da sequência de entrada X. As soluções utilizadas para este fim incluem desde algoritmos probabilísticos/estatísticos clássicos como **Hidden Markov Model**[1] com algoritmo **Viterbi**[2] e **Conditional Random Field**[3] até algoritmos modernos baseados em redes neurais recorrentes (Wang, et al., 2021) ou, mais recentemente, redes baseadas em *Transformers* (Heinzerling & Strube, 2019) ou até mesmo combinações inteligentes de *embeddings* (Wang, et al., 2021).

1.3. Reconhecimento de entidades mencionadas

23. **Reconhecimento de entidades mencionadas**, tradução livre para *named entity recognition* (**NER**) é a tarefa de rotular entidades em um texto com seu tipo correspondente. Aqui, "tipo de entidade" não significa classe gramatical ou sintática, mas a categorização de entidades que em geral referem-se a nomes próprios (podendo, entretanto, referenciar conceitos mais abstratos). Os tipos frequentemente encontrados em modelos que classificam entidades mencionadas incluem pessoa, lugar, data etc.

24. Tradicionalmente, modelos de baseline para reconhecimento de entidades mencionadas faziam uso de algoritmos estatísticos como Conditional Random Field, muitas vezes em combinação com redes neurais recorrentes LSTM ou convolucionais.

25. No momento da escrita deste relatório, soluções consideradas estado-da-arte utilizavam *Transformers* (BERT) (Yamada, Asai, Shindo, Takeda, & Matsumoto, 2020) e combinação de rede neural recorrente do tipo LSTM bidirecional com rede convolucional (Shahzad, Amin, Esteves, & Ngonga Ngomo, 2021).

26. Em termos de *dataset* para reconhecimento de entidades na língua portuguesa, temos o HAREM (Mota & Santos, 2008).

1.4. *Chatbots*

27. Em (Jurafsky & Martin, 2020), **chatbots** são definidos como "sistemas projetados para conversações estendidas, de forma a mimetizar conversações não estruturadas características de

[1] Hidden Markov Model: https://en.wikipedia.org/wiki/Hidden_Markov_model

[2] Viterbi algorithm: https://en.wikipedia.org/wiki/Viterbi_algorithm

[3] Conditional Random Field: https://en.wikipedia.org/wiki/Conditional_random_field

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

interações entre humanos, principalmente para fins de entretenimento, mas também para propósitos práticos como tornar agentes orientados a tarefas mais naturais" (tradução livre).

28. Ainda segundo (Jurafsky & Martin, 2020), as três principais arquiteturas para chatbots incluem sistemas baseados em regras, sistemas de recuperação de informações e geradores *encoder-decoder*.

29. **Sistemas baseados em regras** podem ser definidos como aqueles que utilizam padrões textuais semelhantes a expressões regulares para interpretar as perguntas feitas pelo usuário e gerar as respostas mais adequadas. ELIZA (ELIZA, 2021) e A.L.I.C.E (A.L.IC.E, 2019) são exemplos históricos desse tipo de chatbot. Apesar da simplicidade desses sistemas, uma vez que em geral não empregam nenhum mecanismo de aprendizagem de máquina propriamente dito, até hoje alguns chatbots ainda utilizam essa arquitetura básica, também denominada arquitetura "padrão/ação" (Jurafsky & Martin, 2020), sendo alguns capazes até mesmo de serem aprovados no Teste de Turing[1].

30. **Sistemas de recuperação de informações** se assemelham a motores de busca, pois tratam a pergunta realizada pelo usuário como uma *query*. Tais sistemas em geral utilizam técnicas baseadas em similaridade textual.

31. **Geradores encoder-decoder** são uma tentativa de tratar o problema de geração de diálogos por meio de arquiteturas do tipo *encoder-decoder*[2]. Modelos alternativos incluem o *fine-tuning* de um modelo de linguagem em um *dataset* de diálogos, a exemplo do que foi feito em (Paranjape, et al., 2020), com *fine-tuning* do GPT-2 (Radford, et al., 2019). Segundo (Jurafsky & Martin, 2020), essa classe de solução tem foco na geração de respostas únicas, não sendo apropriadas para um fluxo de diálogo longo, em que teriam que ser aplicadas técnicas adicionais baseadas por exemplo em aprendizagem por reforço ou redes adversariais.

32. A classificação acima está longe de ser exaustiva, uma vez que cada vez mais as soluções para construção de *chatbots* tendem a empregar arquiteturas hibridas que empregam desde sistemas baseados em regras e expressões regulares em combinação com aprendizagem de máquina com uso de dados até técnicas de aprendizagem por reforço para prever a próxima pergunta a ser feita ao usuário ou próxima ação a ser executada (ex.: reserva de voos, hotéis etc.).

33. No momento da escrita deste relatório, trabalhos considerados estado-da-arte incluíam o uso de redes neurais recorrentes (Li, Lin, Collinson, Li, & Chen, 2019) e *Tranformers/*BERT (Lai, Hung Tran, Bui, & Kihara, 2020) (Liu, et al., 2020).

1.5. Outras aplicações supervisionadas

34. A tabela abaixo lista outros tipos de aplicações supervisionadas em processamento de linguagem natural que não foram detalhadas neste relatório, mas que sem dúvida merecem atenção tanto em termos de aplicabilidade como para exploração dos limites do uso de inteligência artificial para lidar com as complexidades inerentes à linguagem humana. Os exemplos com as respectivas descrições foram baseados em (Ruder, NLP-progress, 2021), bem como as referências[3] para as implementações

[1] Dizemos eu um sistema é aprovado no Teste de Turing quando um usuário, ao travar uma conversação com tal sistema, não é capaz de diferenciá-lo de um ser humano.

[2] Nesse tipo de arquitetura neural, resumidamente, uma rede codificadora é treinada para gerar uma representação vetorial intermediária do texto de entrada, enquanto uma rede decodificadora é treinada para resolver a tarefa-alvo (ex.: tradução de textos ou, no caso de chatbots, geração de respostas) a partir da representação intermediária.

[3] Dentre as referências aqui citadas, nem todas foram publicadas oficialmente em conferências ou periódicos, estando algumas em fase "*preprint*". Ainda assim, o objetivo aqui não é entrar no mérito de cada trabalho, mas sim apontar os progressos mais recentes na área, que tem evoluído constantemente.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

estado-da-arte segundo o site. Para a listagem completa de tarefas em PLN com respectivas referências, consultar (Ruder, NLP-progress, 2021)

Tarefa	Descrição	Principais tipos de tecnologias empregados nos últimos dois anos	Referências para exemplos de trabalhos estado-da-arte
Extração de relacionamentos	Extração de relacionamentos semânticos de um texto. Relacionamentos extraídos normalmente ocorrem entre duas ou mais entidades de um certo tipo (ex.: pessoa, organização, local) e recaem em um número de categorias semânticas (ex.: "casado com", "empregado por", "mora em").	*Transformer,* *LSTM*	(Baldini Soares, FitzGerald, Ling, & Kwiatkowski, 2019) (Yamada, Asai, Shindo, Takeda, & Matsumoto, 2020) (Nayak & Ng, 2020)
Geração de paráfrase	Geração de uma sentença de saída que preserva o significado da sentença de entrada, mas contém variações em escolha de palavras e gramática.	*Transformer*	(Bui, Le, To, & Cha, 2021)
Question Answering	Resposta a perguntas a partir da compreensão de um texto escrito.	*Transformer*	(Yang, et al., 2019)
Reconhecimento automático de discurso	Compreensão e/ou transcrição de discurso falado	Redes neurais convolucionais *Transformer* Redes neurais recorrentes	(Hsu, et al., 2021) (Gulati, et al., 2020)
Resolução de correferência	Agrupamento de menções no texto que se referem às mesmas entidades do mundo real	*Transformer*	(Kirstain, Ram, & Levy, 2021) (Attree, 2019)
Senso comum	Tarefas de raciocínio de senso comum requerem que o modelo vá além do reconhecimento de padrões, utilizando ao invés "senso comum" ou conhecimento do mundo para inferir suas predições. Exemplos incluem compreensão de textos, geração de textos que descrevam o que uma imagem dá a entender, entre outros.	*Transformer*	(Yang, et al., 2019)
Similaridade textual semântica	Determina o grau de similaridade entre dois textos.	*Transformer*	(Yang, et al., 2019)
Sumarização	Produção de uma versão menor de um ou vários documentos,	*Transformer,* *Pointer-Generator network,*	(Zhong, et al., 2020) (Dou, Liu, Hayashi, Jiang, & Neubig, 2021)

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

	preservando a maior parte do significado original.	LSTM	
Tradução	Tarefa de traduzir uma sentença em uma linguagem de origem para uma linguagem de destino diferente.	*Transformer*	(Liu, Duh, Liu, & Gao, 2020) (Liu, Liu, Gao, Chen, & Han, 2020)
Visual Question Answering	Dada uma imagem e uma pergunta em linguagem natural, a tarefa consiste em fornecer uma resposta correta em linguagem natural.	*Transformer*	(Hu, Singh, Darrell, & Rohrbach, 2020)

2. Aplicações não supervisionadas

2.1. Agrupamento (*Clustering*)

35. De acordo com (Manning, Raghavan, & Schütze, 2009), *clustering* é a forma mais comum de aprendizado não supervisionado.

36. Embora os algoritmos para clustering citados no Apêndice I tenham sido originalmente concebidos tendo em mente técnicas de representação textual convencionais como TF-IDF[1], estratégias de vetorização mais recentes baseadas em BERT/*Transformers* começam a ser empregadas para representar sentenças/documentos para fins de agrupamento (Reimers, 2021).

1.1.1. Modelagem de tópicos

37. Modelagem de tópicos é outra técnica de aprendizado não supervisionado, também utilizada para agrupar documentos por similaridade. Sua principal aplicação prática consiste em agrupar documentos por "tópico" ou assunto/tema de interesse. Outra aplicação prática de alguns dos algoritmos aqui utilizados inclui sistemas de recomendação (ex.: recomendação de livros, filmes ou músicas/artistas que podem agradar a determinado usuário com base em seu histórico de atividades – sejam elas filmes assistidos, livros comprados etc).

38. Tradicionalmente, as tarefas de modelagem de tópicos têm sido resolvidas com métodos matemáticos especialmente no campo da álgebra linear, coletivamente referidos como métodos de decomposição ou fatoração de matrizes. Os principais exemplos incluem **Latent Semantic Analysis**[2], baseada em uma técnica conhecida como **Singular Value Decomposition**[3], e **Non-negative Matrix Factorization**[4]. Outra técnica empregada – **Latent Dirichlet Allocation**[5] – baseia-se por sua vez em algoritmos estatísticos generativos que objetivam modelar cada documento como uma mistura de tópicos (cada tópico em maior ou menor nível) e cada palavra por sua vez como pertencente a um desses tópicos. Detalhes matemáticos a respeito dessas técnicas podem ser encontrados em (Manning, Raghavan, & Schütze, 2009) e/ou nas respectivas referências.

39. Os algoritmos supracitados são antigos e em sua maioria baseados em vetorizações do *tipo bag of words*, em que cada documento é representado por um vetor em que cada posição está associada a uma palavra do vocabulário. Assim, cada posição do vetor contém o número de vezes em que a respectiva palavra ocorre no documento. Porém, nos últimos anos, técnicas de vetorização com uso de

[1] Tf-idf: https://pt.wikipedia.org/wiki/Tf%E2%80%93idf

[2] Latent semantic analysis: https://en.wikipedia.org/wiki/Latent_semantic_analysis

[3] Singular value decomposition: https://en.wikipedia.org/wiki/Singular_value_decomposition

[4] Non-negative matrix factorization: https://en.wikipedia.org/wiki/Non-negative_matrix_factorization

[5] Latent Dirichlet allocation: https://en.wikipedia.org/wiki/Latent_Dirichlet_allocation

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

arquiteturas de rede neural baseadas em *Transformers* têm se tornado cada vez mais comuns, e começam a ser utilizadas também para representar documentos para fins de modelagem de tópicos, ainda que por vezes os mesmos algoritmos acima ainda sejam utilizados para a finalidade principal de agrupamento. Exemplos incluem BERTopic (Grootendorst, 2020), baseado por padrão em vetorização com BERT (embora suporte outros modelos) e com a vantagem de inferir o número ideal de tópicos, e Contextualized Topic Models (Bianchi, Terragni, & Hovy, Pre-training is a Hot Topic: Contextualized Document Embeddings Improve Topic Coherence, 2021) (Bianchi, Terragni, Hovy, Nozza, & Fersini, Cross-lingual Contextualized Topic Models with Zero-shot Learning, 2021), que suporta uma combinação de *transformers* com *bag of words*, requerendo entretanto que se defina a priori a quantidade de tópicos. Por fim, Top2Vec (Angelov, 2020), para modelagem de tópicos e busca semântica, é outro exemplo de ferramenta recente que suporta, entre outras técnicas de vetorização, o uso de vetorização com BERT aliada a HDBSCAN para geração dos agrupamentos.

3. Outras bases de dados para a língua portuguesa

40. A maior parte da pesquisa em processamento de linguagem natural tem tido como foco *datasets* (bases de textos para treinamento de modelos) em idiomas mais comumente falados como inglês e mandarim, e algumas vezes espanhol, francês ou alemão. Por outro lado, o idioma português ainda tem uma carência de bases rotuladas para as várias tarefas disponíveis. Muitas vezes não é possível resolver tarefas para as quais já existe solução (algoritmos mais recentes baseados em inteligência artificial) devido à escassez de uma base de textos que tenha sido rotulada adequadamente.

41. A tarefa de rotulagem manual também é de alto custo, trabalhosa, passível de erros e de subjetividade, e por isso comumente não é uma opção para grande parte dos pesquisadores da academia e da indústria. Por outro lado, modelos de inteligência artificial ainda dependem fortemente de dados para atingirem melhor desempenho, uma vez que o status atual de maior parte do desenvolvimento em inteligência artificial ainda é baseado em aprendizagem de máquina (*machine learning*) supervisionada por dados.

42. Iniciativas já foram criadas para mitigar o problema de linguagens denominadas *low resource* (de baixos recursos), como transferência de aprendizagem (ver Apêndice II – Redes Neurais Artificiais - seção 0).

43. A seguir, apresentamos uma lista de outras iniciativas e *datasets* para tarefas diversas em PLN para a língua portuguesa:

a) *List of Portuguese Datasets for Machine Learning Projects*: https://metatext.io/datasets-list/portuguese-language

b) *PortugueseGLUE Dataset*: tradução para o português do *benchmark* GLUE (General Language Understanding Evaluation (Wang, et al., 2018), que consiste em dados para treinamento de nove tarefas de compreensão de linguagem): https://metatext.io/datasets/portugueseglue

c) PLN - PUCRS – Grupo de Processamento da Linguagem Natural – RECURSOS E FERRAMENTAS: https://www.inf.pucrs.br/linatural/wordpress/recursos-e-ferramentas/

d) Tradução: exemplos de *datasets* incluem o CAPES Dataset (Soares, Yamashita, & Anzanello, 2018) – um corpus paralelo de resumos (abstracts) de teses e dissertações em português e inglês da base da CAPES – além do projeto OPUS (Tiedemann, 2012).

e) Extração de relações: ReRelEM (Freitas, Santos, Mota, Gonçalo Oliveira, & Carvalho, 2009), Summ-it++ (Antonitsch, et al., 2016).

f) Resolução de correferência: Corref-PT (Fonseca, et al., 1998) e Summ-it++ (Antonitsch, et al., 2016).

g) BERTimbau, um modelo de linguagem pré-treinado com o algoritmo BERT (Devlin, Chang, Lee, & Toutanova, 2018) para a língua portuguesa (Souza, Nogueira, & Lotufo).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

h) Tradução do SQuAD (The Stanford Question Answering Dataset[1]): SQuAD v1.1 Dataset (Portuguese SQuAD v1.1 Dataset, 2019).

i) Reconhecimento de discurso falado e síntese de fala: How2 (Sanabria, et al., 2018) contém uma coleção de vídeos instrucionais com legendas em inglês e traduções para português.

44. Paralelamente, vale ressaltar que o *Center for Artificial Intelligence – C4AI* – iniciativa sediada pela Universidade de São Paulo, em parceria com outras instituições acadêmicas e privadas – está conduzindo projetos com o objetivo de melhorar a disponibilidade de corpus para a língua portuguesa (C4AI, 2020).

[1] SQuAD: https://rajpurkar.github.io/SQuAD-explorer/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE IV – Ciclo de desenvolvimento de soluções de IA

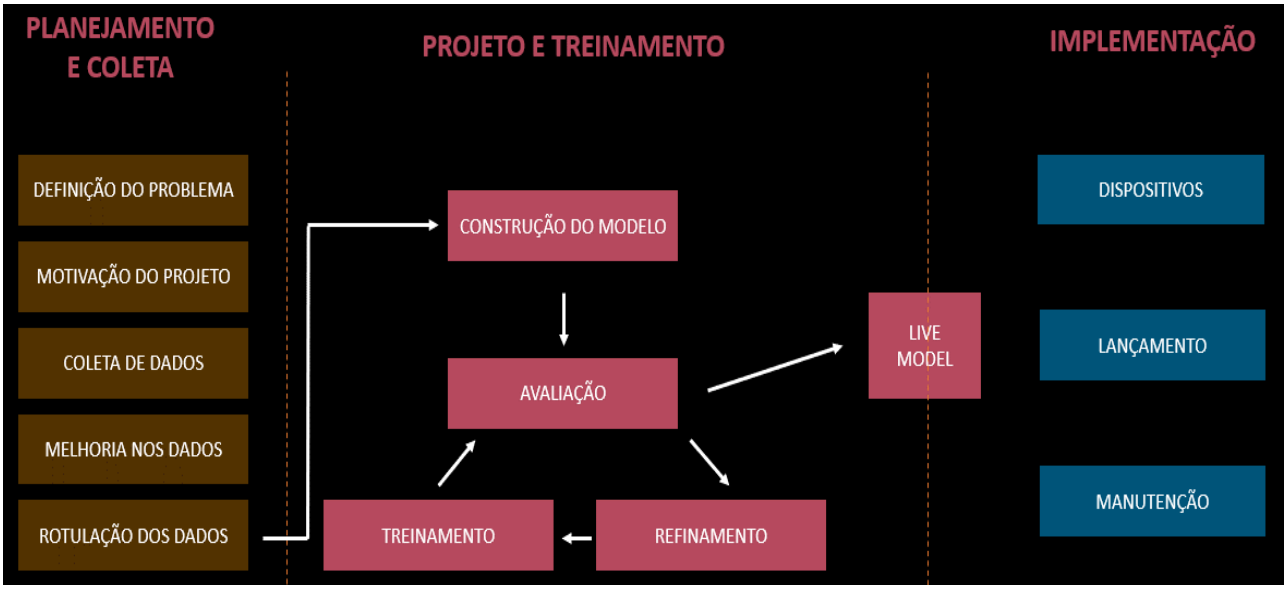


Figura 32 - Ciclo de desenvolvimento de um projeto de IA[1]

0. A fase de **planejamento e coleta** é sem dúvida a parte mais importante do projeto de IA. Este estágio envolve o planejamento e os aspectos motivacionais do projeto. Isso significa que será necessário verificar a lucratividade potencial do futuro modelo de ML e trabalhar o processo de sua adoção, É preciso avaliar se a IA é a tecnologia mais indicada para o atingimento do objetivo pretendido, ou se não existem alternativas tecnológicas mais acessíveis e de menor custo.

1. Também se destaca nesta fase a importância e o protagonismo que os dados assumem para o sucesso do projeto. Entender quais dados o projeto de IA requer é fundamental. Aqui aplica-se o princípio "entra lixo, sai lixo": se os dados coletados forem de baixa qualidade ou insuficientes, o modelo de ML não funcionará, por mais bem construído que seja. Portanto, esta etapa do projeto trata da coleta de dados com os seguintes requisitos: grande volume (quanto mais, melhor); relevante (corresponde ao problema); de alta qualidade (confiável, real e bem estruturado). Esses dados podem ser difíceis de encontrar e coletar para o projeto de inteligência artificial, devendo ser utilizadas fontes internas e externas à organização.

2. Depois de coletar os dados, é importante que eles sejam avaliados previamente. Durante esta etapa do projeto, é necessário reformatar, limpar, analisar e rotular os dados coletados. Para ter uma IA eficiente, a reformatação é necessária para fornecer estrutura aos dados. O anonimato ajuda a proteger os bits sensíveis dos dados. As técnicas de aumento de dados (*data augmentation*) permitem melhorar a diversificação e potencialmente compensar os casos extremos. A normalização dos dados ajuda a evitar vieses. Acima de tudo, a rotulagem dos dados é a atividade mais trabalhosa, uma vez que os dados corretamente rotulados são um elemento essencial para o sucesso do modelo de aprendizagem de máquina a ser desenvolvido.

3. Na etapa de **Projeto e Treinamento**, é o momento em que se deve criar o modelo de aprendizado de máquina, projetar recursos, construir o algoritmo, treiná-lo e ver os primeiros resultados.

[1] Disponível em https://labelyourdata.com/articles/lifecycle-of-an-ai-project-stages-breakdown/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

4. Em seguida, realiza-se o treinamento do modelo proposto. O princípio básico utilizar os dados que foram preparados na fase de planejamento e coleta e colocá-los no algoritmo. Em geral, precisa-se repetir esse processo várias vezes até que se atinja a performance desejada, tanto em tempo de execução quanto na acurácia dos resultados obtidos.

5. Após o treinamento do modelo, realiza-se a etapa de testes, na qual o modelo de aprendizagem de máquina é aplicado a novos dados que não foram utilizados na etapa de treinamento. Após a realização dos testes, é necessário realizar refinamentos no modelo. Esta etapa do projeto é crucial, pois ajuda a eliminar as falhas do modelo e aumentar o seu desempenho. Naturalmente, para ajustar o modelo com precisão, essas três etapas do projeto de treinamento-teste-refinamento devem ser repetidas até obter resultados satisfatórios.

6. Na próxima etapa, após o modelo estar devidamente treinado e refinado, é necessário um esforço extra para que o modelo possa ser implantado no ambiente de produção da organização, e que sejam criados procedimentos que garantam que a qualidade do modelo não será comprometida com o tempo ou com aplicação do mesmo sobre novos dados.

7. A secção seguinte traz a abordagem proposta por meio do *Machine Learning Operations* (ML Ops), que, segundo definição fornecida por (Sweenor, et al., 2020), é o processo de operacionalizar a ciência de dados levando modelos de aprendizagem de máquina para produção – de forma a monitorar seu desempenho e assegurar que eles são justos e estão de acordo com a regulação aplicável.

Machine Learning Operations (ML Ops)

8. Segundo definição fornecida por (Sweenor, et al., 2020), é o processo de operacionalizar a ciência de dados levando modelos de aprendizagem de máquina para produção – de forma a monitorar seu desempenho e assegurar que eles são justos e estão de acordo com a regulação aplicável. (...) ML Ops assegura que modelos continuem a entregar valor à organização, ao mesmo tempo em que forneçam insights críticos para o gerenciamento dos riscos em potencial da tomada de decisões baseada em modelos, assim como mudanças nas condições de negócio ou técnicas" (tradução livre).

9. Ainda de acordo com os mesmos autores, ML Ops originou-se da aplicação de práticas Dev Ops, utilizadas no desenvolvimento tradicional de sistemas, a fluxos de trabalho em ciência de dados e aprendizagem de máquina, englobando não apenas os modelos em si, mas também rotinas complexas de transformação de dados para geração de atributos, métodos estatísticos e outras técnicas analíticas que estejam além do escopo da definição tradicional de aprendizagem de máquina.

10. Em resumo, ML Ops trata de questões como:

a) garantia da justiça (*fairness*) de modelos, com foco em evitar vieses e preconceitos embutidos em distribuições amostrais/populacionais pouco representativas da população real;

b) ciclo de vida, versionamento e aprovação de modelos;

c) papeis envolvidos ao longo do processo;

d) refinamento constante de modelos para que sua acurácia esteja em consonância com mudanças (de negócio ou nos dados reais), o que pode envolver desde o retreinamento de modelos como remodelagem completa (ex.: substituição de um algoritmo por outro), evitando o fenômeno conhecido como *model drift*, e detalhado no item 5 a seguir.

e) considerações a respeito da transparência e interpretabilidade de modelos, em especial para atendimento a requisitos regulatórios;

f) integração de modelos aos demais componentes de um sistema de software.

11. Enquanto metodologias focadas diretamente em ciência de dados e aprendizagem de máquina como CRISP-DM estabelecem diretrizes para o ciclo de desenvolvimento dos modelos em si, de forma a alinhar as ferramentas analíticas às necessidades do negócio, ML Ops complementa essas

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

práticas tratando da operacionalização de um modelo (ie.: entrada em produção e utilização pelo usuário final, levando em consideração ainda questões comuns a outras áreas da engenharia de software, como aquelas relativas a disponibilidade, escalabilidade, manutenção, auditoria e governança). Por esse motivo, a abordagem ML Ops envolve equipes multidisciplinares, uma vez que diferentes papeis como cientista de dados, desenvolvedor de aplicações e administrador de infraestrutura em geral envolvem diferentes perfis de habilidades técnicas, muitas vezes conflitantes.

12. Model Drift

12.1. Conforme já mencionado, os modelos de aprendizado de máquina treinados e em produção devem ser monitorados não apenas pelas métricas comumente utilizadas pela área operacional de TI, como uso adequado de CPU, RAM, disco, rede, entre vários outros, comuns em tópicos relacionados a DevOps. Esses modelos devem também ter sua performance medida ao longo do tempo para responder se o desempenho do modelo degradou (*model drift*) com relação ao que foi medido quando entrou em produção.

12.2. Atualmente, existem duas abordagens usuais para monitorar a degradação de performance de modelos: avaliação contra o *ground truth* e detecção de *drift* da entrada dos dados (*input drift*), conforme ilustrado abaixo ao final do *pipeline* de entrega contínua:

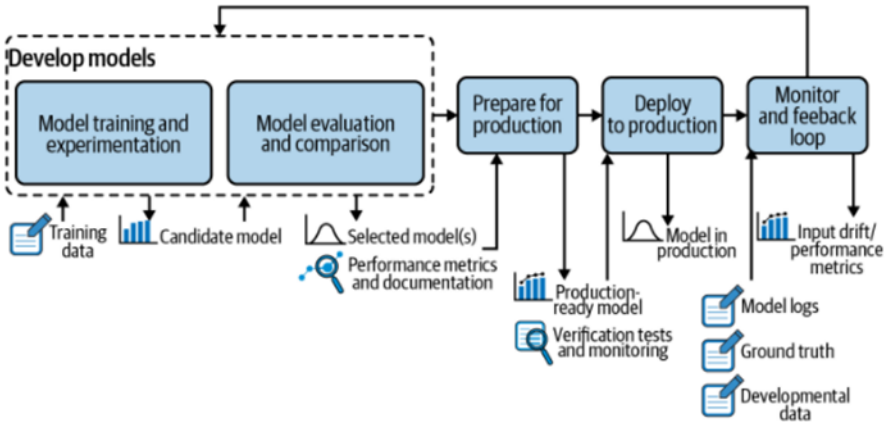


Figure 7-2. Continuous delivery for end-to-end machine learning process

12.3. A abordagem contra o *ground truth* atualmente é a melhor solução, apesar de possíveis limitações práticas, entre as duas para avaliar o *drift* e detectar se o modelo degradou além de um limite considerado aceitável e se, portanto, necessita de retreino. As métricas para avaliar o modelo podem ser tanto estatísticas como, por exemplo, a acurácia do modelo, quanto de negócio como, por exemplo, avaliação de custo-benefício do retreino comparado com o ganho negocial proporcionado pelo novo modelo. A avaliação de *ground truth* nem sempre é possível, pois novos dados rotulados nem sempre estão disponíveis a tempo para evitar que o modelo tenha uma degradação elevada ou então, estão parcialmente disponíveis, inviabilizando seu uso. Além disso, podem ser necessárias transformações de dados complexas ou com temporalidades distintas, reduzindo o custo-benefício ou até mesmo inviabilizando a avaliação.

12.4. A detecção de *drift* da entrada de dados pode ser feita através do monitoramento dos dados de entrada para verificar se houve algum deslocamento da distribuição (*distribution shift*) relacionado a alguma *feature* utilizada pelo modelo ou alguma outra informação que seja correlacionada a alguma *feature* utilizada (variável dependente). Caso a mudança não tenha relação com as *features* utilizadas com o modelo, temos um *drift* conceitual.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE V – Considerações do desenvolvimento de soluções de IA

Qualidade dos Dados

0. Com relação à qualidade dos dados, desde os primórdios da computação, criou-se o termo "*GIGO - Garbage In, Garbage Out*" (Entra Lixo, Sai Lixo), referindo-se à necessidade de que os dados utilizados como entrada nos sistemas computacionais tenham qualidade suficiente para assegurar a boa qualidade dos resultados. Com o aumento das soluções que envolvem tecnologias de IA e aprendizado de máquina, a qualidade dos dados passou a ter um significado ainda mais relevante.

1. Nesse contexto, diversas abordagens foram propostas para endereçar o problema de buscar assegurar que os dados utilizados como base para treinamento e aplicação dos modelos de IA sejam suficientemente seguros e representem fidedignamente a realidade à qual se referem, a exemplo da norma ISO 8000[1], a qual define características de informações e dados que determinam sua qualidade e fornece métodos para gerenciar, medir e melhorar a qualidade de informações e dados.

2. Outro exemplo de aferição da qualidade dos dados é proposto pelo Escritório para Inteligência Artificial do Serviço de Governo Digital do Reino Unido, o qual aponta questões que devem ser consideradas quando da avaliação da qualidade dos dados que serão utilizados na construção dos modelos para aprendizagem de máquina[2]:

Justiça (*Fairness*)[3]

3. Quando do design e implantação de sistemas de IA, é importante sempre se ter em mente que essas tecnologias, não importa o quão neutras possam parecer, são projetadas e produzidas por seres humanos, que estão limitados por seus contextos e preconceitos.

4. O erro humano, o preconceito e o erro de julgamento podem fazer parte do ciclo de vida da inovação e criar vieses em qualquer ponto do processo de entrega do projeto de IA, desde os estágios preliminares de extração de dados, coleta e pré-processamento até as fases críticas de formulação do problema, construção de modelo e implementação.

5. Além disso, as tecnologias baseadas em dados atingem precisão e eficácia ao construir inferências a partir de conjuntos de dados que registram padrões sociais e históricos complexos, que podem conter formas culturalmente cristalizadas de preconceito e discriminação. Não há solução mágica quando se trata de remediar os perigos da discriminação e da injustiça nos sistemas de IA. O problema de justiça e mitigação de viés no design e uso de algoritmos, portanto, não tem solução simples ou estritamente técnica.

6. Embora existam diferentes maneiras de caracterizar ou definir justiça no projeto e uso de sistemas de IA, deve-se considerar o princípio da não discriminação discriminatória como um limite mínimo exigido de justiça. Este princípio nos orienta a não causar danos a terceiros por meio de resultados tendenciosos ou discriminatórios que podem resultar de práticas de inovação em IA

7. Dessa forma, a preocupação com as soluções de IA deve estar presente nas fases de aquisição, tratamento e gerenciamento dos dados, na formulação do projeto, nas saídas planejadas, na implementação e no suporte e monitoramento dos resultados dos modelos.

Accountability

[1] Disponível em https://www.iso.org/standard/63086.html

[2] Disponível em https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector

[3] O resumo apresentado nesta seção é parcialmente baseado em https://www.turing.ac.uk/research/publications/understanding-artificial-intelligence-ethics-and-safety

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

8. Os sistemas de IA devem ser projetados de forma que possam ser auditáveis. Para atingir tal objetivo, o ciclo de desenvolvimento de soluções de IA deve estabelecer uma cadeia contínua de responsabilidade para todas as funções envolvidas no ciclo de vida de design e implementação do projeto, bem como deve ser implementado o monitoramento das atividades que permitam a supervisão e revisão ao longo de todo o projeto.

9. Tais requisitos e responsabilidades são tratados por vários frameworks de referência para a auditoria de soluções de IA, a exemplo dos propostos pelas SAI e pelo GAO, os quais são abordados na seção 5.2 do presente relatório.

Privacidade dos dados

10. A privacidade é um direito essencial à proteção da dignidade humana, devendo ser respeitada, protegida e promovida ao longo do ciclo de vida dos sistemas de IA, tanto no nível pessoal quanto coletivo. É crucial que os dados para IA sejam coletados, usados, compartilhados, arquivados e excluídos de maneiras que sejam consistentes com os valores e princípios estabelecidos nesta Recomendação.

11. Estruturas de proteção de dados e mecanismos de governança adequados estão sendo estabelecidos por agências reguladoras, em nível nacional ou supranacional, protegidos por sistemas judiciais e garantidos durante todo o ciclo de vida dos sistemas de IA, tais como a Lei Geral de Proteção de Dados (Lei 13.709/2018). Este quadro e mecanismos de proteção dizem respeito à coleta, controle e utilização de dados e ao exercício dos seus direitos pelos titulares dos dados e do direito dos indivíduos ao apagamento dos dados pessoais, garantindo um objetivo legítimo e uma base jurídica válida para o tratamento de dados pessoais, bem como para a personalização e despersonalização de dados, transparência, salvaguardas apropriadas para dados confidenciais e supervisão independente eficaz.

12. Os sistemas algorítmicos exigem avaliações completas do impacto na privacidade, que também incluem considerações sociais e éticas de seu uso e um uso inovador da abordagem de privacidade desde o projeto.

13. Nesse contexto, tem-se por outro lado que o compartilhamento dos dados é algo que pode alavancar o desenvolvimento de novas soluções e modelos de aprendizado de máquina por permitir que indivíduos possam ter acesso a dados que não são custodiados por eles. Isso possibilita que mais informações possam ser cruzadas e que resultados mais completos possam ser produzidos. Como exemplo de soluções construídas a partir de compartilhamento de dados produzidos por entidades públicas que tiveram sucesso, temos a Operação Serenata de Amor (https://serenata.ai/) que só foi possível uma vez que a Câmara dos Deputados disponibilizou seus dados para uso da população.

14. A privacidade dos dados é um problema recorrente quando se deseja compartilhá-los. Deve-se levar em conta aspectos relacionados à preservação de privacidade ao disponibilizá-los em algum meio para acesso por outros atores. Mais detalhes sobre a privacidade dos dados são apresentados no Apêndice VI.

15. Acrescenta-se ainda que o TCU realizou trabalho recente referente à implementação dos requisitos da Lei Geral de Proteção de Dados (LGPD) na APF nos órgãos da APF, constante no TC 039.606/2020-1

Interpretabilidade e transparência

16. A transparência dos sistemas de IA é frequentemente uma pré-condição crucial para garantir que os direitos humanos fundamentais e os princípios éticos sejam respeitados, protegidos e promovidos. A transparência é necessária para que a legislação nacional e internacional de responsabilidade funcione de maneira eficaz.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

17. Embora sejam necessários esforços para aumentar a transparência e a interpretabilidade dos sistemas de IA ao longo de seu ciclo de vida para apoiar a governança dos mesmos, o nível de transparência e interpretabilidade deve sempre ser apropriado ao contexto, já que existem algumas compensações entre transparência e explicabilidade e outros princípios como segurança e proteção. As pessoas têm o direito de saber quando uma decisão está sendo tomada com base em algoritmos de IA e, nessas circunstâncias, exigir ou solicitar informações explicativas de empresas do setor privado ou instituições do setor público.

18. Alguns algoritmos de aprendizagem de máquina podem ser classificados como intrinsicamente interpretáveis, a exemplo dos algoritmos de regressão linear, regressão logística, árvores de decisão, Naïve Bayes e *K-nearest neighbour*. Por outro lado, outras classes de algoritmos, como por exemplo modelos baseados em redes neurais, floresta randômica, *ensemble*[1] e máquinas de vetores de suporte, são comumente denominadas de modelos "caixa-preta", por não fornecerem intrinsicamente mecanismos que permitam explicar suas predições com base nos dados de entrada. Ainda assim, existem técnicas que permitem contornar essas dificuldades, tornando possível explicar e interpretar as decisões advindas de modelos caixa-preta, tornando possível que tais algoritmos possam ser empregados de forma mais transparente. Maiores detalhamentos sobre o tema são apresentados no Apêndice VII.

Segurança de modelos

19. Modelos de aprendizagem de máquina, além dos possíveis riscos devido a brechas de segurança inerentes a qualquer sistema computacional, trazem consigo uma nova classe de riscos, advinda de dados que podem ser introduzidos de modo a levar o modelo a gerar respostas erradas ou até mesmo modificar sua lógica de inferência, por meio de retreinamento.

20. Uma classe importante de ataques aos quais modelos de IA estão expostos incluem os **ataques adversariais**, definidos em (Treveil, et al., 2020) como um ataque em geral executado em relação a redes neurais profundas que processam imagens, em que uma modificação a uma imagem que pode parecer insignificante ou mesmo imperceptível aos olhos humanos pode fazer com que o modelo modifique drasticamente sua predição. Em geral, modelos de maior complexidade, em especial aqueles disponíveis abertamente como redes neurais pré-treinadas, costumam ser mais vulneráveis a ataques adversariais do que modelos mais simples.

21. Por outro lado, quando um ataque adversarial ocorre não apenas em tempo de inferência, mas diretamente sobre os dados usados para treinar o modelo, dizemos que se trata de um ataque do tipo ***poisoning attack***, ou ***data poisoning*** (**envenenamento de dados** em tradução livre). Um exemplo largamente divulgado desse tipo de ataque inclui a geração de discursos de ódio por um robô disponibilizado em 2016 em uma rede social, que começou a replicar discursos submetidos por usuários mal-intencionados (Vincent, 2016).

22. Por fim, modelos de inteligência artificial também podem gerar **vazamentos de dados** (***data leakage***) de formas não previstas. Segundo (Treveil, et al., 2020), uma vez que "modelos de aprendizagem de máquina podem fundamentalmente ser considerados um sumário dos dados nos quais eles foram treinados, podem vazar informações mais ou menos precisas dos dados de treinamento" (tradução livre). Na administração pública, trata-se de um risco a ser considerado, por exemplo, em corpus que contenham dados pessoais ou de processos judiciais sigilosos que não tenham recebido um pré-processamento adequado para mascarar esses dados antes de serem empregados para treinar modelos generativos.

23. Assim como ocorre com qualquer sistema computacional, segurança é um requisito a ser considerado por todo o ciclo de desenvolvimento do modelo, com clara definição de papeis,

[1] Ensemble learning: https://en.wikipedia.org/wiki/Ensemble_learning

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

responsabilidades e processos de gerenciamento de riscos. Portanto, no desenvolvimento de sistemas de inteligência artificial, não seria diferente, devendo-se atentar adicionalmente aos tipos de riscos supracitados.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE VI – Privacidade de Dados

1. Em sua forma mais básica, a publicação de dados com preservação de privacidade (*Privacy-Preserving Data Publishing* – PPDP) tem a seguinte forma: Identificador explícito, Quasi identificador, Atributos sensíveis e Atributos não sensíveis.

2. Identificadores explícitos são conjuntos de atributos que contêm informações que identificam explicitamente seus donos, como CPF/CNPJ e nome, entre outros. Quasi identificadores são conjuntos de atributos que podem potencialmente identificar seus donos. Atributos sensíveis são aqueles que identificam informações sensíveis de seus donos, como doenças, deficiências, valor de proposta de empresa em uma licitação antes da abertura das propostas, entre outros. Atributos não sensíveis são todos aqueles que não se encaixam nas três categorias anteriores.

3. Os modelos de privacidade podem ser classificados amplamente em duas categorias com base em seus princípios de ataque. A primeira categoria considera que a ameaça à privacidade ocorre quando se é capaz de vincular um proprietário de registro (pessoa, empresa, entre outros) a um registro em uma tabela de dados publicada (vinculação de registro), a um atributo confidencial em uma tabela de dados publicada (vinculação de atributo) ou à própria tabela de dados publicada (vinculação de tabela). A segunda categoria visa alcançar o princípio não informativo: a tabela publicada deve fornecer poucas informações adicionais ao usuário além do que já é conhecimento prévio dele. Caso contrário, através de análises estatísticas (ataque probabilístico), o atacante poderá aumentar consideravelmente seu conhecimento sobre os registros.

4. A tabela abaixo lista os modelos de ataque e os modelos de privacidade que mitigam os ataques e possibilitam a publicação dos dados levando-se em conta a privacidade:

Modelo de Privacidade	Modelo de Ataque			
	Vinculação de Registro	Vinculação de Atributo	Vinculação de Tabela	Ataque Probabilístico
K-Anonymity	✓			
MultiR *k*-Anonymity	✓			
L-Diversity	✓	✓		
Confidence Bounding		✓		
(α, *k*)-Anonymity	✓	✓		
(X, Y)-Privacy	✓	✓		
(*k*, e)-Anonymity		✓		
(ε, m)-Anonymity		✓		
Personalized Privacy		✓		
T-Closeness		✓		✓
δ-Presence			✓	
(c, t)-Isolation	✓			✓
ε-Differential Privacy			✓	✓
(*d*, γ)-Privacy			✓	✓
Distributional Privacy			✓	✓

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE VII – Algumas soluções para interpretabilidade de modelos de IA

0. De acordo com (Ribeiro, Singh, & Guestrin, 2016), ainda que modelos sejam avaliados por meio de métricas de acurácia em uma base de validação separada, dados do mundo real muitas vezes são significativamente diferentes, e a métrica de avaliação pode não ser indicativa do objetivo do produto. Assim, de acordo com os mesmos autores, inspecionar predições individuais e suas explicações vale a pena – nesse caso, é importante auxiliar os usuários, sugerindo quais instâncias inspecionar, especialmente para grandes bases de dados.

1. Segundo (Nori, Jenkins, Koch, & Caruana, 2019), é crescente a importância de que usuários sejam capazes de compreender modelos, requisito comum em aplicações de alto risco nas áreas de saúde, financeira e ambientes jurídicos, sendo igualmente importante em problemas gerais de aprendizagem de máquina aplicada, exigências regulatórias e interação homem-máquina.

2. Em 2020, o Information Commisioner's Office (ICO)[1], em coautoria com o Instituto Alan Turing[2], lançou um guia para explicabilidade de modelos de IA (ICO; The Alan Turing Institue, 2020), que sugere melhores práticas para avaliar o impacto de sistemas capazes de tomar ou prescrever decisões que impactem diretamente direitos dos cidadãos, além de políticas, procedimentos e técnicas algorítmicas para melhorar a interpretabilidade desses modelos, sejam elas por meio do emprego de algoritmos intrinsicamente interpretáveis ou mesmo métodos para explicar modelos do tipo caixa-preta.

3. A seguir, apresentam-se algumas soluções para interpretabilidade de modelos de IA[3]:

Gráfico de dependência parcial (Molnar, 2021)

4. Um gráfico de dependência parcial mostra "o efeito marginal que um ou dois atributos têm sobre o resultado de um modelo de aprendizagem de máquina" (Molnar, 2021), podendo mostrar, por exemplo, se o relacionamento entre o resultado da predição e um atributo é linear ou mais complexo. Dentre as principais vantagens, pode-se apontar, além da facilidade de implementação, o fato da técnica ser intuitiva e rapidamente compreensível por leigos: a função de dependência parcial em relação a um atributo particular representa a predição média se forçarmos todos os pontos de dados a assumirem aquele determinado valor para o atributo. Já entre as principais desvantagens, está o fato de que o número máximo de atributos em uma função de dependência parcial é dois, devido à inabilidade humana de visualizar um gráfico com número de dimensões maior que três. Além disso, é assumido que os atributos para os quais a dependência parcial é computada não são correlacionados com outros atributos, cenário que não é realista na grande maioria das situações.

LIME – Local Interpretable Model-Agnostic Explanations (Molnar, 2021)

5. LIME, acrônimo para Local Interpretable Model-Agnostic Explanations (Ribeiro, Singh, & Guestrin, 2016) é uma técnica baseada na ideia de modelos denominados "substitutos", ou seja, modelos interpretáveis que são utilizados para explicar predições individuais de modelos de aprendizagem do tipo caixa-preta. No artigo original, os autores propõem uma implementação dessa técnica de acordo com o seguinte roteiro: assumindo-se que se tenha em mãos o modelo original, para o qual podem-se fornecer dados de entrada e obter as respectivas predições, o LIME gera uma nova base de dados rotulada consistindo em amostras obtidas a partir de perturbações de uma determinada instância – no caso, a instância cuja predição deseja-se explicar - e das predições correspondentes fornecidas pelo modelo caixa-preta. Com essa nova base, o LIME treina um novo modelo interpretável (ex.: baseado em árvore

[1] Autoridade independente do Reino Unido encarregada de garantir direitos de informação de interesse público, promovendo abertura pelos órgãos públicos e privacidade de dados para os indivíduos: https://ico.org.uk/

[2] Instituto nacional do Reino Unido para ciência de dados e inteligência artificial: https://www.turing.ac.uk/

[3] Enquanto apresentaremos aqui apenas em linhas gerais as principais técnicas de interpretabilidade, em (Molnar, 2021) são apresentados os fundamentos teóricos e matemáticos que embasam tais métodos, bem como exemplos de implementações em software.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

de decisão ou regressão logística). Denominamos esse modelo gerado de "modelo de explicação". Ele atua, portanto, como uma aproximação do modelo original, mantendo fidelidade ao modelo original ao menos para instâncias parecidas com aquela para a qual se deseja interpretar a predição fornecida. Uma vantagem é que esse é um dos poucos métodos que funciona para dados tabulares, texto e imagens. Como desvantagem, é citada a instabilidade das explicações, que para duas instâncias próximas podem variar consideravelmente. Por fim, é um método ainda em desenvolvimento e muitos problemas ainda precisam ser resolvidos antes que possa ser empregado com segurança.

SHAP - SHapley Additive exPlainations (Lundberg & Lee, 2017)

6. Em resposta ao fato de que vários métodos têm sido recentemente propostos para auxiliar usuários a interpretarem as predições de modelos complexos, e que muitas vezes não é claro como esses métodos estão relacionados e quando um método é preferível em relação a outro, (Lundberg & Lee, 2017) apresentam um arcabouço unificado para interpretar predições – SHAP. SHAP "calcula para cada atributo um valor de importância para uma predição particular". Utiliza em sua formulação conceitos da teoria dos jogos (ex.: valor Shapley[1]). Segundo os autores, este método unifica seis métodos já existentes, porém com melhor desempenho computacional e melhor consistência com a intuição humana quando comparado com abordagens prévias.

Explicação contrafactual

7. Explicação contrafactual refere-se à técnica de gerar perturbações mínimas nos atributos de uma determinada instância de modo a gerar um impacto significativo no resultado a ser retornado pelo algoritmo de IA, modificando, por exemplo, sua decisão final.

Interpretação de redes neurais (Molnar, 2021)

8. Para fazer predições com uma rede neural, os dados de entrada são repassados através de múltiplas camadas de multiplicação com os pesos aprendidos seguidas por transformações não lineares. Uma única predição pode envolver milhões de operações matemáticas dependendo da arquitetura da rede neural, tornando impossível que um humano possa seguir o mapeamento exato dos dados de entrada para a predição. Assim, para interpretar o comportamento e predições de redes neurais, precisa-se de métodos de interpretação específicos. Podemos certamente utilizar métodos agnósticos em relação a modelo, tais como métodos locais (ex.: LIME) ou gráficos de dependência parcial, porém, redes neurais aprendem atributos e conceitos em suas camadas escondidas e precisamos de ferramentas especiais que revelem tais atributos/conceitos. Por fim, os métodos citados nas seções anteriores foram pensados inicialmente para uso com dados tabulares, e em geral imagens e textos requerem diferentes métodos.

9. Redes neurais convolucionais – comumente aplicadas em visão computacional – aprendem atributos e conceitos abstratos a partir dos pixels brutos das imagens. Atributos de alto nível são aprendidos nas camadas escondidas. Já as últimas camadas convolucionais aprendem atributos como objetos ou partes de objetos. Por fim, as camadas responsáveis por uma tarefa de classificação da imagem, por exemplo, aprendem a conectar as ativações geradas pelas camadas convolucionais que identificaram os atributos de alto nível às classes individuais a serem preditas. A abordagem de tornar explícitos os atributos aprendidos é denominada visualização de atributos. Para uma unidade de rede neural – seja essa unidade um único neurônio, um mapa de ativações (um dos canais RGB) ou uma camada inteira – isso é feito encontrando a imagem que maximiza a ativação para aquela unidade. Claramente um problema de otimização matemática. O problema de otimização pode ser endereçado por meio da geração de novas imagens a partir de ruído aleatório. Visualização de atributos é uma técnica comumente utilizada para explicar redes neurais voltadas para reconhecimento de imagens, tendo como principal vantagem ser capaz de comunicar de maneira não técnica como redes neurais funcionam.

[1] Shapley value: https://en.wikipedia.org/wiki/Shapley_value

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Por outro lado, é possível que sejam geradas imagens (explicações) abstratas para as quais não existam palavras ou conceitos mentais. Além disso, o número de canais de ativação em redes convolucionais é muito alto, na casa dos milhares, tornando impossível para um ser humano visualizar todos os canais. Por fim, outra desvantagem citada em (Molnar, 2021) refere-se ao fato de que visualização de atributos pode gerar uma ilusão de interpretabilidade, uma vez que, ainda que olhemos para centenas ou milhares de visualizações, ainda assim não podemos de fato compreender a rede neural, pois seus canais interagem de uma forma complexa, com múltiplos neurônios podendo aprender atributos muito similares, e para muitos dos atributos aprendidos não há um significado equivalente a conceitos humanos.

10. Já os métodos de atribuição de pixels destacam os pixels que foram relevantes para a classificação de uma certa imagem por uma rede neural. A vantagem dessa técnica é que as explicações são visuais, uma vez que é fácil reconhecer imediatamente as regiões mais importantes de uma imagem. Já entre as desvantagens, (Molnar, 2021) aponta o fato de que é difícil saber se uma explicação é correta, e boa parte dessa avaliação tende a ser apenas qualitativa, ou subjetiva. Também é reportado (Molnar, 2021) que tais métodos podem ser frágeis e pouco confiáveis, citando estudos que demonstram que pequenas perturbações em uma imagem podem fazer com que pixels completamente diferentes sejam apontados como possíveis explicações. Em geral, trata-se de uma área de pesquisa que ainda requer progressos adicionais.

11. Por fim, a abordagem baseada em conceitos endereça as limitações dos métodos anteriores, baseados nos atributos dos dados de entrada (ex.: pixels). Aqui, um conceito pode ser qualquer abstração, tal como cor, objeto ou mesmo ideia. Dado qualquer conceito definido pelo usuário, ainda que a rede neural não tenha sido explicitamente treinada com aquele conceito, essa abordagem detecta o conceito, que pode estar embutido na rede, podendo gerar explicações que não são limitadas pelo espaço de atributos do modelo. Um exemplo dessa técnica refere-se ao TCAV, acrônimo para *Testing with Activation Vectors,* cuja proposta é, para cada conceito dado, medir o grau da influência dele na predição do modelo para uma certa classe. Em (Molnar, 2021) bem como no artigo original (Kim, et al., 2018), são fornecidos os detalhes matemáticos e exemplos de como a técnica funciona.

Outras técnicas

12. Outras soluções e ferramentas baseadas em software que merecem destaque incluem InterpretML (Nori, Jenkins, Koch, & Caruana, 2019), algoritmo Explainable Boosting Machine (Nori, Jenkins, Koch, & Caruana, 2019) e What-If Tool (Wexler, et al., 2019).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE VIII – Detalhamento de Soluções de IA desenvolvidas na APF

1. TRIBUNAL SUPERIOR DO TRABALHO (TST)

BEM-TE-VI – Gestão do Acervo Processual			
Objetivo		**Principais funcionalidades**	
Gestão do acervo processual do TST		✓ *Triagem e encaminhamento de processos para analistas com base em modelo preditivo* ✓ *Geração de modelos de instrução*	
Tipo de aprendizado	**Entradas do modelo de IA**	**Saídas**	**Origem dos dados de treinamento**
Supervisionado	*Modelo de Linguagem jurídico Trabalhista* *Dados históricos de julgados*	✓ *Modelos preditivos para os gabinetes dos Ministros*	*Dados de TRTs* *Sistemas PJe e SIJ*
Algoritmos e técnicas			**Etapa**
Word2vec *AutoML, XGBoost*			*Produção*

Síntese da solução

A solução nasceu em decorrência do grande volume de processos constantes no acervo do TST e da necessidade de apoio aos gabinetes. Em média, tem-se o acervo de onze mil processos por gabinete, sendo julgados em média mil processos por gabinete por mês.

O sistema Bem-te-vi é uma solução que tem o objetivo de gerenciar o acervo processual do TST. Entre suas funcionalidades, está a realização de triagem prévia de processos do acervo, sugestão de encaminhamento para instrução a assessores mais experientes de acordo com a matéria, e previsão de denegação de transcendência e provimento/conhecimento do processo.

Funcionalidades da solução

- Interface para gestão do acervo processual, a qual possibilita a filtragem e busca de informações em todo o acervo do gabinete;
- Modelo preditivo que sugere:
  - denegação de transcendência do processo;
  - assessor mais experiente para analisar a matéria;
  - decisão do processo (provimento e conhecimento);
  - decisão monocrática ou colegiado.
- Possibilidade do usuário estabelecer intervalos de confiança nos modelos preditivos;

Detalhes do tipo de aprendizado empregado

O processo teve início com a construção de um *Data Lake* a partir de dados dos TRTs e do TST (Sistema PJe), consistindo de aproximadamente cinquenta milhões de documentos. Em seguida, a partir do *Data Lake,* utilizou-se a técnica Word2vec (biblioteca Gensim) para geração de um Modelo de Linguagem Jurídico Trabalhista. Posteriormente, utilizou-se o modelo de linguagem

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

construído mais dados históricos dos gabinetes dos ministros (últimos 2 anos) para geração do modelo preditivo final (Spark + H2O).

Entrada do modelo

- Informações do Modelo de Linguagem Jurídico Trabalhista, o qual foi construído a partir de dados oriundos dos 27 Tribunais Regionais do Trabalho (Sistema PJe), utilizando algoritmo de processamento de linguagem natural (Word2vec);
- Dados históricos de julgados dos ministros dos últimos 2 anos

Saída do modelo

Para cada um dos 24 gabinetes de ministro no TST, a solução gera 4 modelos que permitem prever denegação por transcendência, assessor mais experiente para instrução do processo, decisão do processo (conhecimento/provimento) e o tipo de decisão (monocrática/colegiada).

Formato utilizado dos dados

Para construção do *Data Lake*, dados contidos em peças do sistema PJe.

Tipo de armazenamento

Local *(Hadoop)*

Origem dos dados de treinamento

- Documentos do Sistema PJe (TRTs e TST)
- Julgados dos gabinetes dos ministros

Algoritmos e técnicas

- Foi utilizada o modelo Word2Vec (biblioteca Gensim) na geração do Modelo de Linguagem Jurídico Trabalhista.
- Utilização da funcionalidade AutoML da aplicação H2O para geração dos 94 modelos. Em geral, é selecionado o algoritmo XGBoost.

Quantas pessoas estiveram envolvidas (e seus papéis)?

- 2 colaboradores em tempo integral e 1 colaborador em tempo parcial.

Duração do projeto de desenvolvimento

- Primeira entrega realizada em 2 anos (início do projeto em 2018).

Benefícios

- Constatou-se alta correlação entre o uso do Bem-te-vi e a produtividade dos gabinetes.
- Maior uniformização das decisões dos gabinetes devido à possibilidade de agrupamento dos processos.

Fontes utilizadas para construir a solução

Surgiu em uma oficina de pensamento criativo, onde a necessidade de distribuição e triagem de processos foi priorizada. Contato com grupo de estudos do TCU. Artigos científicos, sites das ferramentas no GitHub.

Contatos

siad@tst.jus.br

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

2. COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES)

Mimir			
Objetivo		**Principais funcionalidades**	
Desambiguação de dados		✓ *Automatização do processo de preenchimento das plataformas de governo relacionadas à pesquisa científica* ✓ *Apoio à integração entre sistemas e exposição de dados* ✓ *Higienização, unificação e desambiguação de dados* ✓ *Aprendizado ativo*	
Tipo de aprendizado	**Entradas do modelo de IA**	**Saídas**	**Origem dos dados de treinamento**
Supervisionado	*Campos a serem desambiguados*	✓ *Saída de cada algoritmo (escore de similaridade)* ✓ *Média das saídas de cada algoritmo*	*Diversas fontes de dados relativos à pesquisa científica*
Algoritmos e técnicas		**Etapa**	
✓ *Algoritmos de similaridade e classificação* ✓ *Rede neural* ✓ *Ensemble*		*Em homologação*	

Síntese da solução

A CAPES tem parceria com outros órgãos (ex.: Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, Instituo Brasileiro de Informação em Ciência e Tecnologia - Ibict, Rede Nacional de Pesquisa - RNP, Confederação Nacional de Fundações de Amparo à Pesquisa - Confap etc.), surgindo daí a necessidade de desenvolver soluções comuns, com vista a criar um hub nacional de integração de dados para educação, ciência, tecnologia e pesquisa.

Dentro desse contexto, fez-se necessária a integração entre as diferentes plataformas de informação. Um exemplo de uma dessas plataformas é a Sucupira (CAPES, 2016), que lida com dados de pós-graduação. Devido ao preenchimento de dados de diferentes formas pelas várias instituições, surgiu como desafio a necessidade de automatizar o preenchimento das plataformas de governo, contando adicionalmente com o auxílio de ferramentas de apoio à integração entre sistemas.

A solução Mimir, desenvolvida pela CAPES, tem por objetivo principal processar dados relativos à pesquisa científica disponibilizados por várias instituições brasileiras, de modo a unificar e desambiguar informações relativas a entidades como pesquisadores, alunos e publicações.

Funcionalidades da solução

- Automatização do processo de preenchimento das plataformas de governo relacionados à pesquisa cientifica, substituindo o preenchimento manual;
- Apoio à integração entre sistemas e exposição de dados: desenvolvimento de ferramentas de apoio à integração de sistemas e para exposição de dados entre instituições parceiras;

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Higienização, unificação e desambiguação de dados: Cada instituição tinha ainda um modelo de dados distinto, com diferentes semânticas. Foi desenvolvido um módulo que poderia se integrar às APIs das instituições, a planilhas ou a páginas Web.
- Aprendizado ativo: Recurso de desambiguação manual, fornecendo *feedback* de aprendizado ao módulo de inteligência artificial.

Detalhe tipo de aprendizado

Existe a premissa de que algumas bases possuem informações únicas ("padrão ouro"). Quando informações coincidem com essas bases, considera-se a informação fidedigna. Tais bases portanto são utilizadas para avaliar a acurácia dos algoritmos, embora esse critério ainda não esteja completamente definido.

Tipo de armazenamento

Os dados são armazenados na nuvem da Rede Nacional de Pesquisa, em um banco de dados Redis[1].

Algoritmos e técnicas

- Vários algoritmos que executam cálculo de similaridade (ex.: similaridade de cossenos, distância Jaccard[2], distância de Levenshtein[3], entre vários outros).
- Rede neural, tanto para cálculo de similaridade como classificação, ainda em desenvolvimento em parceria com outras instituições.

Quantas pessoas estiveram envolvidas (e seus papéis)?

A equipe foi composta por 1 (um) gerente de projeto, 2 (dois) desenvolvedores, além de 5 (cinco) pesquisadores em tempo parcial.

Duração do projeto de desenvolvimento

O projeto teve início em outubro de 2019, com a primeira etapa durando 8 (oito) meses, sendo que a parte relativa ao desenvolvimento da rede neural foi executada entre agosto de 2020 a janeiro de 2021, porém, devido à prospecção de novos algoritmos, a equipe considera que a rede neural ainda está em desenvolvimento.

Como os usuários desenvolveram confiança nos resultados produzidos pelo modelo?

A equipe de estatísticos ajudou na análise, comparando com os resultados obtidos por um sistema comercial, chegando à conclusão de que a acurácia era próxima.

Benefícios

A equipe reportou que espera a realização dos seguintes benefícios:

- redução do tempo dos pesquisadores em prover informações;
- melhoria da qualidade da informação;
- redução do trabalho de limpeza dos dados;
- redução do tempo de análise de produção cientifica por parte do pessoal da Qualis[4].

---

[1] Redis: https://redis.io/

[2] Jaccard index: https://en.wikipedia.org/wiki/Jaccard_index

[3] Levenshtein distance: https://en.wikipedia.org/wiki/Levenshtein_distance

[4] Qualis: https://pt.wikipedia.org/wiki/Qualis

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Dificuldades enfrentadas

- Conseguir massa de dados para testar, uma vez que algumas entidades parceiras não queriam enviar os dados por receio de violação à Lei Geral de Proteção de Dados – LGPD.
- Cortes de orçamento.

Lições aprendidas

- Importância da colaboração com entes externos, citando como exemplo um hackaton sobre desambiguação de livros.
- Importância da parceria com professores universitários, uma vez que fizeram uso da contribuição de pesquisadores na área.

Referências

https://www.conectibrasil.org/

Contatos

E-mail: dti@capes.gov.br. Telefone: (61) 2022-6100

3. TRIBUNAL DE CONTAS DA UNIÃO (TCU)

ZELLO – CHATBOT COM CIDADÃOS			
Objetivo		**Principais funcionalidades**	
Chatbot para levar informações do TCU para o cidadão em diversos canais e redes sociais		✓ Interação via portal do TCU ✓ Interação via aplicativo móvel do TCU ✓ Interação via redes sociais ✓ Pergunta e resposta: dúvidas sobre funcionamento do TCU e como interagir e obter informações ✓ Consultas processuais ✓ Consultas a contas julgadas irregulares ✓ Emissão de certidões	
Tipo de aprendizado	**Entradas do modelo de IA**	**Saídas**	**Origem dos dados de treinamento**
Supervisionado	Textos do usuário Intenções dos textos Estados anteriores do *Chatbot*	✓ Intenção do usuário ✓ Entidades nomeadas ✓ Ação do *Chatbot*	Geração de textos manuais Retroalimentação com textos da produção
Algoritmos e técnicas		**Etapa**	
Rede neural LSTM Reconhecimento de entidades mencionadas		Em produção	

Síntese da solução

O *Chatbot* Zello é um assistente para interação com o cidadão através de canais como o portal do TCU, redes sociais e aplicativo móvel do TCU. O desenvolvimento dele foi iniciado em 2018 e está em constante evolução, suportado por um processo de curadoria contínua mantido por uma equipe multidisciplinar de áreas técnicas e de negócio do Tribunal.

Os desafios inicialmente propostos foram responder a um serviço de grande demanda em período eleitoral, acessível somente via portal, que era o de gestores com contas julgadas irregulares para fins eleitorais, bem como levar informações do TCU para as redes sociais, como Twitter, para que os usuários não precisassem sair daquele contexto para interagir com o Tribunal. Posteriormente, foram incluídas consultas a processos, emissões de certidões e perguntas gerais sobre o funcionamento do TCU.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Funcionalidades da solução

- Consultas a processos públicos
- Emissão de certidões
- Consulta a contas de gestores julgadas irregulares para fins eleitorais
- Perguntas e respostas sobre o funcionamento do Tribunal e como interagir com ele.

Detalhe tipo de aprendizado

Para a rotulagem dos dados, as entradas são geradas simulando conversas prováveis com o *Chatbot*. Seus rótulos (intenções) são manualmente informados para que o *Chatbot* possa definir uma ação apropriada baseado na intenção prevista.

Além dos dados gerados, há monitoramento das interações via ferramenta web desenvolvida pela própria equipe para capturar interações e selecionar casos em que a inferência de intenção não foi satisfatória, para que passe pelo processo de curadoria e tenha o exemplo retroalimentado no modelo (para o próximo treino), com intuito de melhorar a predição.

Entrada do modelo

Detecção de Intenção: textos de mensagens de usuário vetorizados usando um recurso do Rasa que implementa a representação em *bag-of-words* das mensagens por contagem de *tokens*, que aqui representam palavras.

Fluxo Conversacional (Rasa *Policies*): Intenção atual detectada e entidades extraídas no primeiro modelo e últimas histórias (estados da conversação antes da ação tomada com textos convertidos para *bag-of-words*) obtidas a partir do histórico de interações do Rasa.

Saída do modelo

- Detecção de Intenção: Classificador multiclasse.
- Entidades nomeadas (modelo previamente treinado no TCU)
- Fluxo Conversacional (Rasa *Policies*): Ação que o *Chatbot* executará. A ação é programada pelo desenvolvedor, e pode enviar um texto pré-definido ou solicitar a entrada de dados necessária para complementar informação para que o *Chatbot* prossiga no atendimento (ex.: número de processo).

Formato utilizado dos dados

Textual com *Markdown*[1] , contendo intenções e seus textos de exemplo e histórias com suas ações baseadas em intenções.

Tipo de armazenamento

Sistema de arquivos para o texto com *Markdown*.

Origem dos dados de treinamento

Textos gerados manualmente e capturados de diálogos em produção.

Algoritmos e técnicas

- Para *Natural Language Understanding*:
  - Extração de Entidades Nomeadas
  - *Bag-of-words*: Vetor esparso usando contagem de tokens, que representam palavras.
- Para Fluxo Conversacional:

[1] Markdown: https://pt.wikipedia.org/wiki/Markdown

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Keras Policy: Rede Neural implementada em Keras[1] pelo Rasa para selecionar a próxima ação usando camada LSTM (Hochreiter & Schmidhuber, 1997) para recorrência.

Quantas pessoas estiveram envolvidas (e seus papéis)?

- Equipe inicial:
  - 1 servidor para desenvolvimento de aplicação, integração e IA;
  - 1 terceirizado para desenvolvimento de aplicação, integração e IA;
  - 1 PO para priorizar demandas estratégicas.
- Equipe atual:
  - 1 terceirizado para desenvolvimento de aplicação, integração e IA;
  - 2 representantes de negócio (Ouvidoria e Centro de Atendimento do Cidadão)
  - 1 representante da área de comunicação para modelagem de diálogos

Duração do projeto de desenvolvimento

O projeto iniciou em 2018 e durou aproximadamente 1 ano e 8 meses. Primeira versão em produção foi com 2 meses de desenvolvimento.

Benefícios

- Interface fácil e intuitiva para acessar alguns serviços do TCU.
- Organização do conhecimento: integração da Ouvidoria com o Centro de Atendimento ao Cidadão gerando uniformidade para o que o TCU repassa de informação.
- Ganhos de economia no atendimento ao usuário.

Fontes utilizadas para construir a solução

Documentação do Rasa, fóruns abertos para tirar dúvidas, consultas de exemplo de código aberto (como o *Chatbot* em Rasa usado pela UnB).

Contatos

sti@tcu.gov.br – STI - Secretaria de Soluções de Tecnologia da Informação

4. AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA (ANEEL)

AGUIA			
Objetivo		**Principais funcionalidades**	
Monitoramento de obras em construção de usinas solares fotovoltaicas utilizando sensoriamento remoto com imagens de satélite		✓ Identificação dos principais marcos de implantação dos empreendimentos de geração solar (início das obras e instalação dos painéis) ✓ Apoiar a avaliação humana e decisão dos fiscais por meio do uso de ferramentas visuais (imagens de satélite e dashboards)	
Tipo de aprendizado	**Entradas do modelo de IA**	**Saídas**	**Origem dos dados de treinamento**
Supervisionado	Imagens de satélite (resolução: 10 metros)	✓ Imagem com Segmentação Semântica	Imagens disponibilizadas pelos Satélites da Sentinel
Algoritmos e técnicas		**Etapa**	
Deep Learning **U-Net (EfficientNet)**		Em produção	

[1] https://keras.io/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Síntese da solução

A Aneel tinha a necessidade de viabilizar ferramenta que auxiliasse nas atividades de acompanhamento e monitoramento de empreendimentos que possuem cronograma de implantação (outorga) em execução, visando ao fortalecimento da atividade de monitoramento, bem como a eficiência de recursos e processos por meio de IA aplicada ao Geoprocessamento.

A primeira aplicação desenvolvida foi o "Aguia" (Acompanhamento Geoespacial de Usinas em Implantação), que é uma aplicação Web utilizada para monitoramento de obras em usinas fotovoltaicas em 2019. A aplicação exibe o mapa do país e possibilita que se informem os empreendimentos a serem acompanhados. Nesses empreendimentos, é possível observar a evolução da obra ao longo do tempo e é possível detectar visualmente as placas de usinas fotovoltaicas instaladas na obra, bem como usinas hidrelétricas e outras fontes de geração. Partindo desse conhecimento prático, buscou-se realizar a própria análise visual de forma automática com o uso de inteligência artificial, iniciando a nova versão da aplicação para usinas fotovoltaicas em 2020.

Funcionalidades da solução

- Visualização de obras georreferenciadas em andamento e seu histórico.
- Busca por empreendimentos em andamento cadastrados.
- Detecção de placas instaladas em usinas fotovoltaicas.

Detalhe tipo de aprendizado

As imagens são segmentadas por um especialista utilizando-se de ferramenta proprietária, o ArcGIS.

Entrada do modelo

- Imagens de satélite gratuitas (tamanho do pixel: 10 metros)
- Variações das imagens (Data Augmentation)

Saída do modelo

- Imagem segmentada identificando:
  - cobertura vegetal anterior ao início das obras;
  - solo exposto com vegetação já removida indicando o início das obras;
  - placas de painéis solares instaladas.

Formato utilizado dos dados

GeoTIFF

Tipo de armazenamento

Sistema de arquivos

Origem dos dados de treinamento

Satélites da constelação Sentinel 2.

Algoritmos e técnicas

Para segmentação da imagem de satélite foi usada a rede Convolucional U-Net com EfficientNet.

Quantas pessoas estiveram envolvidas (e seus papéis)?

- 4 pessoas (1 especialista em I.A, 3 especialistas em Geoprocessamento)

Duração do projeto de desenvolvimento

O projeto da parte de IA foi iniciado em 2020 e durou aproximadamente 1 ano e 6 meses.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Benefícios

- Automatização do processo, reduzindo a necessidade de se fazer a avaliação manual das imagens;
- Redução da assimetria da informação;
- Fortalecimento do *accountability* dos agentes setoriais;
- Otimização do uso dos recursos humanos, sobretudo de servidores públicos mais especializados;
- Aumento da confiabilidade e transparência do processo e dos dados que são utilizados nos modelos do setor, o que inclusive forma preço.

Fontes utilizadas para construir a solução

Solução oriunda da evolução de ferramenta não automatizada para fiscalização remota de obras, intercâmbio de estagiário da UnB e de pesquisa de doutorado de um dos componentes da equipe.

Referências

- Publicação: https://bibliotecadigital.fgv.br/ojs/index.php/cgpc/article/view/80456
- Link da aplicação: https://sigel.aneel.gov.br/aguia/

Contatos

alexorlandi@aneel.gov.br

5. TRIBUNAL REGIONAL DO TRABALHO – 5ª REGIÃO

GEMINI			
Objetivo		**Principais funcionalidades**	
Agrupamento de processos similares para instrução		✓ *agrupamentos de processos em pilhas que contém temas semelhantes,* ✓ *Apresentação de processos similares já julgados*	
Tipo de aprendizado	**Entradas**	**Saídas**	**Origem dos dados**
Não Supervisionado	*Inteiro teor de peças de processos do sistema PJe*	✓ *Clusters com processos similares*	*Sistema PJe*
Algoritmos e técnicas		**Etapa**	
TF-IDF *K-means*		Em produção	

Síntese da solução

A solução surgiu a partir de um *hackaton* da Justiça do Trabalho, realizado em 2019, e nasceu em decorrência da necessidade de os gabinetes do TRT distribuírem de forma mais ágil e direcionada os processos semelhantes para os assessores, para que as instruções possam ser realizadas com ganho de produtividade. Utiliza processamento de linguagem natural (PLN) para analisar peças específicas dos processos de 2ª instância, e em seguida agrupar os processos a partir da similaridade das referidas peças. Os assessores nos gabinetes podem selecionar as peças que serão consideradas pelo modelo para a criação dos agrupamentos.

Atualmente a solução já está incorporada como um módulo do sistema de Processo Judicial Eletrônico (PJe), estando disponível para utilização pelos TRTs. O projeto Gemini está sendo desenvolvido em conjunto pelos TRT5, TRT7, TRT15, TRT 20 e CSJT

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Funcionalidades da solução

Atualmente, a solução desenvolvida encontra-se em produção como um módulo do sistema PJe, com as seguintes funcionalidades implementadas:

- Modelo preditivo que apresenta, com base na similaridade das peças selecionadas, agrupamentos de processos em pilhas que contêm temas semelhantes, apresentando ainda o percentual de semelhança entre os processos de cada agrupamento;
- Apresentação de processos similares já julgados

Detalhe do tipo de aprendizado

A solução de PLN utiliza o método TF-IDF para transformação das palavras em representação numérica. É utilizado o Método do Cotovelo para determinação da quantidade de grupos do algoritmo K-means, sendo que o mesmo modelo de aprendizado de máquina é utilizado tanto para os processos pendentes de julgamento quanto para exibição de processos julgados.

Atualmente, a solução está em processo de substituição do método TF-IDF para Word2Vec.

A partir da validação por especialistas dos rótulos apresentados nos agrupamentos, pretende-se posteriormente implementar modelo supervisionado para classificação dos processos.

Entrada do modelo

Inteiro teor das peças dos processos

Saída do modelo

- Agrupamentos de processos por similaridade.
- Exibição de processos similares já julgados.

Tipo de armazenamento

Local. A transformação do texto é armazenada em sistema de arquivos. Os dados gerados dão armazenados em bancos de dados.

Origem dos dados de treinamento

Documentos do Sistema PJe

Algoritmos e técnicas

Foi utilizada o método TF-IDF para NLP e K-means para agrupamentos.

Quantas pessoas estiveram envolvidas (e seus papéis)?

3 colaboradores em tempo parcial na solução de Inteligência Artificial

Duração do projeto de desenvolvimento

10 meses

Benefícios

Não houve consolidação dos benefícios alcançados com o uso da solução, porém foi verificado aumento de produtividade em gabinetes que a utilizam.

Fontes utilizadas para construir a solução

Pesquisa de pós-graduação patrocinada pelo TRT.

Contatos

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

raphael.oliveira@trt5.jus.br / csjt@csjt.jus.br

6. INSTITUTO NACIONAL DE METROLIGIA, QUALIDADE E TECNOLOGIA (INMETRO)

Identificação de irregularidades através de imagens			
Objetivo		**Principais funcionalidades**	
Monitoramento de inspeções de equipamentos para transporte de produtos perigosos		✓ *Avaliação da completude de registros fotográficos* ✓ *Detecção de fotos clonadas*	
Tipo de aprendizado	**Entradas**	**Saídas**	**Origem dos dados de treinamento**
Supervisionado	*Imagens com tamanho mínimo de 100Kb*	*Lista de objetos detectados e classificação dos respectivos tipos de possíveis irregularidades, com grau de confiança atribuído*	*Dados da própria organização*
Algoritmos e técnicas		**Etapa**	
✓ *ResNet* ✓ *YOLOv3*		*Em ambiente de experimentação com usuário final*	

Síntese da solução

Equipamentos que transportam produtos perigosos devem passar por inspeções regulares por organismos acreditados, que devem enviar ao Inmetro fotos que comprovem que as inspeções foram bem-feitas. Em decorrência da baixa eficiência no monitoramento remoto das acreditações dos organismos de inspeção através dos registros fotográficos, havia a necessidade de identificar, automaticamente e de forma mais tempestiva, indícios de fraudes e irregularidades em inspeções em imagens.

Dessa forma, o Inmetro desenvolveu uma solução para identificar suspeitas de irregularidades em imagens de inspeções de equipamentos para transporte de produtos perigosos (ex.: tanques com gasolina ou outros produtos que ofereçam riscos à saúde). As suspeitas são analisadas por um avaliador humano, que pode aplicar sanções ao organismo de inspeção.

Funcionalidades da solução

- Avaliação da completude dos registros fotográficos, que verifica se todas as fotos obrigatórias foram enviadas através de classificação de imagens e detecção de objetos.
- Detecção de fotos clonadas, por meio do cálculo de hash das imagens.

Por fim, a solução é uma ferramenta para indicar aos humanos indícios de irregularidades, e não os substituir. Trata-se de uma solução desenvolvida em resposta à baixa eficiência dos humanos para identificarem irregularidades através de imagens – a análise humana cobria apenas 0,1% do universo de irregularidades.

Detalhe tipo de aprendizado

Antes do treinamento, as imagens foram analisadas e classificadas manualmente, sendo que imagens de baixa qualidade foram descartadas. Além da classificação, era realizado o traçado dos objetos para treinar o algoritmo de detecção de objetos, por meio da ferramenta Microsoft VoTT (Microsoft, 2021). Um total de cerca de 3000 imagens foram rotuladas para constituir a base de treinamento.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Tipo de armazenamento

Imagens são armazenadas na infraestrutura do próprio Inmetro.

Origem dos dados de treinamento

Base histórica de imagens.

Algoritmos e técnicas

Para as tarefas de classificação de imagens e detecção de objetos, ResNet (He, Zhang, Ren, & Sun, 2016) e YOLOv3 (Redmon & Farhadi, 2018) foram escolhidos, respectivamente, pois, segundo a equipe, "são arquiteturas tradicionais e para o problema em questão abordagens genéricas seriam suficientes, pois os objetos não possuem particularidades que justifiquem a adoção de arquiteturas de rede neural específicas". Para o treinamento para a tarefa específica, foi feito o *fine-tuning* adicional das redes[1].

Quantas pessoas estiveram envolvidas (e seus papéis)?

Em média 1 (um) colaborador para a parte de codificação. Envolvimento adicional de 1 (um) usuário cliente e 1 (um) pesquisador que era coorientador do colaborador responsável pela codificação.

Duração do projeto de desenvolvimento

24 meses.

Benefícios

O método de identificação de irregularidades apresentou, segundo a equipe desenvolvedora, "excelentes resultados, com F1-score de 0,804 e exatidão (acurácia) de 0,98 para detecção final de irregularidades e F1-score de 0,994 e 0,92 para as etapas intermediárias de classificação de imagens e detecção de objetos, respectivamente". Ainda segundo a equipe, isso representa economia.

A equipe citou ainda como benefício em potencial o aumento na assertividade na identificação de irregularidades em um cenário em que o serviço público está com escassez de recursos humanos.

Fontes utilizadas para construir a solução

O projeto surgiu a partir de uma dissertação de mestrado.

Dificuldades enfrentadas

- Falta de GPU (*graphical processing unit*) para treinamento na infraestrutura interna do Inmetro.
- Falta de contrato para anotação manual da base de treinamento.
- Não há controle sobre a qualidade das imagens enviadas ao Inmetro.

Lições aprendidas

- Importância do uso do Google Colab[2] ao menos na fase de pesquisa;
- Importância do engajamento dos usuários para que eles "comprem a ideia" da ferramenta – foi verificado que as pessoas estão entusiasmadas e enxergam os benefícios;

[1] Usualmente, redes como ResNet e YOLO são pré-treinadas na base ImageNet (Stanford Vision Lab, Stanford University, Princeton University, 2021).

[2] O Google Colab é um ambiente gratuito para treinamento de modelos em nuvem.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Importância de patrocínio pela direção de TI e pelo cliente.

Contatos

ctinf@inmetro.gov.br

7. TRIBUNAL REGIONAL FEDERAL – 3ª REGIÃO

Sigma e Sinara			
Objetivo		**Principais funcionalidades**	
Produtividade e padronização de instruções judiciais com reuso de modelos mais adequados às características do processo		✓ *Seleção do modelo mais adequado para novo processo* ✓ *Preenchimento de dados extraídos das peças processuais anteriores*	
Tipo de aprendizado	**Entradas do modelo de IA**	**Saídas**	**Origem dos dados de treinamento**
Supervisionado	*Conjunto de entidades e relações extraídas de peças do processo*	*Classificação e ranqueamento de modelos que podem ser utilizados na redação*	*Dados da própria organização, anotados manualmente por servidores voluntários*
Algoritmos e técnicas		**Etapa**	
✓ *NER (Named Entity Recognition) e Relationship Extraction da biblioteca Spacy e LSTM* ✓ *Classificação por "Bag-of-Entities"*		*Em produção, implantado no Processo Judicial Eletrônico*	

Síntese da solução

O Tribunal Regional Federal da 3ª Região (SP e MS) desenvolveu um sistema que sugere um modelo de instrução ou parágrafo mais adequado às características de um dado processo, enquanto suporte à redação de instrução judicial, visando ganhos de produtividade e padronização. O sistema extrai entidades significativas de peças anteriores do processo, principalmente referências a dispositivos normativos e precedentes citados em decisões anteriores do próprio processo. Essas entidades são então utilizadas para classificação de assuntos e ranqueamento de possíveis modelos ou parágrafos padronizados que seriam os mais adequados na elaboração da minuta.

Funcionalidades da solução

- Extração automática de características mais relevantes de um processo à partir do reconhecimento de entidades nomeadas (NER) e da extração de relações entre essas entidades, no texto de peças do processo. O sistema Sinara é responsável por essa extração de informações de textos jurídicos, identificando, notadamente, os dispositivos normativos e precedentes citados em decisões anteriores do processo.
- O sistema Sigma utiliza então as entidades e relações extraídas pelo Sinara para priorizar os modelos mais relevantes em função da similaridade com usos anteriores.

A solução sugere assim uma lista ranqueada de modelos considerados mais adequados, cuja seleção final é realizada por um analista.

Detalhe tipo de aprendizado

Foi utilizado aprendizado supervisionado para reconhecimento de entidades e relações, seguido de simples contagem de ocorrências dessas entidades em usos anteriores dos modelos de parágrafos criando-se assim um ranqueamento de modelos em função de simples contagem dos diversos tipos de entidades e relações.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Tipo de armazenamento

Texto das peças processuais e das entidades e relações identificadas em cada texto são armazenados em arquivos e banco de dados do órgão.

Origem dos dados de treinamento

Histórico processual do órgão, com anotação inicial de 400 parágrafos para o NER, e progressiva expansão das anotações até as cerca de 5000 que foram realizadas por voluntários do órgão, utilizando a ferramenta WebAnno para suporte às anotações. Diversos workshops foram realizados para mobilizar os voluntários e ensinar-lhes a anotar as entidades e relações, incluindo até magistrados nesse esforço de anotação.

A construção do conjunto de dados de treinamento priorizou processos nos temas mais frequentes, identificados pela classificação de assuntos, realizada por especialistas na hora da distribuição, selecionando diversas peças processuais diferentes (petição inicial, sentença, acórdão, etc.)

Algoritmos e técnicas

Utilizou-se o NER padrão do Spacy 2.0 (BI-LSTM), com testes de BERT e extração simultânea de NER e RE, no Spacy 3.0, mas o NER do Spacy 2.0 demonstrou ser mais rápido.

Ambas as funcionalidades do Sinara e do Siga as ferramentas foram integradas ao Sinapses do CNJ, para fácil reuso por outros órgãos do Judiciário.

De forma bastante original, o NER é utilizado para extrair dispositivos legais (número de leis, artigo, parágrafo), precedentes, súmulas e demais referências; sendo que apenas os dados extraídos pelo NER serão utilizados na classificação: todo o restante do texto é descartado. Essa abordagem é assim muita segura do ponto de vista de anonimização dos dados, visto que apenas as referências jurídicas são utilizadas, sem qualquer outra característica do caso concreto.

Quantas pessoas estiveram envolvidas (e seus papéis)?

Um único colaborador para o desenvolvimento do sistema Sigma e um em tempo parcial para o Sinara. 30 a 40 colaboradores voluntários para a anotação dos dados, com curadoria.

Duração do projeto de desenvolvimento

Cerca de 24 meses.

Benefícios

O algoritmo de ranqueamento é facilmente explicável, indicando modelos de parágrafos com alta chance de serem adequados. Os ganhos de produtividade gerados pela adoção do Sigma permitiram zerar a carga de processos tributários da presidência do TRF3, facilitando a averificação dos pressupostos dos recursos.

O Sigma ganhou prêmio de melhores práticas e está como finalista no Prêmio Inovare.

Fontes utilizadas para construir a solução

O projeto apoiou-se em funcionalidades disponíveis na biblioteca Spacy (NER e RE), mas pretende tornar-se mais independente do Spacy por questões de flexibilidade.

Dificuldades enfrentadas

- Falta de pessoal em quantidade e capacidade para desenvolvimento de mais projetos na área.
- Falta de capacidade de processamento (GPUs) para desenvolvimento de modelos mais complexos.
- Riscos de regulamentação, especialmente com a LGPD que proíbe genericamente diversos usos de dados privados, de forma pouco específica, o que gera insegurança quanto à

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

responsabilização de desenvolvedores em IA. Por exemplo, seria tecnicamente viável construir um BERT jurídico, mas haveria um risco para os desenvolvedores, de que o modelo memorizasse algum dado pessoal. Requer-se uma estratégia de anonimização de referência para a segurança jurídica dos desenvolvedores.

- Falta de entidades revisoras de modelos de IA.
- Falta de uma comunidade de desenvolvedores de IA no judiciário.

Lições aprendidas

- Importância da gestão de projeto de IA.
- Organização de voluntários para anotação dos dados de treinamento.
- Documentação do projeto e do sistema final desenvolvido.
- Recomenda utilizar um órgão de validação ético e jurídico, como mitigador do risco de responsabilização dos desenvolvedores. Ver publicação do TRF3: https://www.trf3.jus.br/documentos/adeg/Inova/LIAA-3R/Aquivos/Manual_de_Conformidade_-_final_13.abril.pdf

Contatos

Laboratório LIAA-3R: ilabtrf3@trf3.jus.br

Fábio Collado: fcollado@trf3.jus.br

8. TRIBUNAL DE CONTAS DA UNIÃO (TCU)

Assistente Conjur			
Objetivo		**Principais funcionalidades**	
Sugestão de teses para resposta a questionamentos judiciais e geração de minuta de pareceres		✓ Seleção de teses a partir de petições iniciais ou mandados de segurança ✓ Geração de minuta de pareceres	
Tipo de aprendizado	**Entradas**	**Saídas**	**Origem dos dados**
Supervisionado e Não supervisionado	Documentos judiciais, Pareceres anteriores, Tópicos inferidos, Base de nomes	✓ Tópicos ✓ *Embeddings* do Word2Vec ✓ Tese do parecer ✓ Concordância de gênero para pessoas físicas e jurídicas	Pareceres anteriores, petições iniciais, mandados de segurança e base de pessoas da Receita Federal
Algoritmos e técnicas		**Etapa**	
Latent Dirichlet Allocation (LDA) Word2Vec Random Forest Long short-term memory (LSTM)		Em produção	

Síntese apresentação geral da solução

A solução desenvolvida visa compensar a diminuição de pessoal e o volume crescente de trabalho, além da uniformização dos pareceres produzidos pela Conjur.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

A aplicação de Instrução Assistida de Pareceres sobre Processos Judiciais é utilizada na Consultoria Jurídica (Conjur) do Tribunal de Contas da União para sugerir as melhores teses para resposta a questionamentos judiciais contra a União, via STF ou AGU, acerca de decisões do TCU.

Funcionalidades da solução

- Seleção de teses a partir de petições iniciais ou mandados de segurança.
- Geração de minutas de pareceres.
- Geração de texto com concordância de gênero de pessoas físicas e jurídicas.

Tipo de aprendizado

- Não supervisionado para os modelos de geração de tópicos e de geração de embeddings.
- Supervisionado para os modelos de predição de tese e de predição de gênero para pessoas.

Detalhe tipo de aprendizado

Não supervisionado:

Para a geração de tópicos, foi utilizado o *Latent Dirichlet Allocation* (LDA) treinado sobre os pareceres anteriores.

Para a geração de *embeddings* para ranqueamento, foi utilizado o Word2Vec usando distância de cosseno treinado sobre os pareceres anteriores.

Supervisionado:

Para a predição de tese foram utilizados os pareceres anteriores já rotulados.

Para a predição de gênero, foram usados os nomes de pessoas da base da Receita Federal e como rótulos, o sexo também presente nesta base.

Entrada do modelo

Todos os modelos usam como entrada a petição inicial ou o mandado de segurança.

Saída do modelo

O gerador de tópicos tem como saída 100 tópicos possíveis e suas probabilidades.

O gerador de *embeddings* tem como saída a representação Word2Vec em 300 dimensões do documento.

A predição de tese é um classificador multiclasse.

A predição de gênero é um classificador multiclasse.

Formato utilizado dos dados

Documentos em PDF e dados tabulares.

Tipo de armazenamento

Sistema de arquivos para os PDFs e bases de dados relacionais e de grafo.

Origem dos dados de treinamento

Petições iniciais de mandados de segurança.

Algoritmos e técnicas

- Para geração de tópicos:
  - *Latent Dirichlet Allocation* (LDA).
- Para geração de *embeddings*:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

  - Word2Vec.
- Para a predição de teses:
  - Random Forest
- Para a predição de gênero:
  - *Long short-term memory* (LSTM).

Quantas pessoas estiveram envolvidas (e seus papéis)?

- Equipe:
  - 1 servidor para desenvolvimento da aplicação Web;
  - 1 terceiro para desenvolvimento da aplicação Web;
  - 1 servidor para treino dos modelos de aprendizado de máquina;
  - 1 terceiro para atividades de User Experience (UX);
  - 1 terceiro para prototipação de telas;
  - 1 Product Owner (PO) para priorizar demandas na área de negócio.

Duração do projeto de desenvolvimento

O projeto iniciou em maio de 2019 e durou aproximadamente 1 ano e 2 meses para a primeira entrega.

Benefícios

- Ganho de eficiência;
- Melhoria de qualidade decorrente da padronização;
- Melhoria na imagem do Tribunal e no relacionamento institucional com a AGU e o STF pelo atendimento célere.

Fontes utilizadas para construir a solução

Exemplos de código-fonte na Web para sumarização extrativa de texto usando aprendizado de máquina.

Contatos

sti@tcu.gov.br – STI - Secretaria de Soluções de Tecnologia da Informação

conjur@tcu.gov.br – Conjur - Consultoria Jurídica

9. SECRETARIA DO TESOURO NACIONAL (STN)

PropLegis			
Objetivo		**Principais funcionalidades**	
Identificação automática de padrões em proposições legislativas		✓ *Rastreio diário de proposições na Câmara e no Senado para classificar se são de interesse do Tesouro Nacional* ✓ *Classificação das proposições legislativas de interesse por área de negócio*	
Tipo de aprendizado	**Entradas do modelo de IA**	**Saídas**	**Origem dos dados de treinamento**
Supervisionado	*Textos vetorizados em TF-IDF*	✓ *Classificação se uma proposição legislativa é de interesse ou não* ✓ *Atribuição de probabilidades de cada proposição*	*APIs de acesso às propostas legislativas*

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

	referir-se a determinados assuntos.
Algoritmos e técnicas	**Etapa**
✓ *Modelagem de tópicos com LDA e NMF* ✓ *Classificação com uso de ensemble (SVM, random forest, redes neurais)*	*Em produção*

Síntese apresentação geral da solução

Uma área específica da STN é responsável por identificar proposições legislativas, até então por meio de esforço manual. O desafio envolvia tornar essa tarefa mais proativa, de forma a permitir que se automatizasse a identificação de padrões nas proposições.

Nesse sentido, a Secretaria do Tesouro Nacional – STN desenvolveu, em parceria com o Serpro, uma solução capaz de identificar automaticamente proposições legislativas que possam envolver assuntos de interesse do órgão e posteriormente classificá-las por assunto de acordo com as respectivas áreas de negócio.

Funcionalidades da solução

O projeto foi dividido em duas etapas:

- A primeira missão era acompanhar diariamente as proposições legislativas na Câmara dos Deputados e no Senado Federal e classificar se eram ou não de interesse da STN;
- A segunda parte lidava com encaminhamento automático para as áreas de negócio da STN, envolvendo, portanto, a identificação de padrões relacionados a assuntos.

Tipo de aprendizado

Supervisionado (classificação de interesse ou não) e não supervisionado (modelagem de tópicos)

Detalhe tipo de aprendizado

As áreas de negócio anotaram manualmente a base para treinamento dos modelos.

Entrada do modelo

Textos vetorizados com TF-IDF.

Saída do modelo

A primeira classificação utiliza medida de probabilidade para aferir se a proposição é de interesse ou não – considerando o limiar acima de 75% para classificar como proposição de interesse. Já a modelagem de tópicos retorna a probabilidade de uma proposição estar relacionada a cada um dos possíveis assuntos e suas respectivas áreas de negócio.

Tipo de armazenamento

Banco de dados relacional e armazenamento proprietário de arquivos (nuvem privada pertencente ao Serpro) para salvar as proposições em formato PDF.

Origem dos dados de treinamento

APIs de acesso às propostas legislativas.

Algoritmos e técnicas

- *Latent Dirichlet Allocation* (LDA) para modelagem de tópicos;
- *Non-negative Matrix Factorization* (NMF) para modelagem de tópicos em textos mais curtos;

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- *Ensemble* (comitê) de classificadores dos tipos máquinas de vetores de suporte, *floresta randômica*, redes neurais, entre vários outros algoritmos;
- Métrica para cálculo de similaridade textual: similaridade por cossenos, comparando tanto vetores em TF-IDF como *word embeddings*[1] e vetores de tópicos, gerando em seguida uma combinação desses 3 (três) cálculos;
- Classificação por áreas: uso de várias representações vetoriais e vários algoritmos de classificação, cujos resultados são combinados por meio de *ensemble*. Os detalhes de quais algoritmos foram utilizados não foram divulgados pela equipe sob a alegação de questões contratuais envolvendo a propriedade intelectual do software;
- Aprendizado ativo, por meio de *web service* utilizado pela área de negócio para informar erros de predição.

Quantas pessoas estiveram envolvidas (e seus papéis)?

Um gerente de projeto pela STN, 2 *product owners* (POs) auxiliados por suas respectivas equipes (ao todo 4 a 5 pessoas). Por parte do Serpro, 2 especialistas para a parte de inteligência artificial.

Duração do projeto de desenvolvimento

O projeto está em produção desde setembro de 2020. Duração total informada entre 1 ano e 1 ano e meio para liberação da primeira versão. A área de negócio ainda não priorizou a próxima etapa, e por isso o projeto não está no momento em andamento.

Benefícios

- Atendimento ao interesse da STN em agir de forma proativa, como parte de sua missão institucional.
- Eficiência: identificação das áreas candidatas a receberem uma demanda para análise, detecção de similaridade com demandas antigas para orientar os servidores a tratarem novas demandas de forma a evitar retrabalho e reduzir o tempo gasto.

Fontes utilizadas para construir a solução

A solução foi orientada diretamente pela área de negócio a partir de um problema existente. A intenção não era fazer um laboratório de inovação. As fontes consultadas incluíram apenas especialistas internos e equipe do Serpro.

Dificuldades enfrentadas

- Métricas para especificação do esforço a ser envolvido.
- Uso de dados não estruturados (ex.: PDF).

Lições aprendidas

- Órgãos públicos precisam ter cuidado com laboratórios de inovação sem clareza do que será entregue para não haver desperdício de recursos.
- Importância dos especialistas técnicos em inteligência artificial.
- Projetos de ciência de dados são mais difíceis de estimar o esforço, porque a incerteza é inerente – ao longo do caminho descobre-se que é necessário seguir por outro caminho.

Contatos

ti.stn@tesouro.gov.br

[1] Word embedding: https://en.wikipedia.org/wiki/Word_embedding

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

10. CÂMARA DOS DEPUTADOS

Plataforma Ulysses	
Objetivo	**Principais funcionalidades**
Conjunto de algoritmos de inteligência artificial construídos para sustentar funcionalidades diversas necessárias à plataforma digital legislativa	✓ *Distribuição automática de solicitações aos grupos de consultores* ✓ *Pesquisa por estudos técnicos e projetos de leis similares às solicitações recebidas do Gabinete Digital* ✓ *Classificação automática de todo o conteúdo digital disponível no sítio eletrônico da Câmara dos Deputados*

A Câmara dos Deputados possui um serviço de ciência de dados que trata de *analytics*, governança de dados e inteligência artificial. Neste último caso, a ideia é que módulos de IA sejam acoplados a sistemas e soluções dentro da estratégia de tecnologia da informação, não tendo foco portanto no desenvolvimento de sistemas completos.

O órgão conta atualmente com uma plataforma denominada ULYSSES, definida em sua estratégia de transformação digital (Câmara dos Deputados, 2021) como um "conjunto de algoritmos de inteligência artificial construídos para sustentar funcionalidades diversas necessárias à plataforma digital legislativa". Por esta razão, a plataforma é descrita neste relatório não como um case isolado, mas sim como um conjunto de soluções inteligentes voltadas para diferentes problemas.

Alguns módulos da plataforma ULYSSES incluem:

- **Algoritmo para distribuição automática de solicitações aos grupos de consultores**: objetiva distribuir entre as áreas da consultoria legislativa as ideias que os parlamentares apresentam para projetos de lei, uma vez que tais demandas requerem elaboração de estudos técnicos preliminares. Na Câmara dos Deputados, os consultores legislativos são distribuídos por área temática, mas temas relativos a novos projetos de lei eventualmente podem envolver várias áreas temáticas. As ideias de projetos de lei são apresentadas pelos deputados como texto livre, e havia originalmente um processo manual pelo qual um servidor designado precisava ler esse texto para inferir por quais grupos temáticos a ideia precisaria passar. A necessidade era implementar um algoritmo de classificação capaz de predizer as 3 (três) principais áreas candidatas. Com a implantação do algoritmo de classificação, atualmente um servidor precisa assinalar se aceita ou não as sugestões do classificador, porém se espera que na próxima versão não haja tal intervenção humana, sendo as proposições diretamente direcionadas às respectivas áreas temáticas. Assim foi decidido porque, com a taxa de erro do algoritmo com relação à alternativa mais provável – cerca de 10% segundo a equipe entrevistada – é aceito que, ao chegar ao respectivo consultor, ele possa redirecionar para a segunda alternativa mais provável. Até o momento, de acordo com feedback fornecido pelos usuários, quando o algoritmo apresenta margem de dúvida (uma vez que retorna como saída a probabilidade ou grau de confiança para cada uma das três alternativas), o mesmo ocorre com o julgamento humano, o que torna, portanto, o desempenho do modelo próximo ao desempenho apresentado por humanos na execução da tarefa. O módulo suporta ainda aprendizagem ativa (usuário pode fornecer feedback para a classificação).
- **Algoritmo para pesquisa por estudos técnicos e projetos de leis similares às solicitações recebidas do Gabinete Digital**: a linguagem utilizada pelos parlamentares varia bastante – muitas vezes diferentes deputados propõem o mesmo projeto utilizando termos completamente diferentes. Para isso, é necessário identificar automaticamente essas similaridades para evitar retrabalho e somar conhecimento a partir de visões complementares. A partir daí, optou-se por gerar respostas automáticas a serem enviadas aos gabinetes, notificando no caso de existência de proposições legislativas similares. É importante ressaltar que o deputado pode discordar, mas para fins de registro a comunicação é documentada.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- **Páginas temáticas**: classificação automática de todo o conteúdo digital disponível no sítio eletrônico da Câmara dos Deputados em 14 (catorze) temas pré-definidos. De acordo com material suplementar fornecido pela equipe entrevistada, "cada página temática apresenta todo o conteúdo, como enquetes eletrônicas, notícias, projetos de leis, discursos, áudios, vídeos, legislação e comissões" relacionado ao respectivo tema, facilitando a navegabilidade por parte do usuário interessado no tema específico. A classificação é feita por meio de aprendizagem de máquina supervisionada, e há suporte a aprendizado ativo[1], em que erros de classificação podem ser corrigidos por profissionais especializados, retroalimentando o treinamento do algoritmo.

A equipe entrevistada revelou que esse tipo de plataforma digital requer um programa de governança de dados, considerado por eles como pré-requisito essencial para o desenvolvimento de uma IA confiável.

Para a implementação das soluções, foi estabelecido convênio com a Universidade de São Paulo (USP) para a elaboração dos modelos que envolvessem processamento de linguagem natural. São utilizadas bases textuais que incluem projetos de lei e discursos de parlamentares. Na gestão do convênio com a USP, ocorrem reuniões semanais para transferência de conhecimento, onde também são envolvidos gestores de negócio.

Sobre segurança dos dados de treinamento, a equipe informou a existência de aditivos contratuais que prescrevem que a base de treinamento só pode utilizar projetos de lei que já tenham sido publicados, e que os dados precisam ser anonimizados para eliminar menções a nomes de deputados, consultores, números telefônicos etc. Atualmente, essa anonimização é feita por meio de expressões regulares. A avaliação da efetividade da anonimização é feita pelas consultorias por meio de amostragem, que vem demonstrando que, até o momento, o pré-processamento implementado tem sido suficiente. Os dados precisam ser processados em um ambiente isolado na infraestrutura presente na Câmara dos Deputados, sem conexão com a internet.

A equipe relatou algumas dificuldades nesse tipo de adoção tecnológica, dentre elas:

- Receio por parte dos parlamentares de que os robôs tomassem seus lugares na representação perante os cidadãos. Para mitigar esse tipo de receio, os algoritmos são divididos em dois grupos: aqueles voltados para a sociedade – que em nenhuma hipótese substituem a autonomia parlamentar – e aqueles voltados para a esfera administrativa do órgão. Apenas neste último grupo ciclos de ação podem ser completados sem intervenção humana.
- Alta curva de aprendizagem na formação de pessoal, o que requer negociação constante por patrocínio.
- Manutenção de sistemas de IA concorrendo com manutenção de sistemas transacionais, em um cenário de escassez de recursos humanos, e considerando ainda que os primeiros são mais difíceis de manter do que os segundos.

[1] Aprendizado ativo ou *Active learning*: https://en.wikipedia.org/wiki/Active_learning_(machine_learning)

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE IX – Organizações participantes da pesquisa

ABDI	Agência Brasileira de Desenvolvimento Industrial
ABGF	Agência Brasileira Gestora de Fundos Garantidores e Garantias S.A.
ABIN	Agência Brasileira de Inteligência
AEB	Agência Espacial Brasileira
AGU	Advocacia-Geral da União
AMAZUL	Amazônia Azul Tecnologias de Defesa S.A.
AmGT	Amazonas Geração e Transmissão de Energia S.A.
ANA	Agência Nacional de Águas
ANAC	Agência Nacional de Aviação Civil
ANATEL	Agência Nacional de Telecomunicações
ANATER	Agência Nacional de Assistência Técnica e Extensão Rural
ANCINE	Agência Nacional do Cinema
ANEEL	Agência Nacional de Energia Elétrica
ANM	Agência Nacional de Mineração
ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
ANS	Agência Nacional de Saúde Suplementar
ANTAQ	Agência Nacional de Transportes Aquaviários
ANTT	Agência Nacional de Transportes Terrestres
ANVISA	Agência Nacional de Vigilância Sanitária
APEX Brasil	Agência Brasileira de Promoção de Exportações e Investimentos
APS	Associação das Pioneiras Sociais
BANCO DA AMAZÔNIA	Banco da Amazônia S.A.
BCB	Banco Central do Brasil
BN	Fundação Biblioteca Nacional
BNB	Banco do Nordeste do Brasil S.A.
BNDES	Banco Nacional de Desenvolvimento Econômico e Social
CADE	Conselho Administrativo de Defesa Econômica
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CBTU	Companhia Brasileira de Trens Urbanos
CDC	Companhia Docas do Ceará
CDP	Companhia Docas do Pará
CDRJ	Companhia Docas do Rio de Janeiro
CEASAMINAS	Centrais de Abastecimento de Minas Gerais S.A.
CEFET-MG	Centro Federal de Educação Tecnológica de Minas Gerais
CEFET-RJ	Centro Federal de Educação Tecnológica Celso Suckow da Fonseca
CEITEC	Centro Nacional de Tecnologia Eletrônica Avançada S.A
CEX	Comando do Exército
CGT Eletrosul	Comp. de Geração E Transmissão de Energia Elétrica do Sul do Brasil
CGU	Controladoria-Geral da União
CHESF	Companhia Hidro Elétrica do São Francisco
CJF	Conselho da Justiça Federal
CM	Comando da Marinha
CMB	Casa da Moeda do Brasil
CNEN	Comissão Nacional de Energia Nuclear
CNJ	Conselho Nacional de Justiça
CNMP	Conselho Nacional do Ministério Público
CNPQ	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CODERN	Companhia Docas do Rio Grande do Norte
CODESA	Companhia Docas do Espírito Santo

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

CODEVASF	Companhia de Desenvolvimento dos Vales do São Francisco e do Parnaíba
COMAER	Comando da Aeronáutica
CONAB	Companhia Nacional de Abastecimento
Conceição	Hospital Nossa Senhora da Conceição S.A.
CPII	Colégio Pedro II
CPRM	Companhia de Pesquisa de Recursos Minerais
CSJT	Conselho Superior da Justiça do Trabalho
CVM	Comissão de Valores Mobiliários
DATAPREV	Empresa de Tecnologia e Informações da Previdência Social
DNIT	Departamento Nacional de InfraEstrutura de Transportes
DNOCS	Departamento Nacional de Obras Contra as Secas
DPU	Defensoria Pública da União
EBSERH	Empresa Brasileira de Serviços Hospitalares
ECT	Empresa Brasileira de Correios e Telégrafos
ELETROBRAS	Centrais Elétricas Brasileiras S.A.
ELETRONORTE	Centrais Elétricas do Norte do Brasil S.A.
ELETRONUCLEAR	Eletrobrás Termonuclear S.A.
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
EMGEPRON	Empresa Gerencial de Projetos Navais
ENAP	Fundação Escola Nacional de Administração Pública
EPE	Empresa de Pesquisa Energética
FAZENDA	Secretaria Especial da Fazenda
FCP	Fundação Cultural Palmares
FINEP	Financiadora de Estudos e Projetos
FIOCRUZ	Fundação Oswaldo Cruz
FNDE	Fundo Nacional de Desenvolvimento da Educação
FUB	Fundação Universidade de Brasília
FUNAG	Fundação Alexandre de Gusmão
FUNAI	Fundação Nacional do Índio
FUNARTE	Fundação Nacional de Artes
FUNASA	Fundação Nacional de Saúde
Fundacentro	Fundação Jorge Duprat Figueiredo de Segur. e Medicina do Trabalho
FURG	Fundação Universidade Federal do Rio Grande
FURNAS	Furnas Centrais Elétricas S.A.
Ibama	Instituto Brasileiro do Meio Amb. e dos Recursos Naturais Renováveis
IBGE	Fundação Instituto Brasileiro de Geografia e Estatística
ICMBIO	Instituto Chico Mendes de Conservação da Biodiversidade
IF Baiano	Instituto Federal de Educação, Ciência e Tecnologia Baiano
IF Farroupilha	Instituto Federal de Educação, Ciência e Tecnologia Farroupilha
IF Goiano	Instituto Federal de Educação, Ciência e Tecnologia Goiano
IF Sertão-PE	Instituto Federal de Educação, Ciência e Tecnologia do Sertão Pernambucano
IF Sudeste	Instituto Federal de Educ. Ciência e Tec. do Sudeste de Minas Gerais
IFAC	Instituto Federal de Educação, Ciência e Tecnologia do Acre
IFAL	Instituto Federal de Educação, Ciência e Tecnologia de Alagoas
IFAM	Instituto Federal de Educação, Ciência e Tecnologia do Amazonas
IFAP	Instituto Federal de Educação, Ciência e Tecnologia do Amapá
IFB	Instituto Federal de Educação, Ciência e Tecnologia de Brasília
IFBA	Instituto Federal de Educação, Ciência e Tecnologia da Bahia
IFC	Instituto Federal de Educação, Ciência e Tecnologia Catarinense
IFCE	Instituto Federal de Educação, Ciência e Tecnologia do Ceará
IFES	Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo
IFF	Instituto Federal de Educação, Ciência e Tecnologia Fluminense

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

IFG	Instituto Federal de Educação, Ciência e Tecnologia de Goiás
IFMA	Instituto Federal de Educação, Ciência e Tecnologia do Maranhão
IFMG	Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais
IFMS	Instituto Federal de Educação, Ciência e Tecnologia do Mato Grosso do Sul
IFMT	Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
IFNMG	Instituto Federal de Educação, Ciência e Tecnologia do Norte de Minas Gerais
IFPA	Instituto Federal de Educação, Ciência e Tecnologia do Pará
IFPB	Instituto Federal de Educação, Ciência e Tecnologia da Paraíba
IFPE	Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco
IFPI	Instituto Federal de Educação, Ciência e Tecnologia do Piauí
IFPR	Instituto Federal de Educação, Ciência e Tecnologia do Paraná
IFRJ	Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro
IFRN	Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte
IFRO	Instituto Federal de Educação, Ciência e Tecnologia de Rondônia
IFRR	Instituto Federal de Educação, Ciência e Tecnologia de Roraima
IFRS	Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul
IFS	Instituto Federal de Educação, Ciência e Tecnologia de Sergipe
IFSC	Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina
IFSP	Instituto Federal de Educação, Ciência e Tecnologia de São Paulo
IFSul	Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense
IFSULDEMINAS	Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais
IFTM	Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro
IFTO	Instituto Federal de Educação, Ciência e Tecnologia do Tocantins
INB	Indústrias Nucleares do Brasil S.A.
INCRA	Instituto Nacional de Colonização e Reforma Agrária
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
INFRAERO	Empresa Brasileira de Infraestrutura Aeroportuária
INMETRO	Instituto Nacional de Metrologia, Qualidade e Tecnologia
INPI	Instituto Nacional da Propriedade Industrial
INSS	Instituto Nacional do Seguro Social
IPEA	Fundação Instituto De Pesquisa Econômica Aplicada
IPHAN	Instituto do Patrimônio Histórico e Artístico Nacional
ITI	Instituto Nacional de Tecnologia da Informação
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
MC	Ministério da Cidadania
Mcom	Ministério Das Comunicações
MCTI	Ministério da Ciência, Tecnologia e Inovações
MD	Ministério da Defesa
MDR	Ministério do Desenvolvimento Regional
ME	Ministério da Economia
MEC	Ministério da Educação
MINFRA	Ministério da Infraestrutura
MJSP	Ministério da Justiça e Segurança Pública
MMA	Ministério do Meio Ambiente
MME	Ministério de Minas e Energia
MMFDH	Ministério da Mulher, da Família e dos Direitos Humanos
MPU	Ministério Público da União
MRE	Ministério das Relações Exteriores
MS	Ministério da Saúde
MTur	Ministério do Turismo
PETROBRAS	Petróleo Brasileiro S.A.
PF	Polícia Federal

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

PGFN	Procuradoria-Geral da Fazenda Nacional
PR	Presidência da República
PREVIC	Superintendência Nacional de Previdência Complementar
PRF	Polícia Rodoviária Federal
RFB	Secretaria Especial da Receita Federal do Brasil
SENAI-CETIQT	Centro de Tecnologia da Indústria Química e Têxtil do Senai/RJ
SPA	Autoridade Portuaria de Santos S.A
STM	Superior Tribunal Militar
SUDAM	Superintendência de Desenvolvimento da Amazônia
SUDECO	Superintendência do Desenvolvimento do Centro-Oeste
SUDENE	Superintendência de Desenvolvimento do Nordeste
SUFRAMA	Superintendência da Zona Franca de Manaus
SUSEP	Superintendência de Seguros Privados
Telebras	Telecomunicações Brasileiras S.A.
TJDFT	Tribunal de Justiça do Distrito Federal e dos Territórios
TRANSPETRO	Petrobras Transporte S.A.
TRE-AC	Tribunal Regional Eleitoral do Acre
TRE-AL	Tribunal Regional Eleitoral de Alagoas
TRE-AM	Tribunal Regional Eleitoral do Amazonas
TRE-AP	Tribunal Regional Eleitoral do Amapá
TRE-BA	Tribunal Regional Eleitoral da Bahia
TRE-CE	Tribunal Regional Eleitoral do Ceará
TRE-DF	Tribunal Regional Eleitoral do Distrito Federal
TRE-ES	Tribunal Regional Eleitoral do Espírito Santo
TRE-GO	Tribunal Regional Eleitoral de Goiás
TRE-MA	Tribunal Regional Eleitoral do Maranhão
TRE-MG	Tribunal Regional Eleitoral de Minas Gerais
TRE-MS	Tribunal Regional Eleitoral do Mato Grosso do Sul
TRE-MT	Tribunal Regional Eleitoral do Mato Grosso
TRE-PA	Tribunal Regional Eleitoral do Pará
TRE-PB	Tribunal Regional Eleitoral da Paraíba
TRE-PE	Tribunal Regional Eleitoral de Pernambuco
TRE-PI	Tribunal Regional Eleitoral do Piauí
TRE-PR	Tribunal Regional Eleitoral do Paraná
TRE-RJ	Tribunal Regional Eleitoral do Rio de Janeiro
TRE-RN	Tribunal Regional Eleitoral do Rio Grande do Norte
TRE-RO	Tribunal Regional Eleitoral de Rondônia
TRE-RR	Tribunal Regional Eleitoral de Roraima
TRE-RS	Tribunal Regional Eleitoral do Rio Grande do Sul
TRE-SC	Tribunal Regional Eleitoral de Santa Catarina
TRE-SE	Tribunal Regional Eleitoral do Sergipe
TRE-SP	Tribunal Regional Eleitoral de São Paulo
TRE-TO	Tribunal Regional Eleitoral do Tocantins
TRF1	Tribunal Regional Federal da 1ª Região
TRF2	Tribunal Regional Federal da 2ª Região
TRF3	Tribunal Regional Federal da 3ª Região
TRF4	Tribunal Regional Federal da 4ª Região
TRF5	Tribunal Regional Federal da 5ª Região
TRT1	Tribunal Regional do Trabalho da 1ª Região/Rj
TRT10	Tribunal Regional do Trabalho da 10ª Região/Df E To
TRT11	Tribunal Regional do Trabalho da 11ª Região/Am E Rr
TRT12	Tribunal Regional do Trabalho da 12ª Região/Sc

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

TRT13	Tribunal Regional do Trabalho da 13ª Região/Pb
TRT14	Tribunal Regional do Trabalho da 14ª Região/Ac E Ro
TRT15	Tribunal Regional do Trabalho da 15ª Região - Campinas/Sp
TRT16	Tribunal Regional do Trabalho da 16ª Região/Ma
TRT17	Tribunal Regional do Trabalho da 17ª Região/Es
TRT18	Tribunal Regional do Trabalho da 18ª Região/Go
TRT19	Tribunal Regional do Trabalho da 19ª Região/Al
TRT2	Tribunal Regional do Trabalho da 2ª Região/Sp
TRT20	Tribunal Regional do Trabalho da 20ª Região/Se
TRT21	Tribunal Regional do Trabalho da 21ª Região/Rn
TRT22	Tribunal Regional do Trabalho da 22ª Região/Pi
TRT23	Tribunal Regional do Trabalho da 23ª Região/Mt
TRT24	Tribunal Regional do Trabalho da 24ª Região/Ms
TRT3	Tribunal Regional do Trabalho da 3ª Região/Mg
TRT4	Tribunal Regional do Trabalho da 4ª Região/Rs
TRT5	Tribunal Regional do Trabalho da 5ª Região/Ba
TRT6	Tribunal Regional do Trabalho da 6ª Região/Pe
TRT7	Tribunal Regional do Trabalho da 7ª Região/Ce
TRT8	Tribunal Regional do Trabalho da 8ª Região/Pa E Ap
TRT9	Tribunal Regional do Trabalho da 9ª Região/Pr
TSE	Tribunal Superior Eleitoral
TST	Tribunal Superior do Trabalho
UFABC	Fundação Universidade Federal do ABC
UFAC	Fundação Universidade Federal do Acre
UFAL	Universidade Federal de Alagoas
UFAM	Fundação Universidade do Amazonas
UFBA	Universidade Federal da Bahia
UFC	Universidade Federal do Ceará
UFCA	Universidade Federal do Cariri
UFCG	Universidade Federal de Campina Grande
UFCSPA	Fundação Universidade Federal de Ciências da Saúde de Porto Alegre
UFERSA	Universidade Federal Rural do Semiárido
UFES	Universidade Federal do Espírito Santo
UFF	Universidade Federal Fluminense
UFFS	Universidade Federal da Fronteira Sul
UFG	Universidade Federal de Goiás
UFGD	Fundação Universidade Federal da Grande Dourados
UFJF	Universidade Federal de Juiz de Fora
UFLA	Universidade Federal de Lavras
UFMA	Fundação Universidade Federal do Maranhão
UFMG	Universidade Federal de Minas Gerais
UFMS	Fundação Universidade Federal de Mato Grosso do Sul
UFMT	Fundação Universidade Federal de Mato Grosso
UFOB	Universidade Federal do Oeste da Bahia
UFOP	Fundação Universidade Federal de Ouro Preto
UFOPA	Universidade Federal do Oeste do Pará
UFPA	Universidade Federal do Pará
UFPB	Universidade Federal da Paraíba
UFPE	Universidade Federal de Pernambuco
UFPel	Universidade Federal de Pelotas
UFPI	Fundação Universidade Federal do Piauí
UFPR	Universidade Federal do Paraná

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

UFRA	Universidade Federal Rural da Amazônia
UFRB	Universidade Federal do Recôncavo da Bahia
UFRGS	Universidade Federal do Rio Grande do Sul
UFRJ	Universidade Federal do Rio de Janeiro
UFRN	Universidade Federal do Rio Grande do Norte
UFRPE	Universidade Federal Rural de Pernambuco
UFRR	Universidade Federal de Roraima
UFRRJ	Universidade Federal Rural do Rio de Janeiro
UFS	Fundação Universidade Federal de Sergipe
UFSB	Universidade Federal do Sul da Bahia
UFSC	Universidade Federal de Santa Catarina
UFSCar	Fundação Universidade Federal de São Carlos
UFSJ	Fundação Universidade Federal de São João Del Rei
UFSM	Universidade Federal de Santa Maria
UFT	Fundação Universidade Federal do Tocantins
UFTM	Universidade Federal do Triângulo Mineiro
UFU	Fundação Universidade Federal de Uberlândia
UFV	Fundação Universidade Federal de Viçosa
UFVJM	Universidade Federal dos Vales do Jequitinhonha e Mucuri
UNIFAL-MG	Universidade Federal de Alfenas
UNIFAP	Fundação Universidade Federal do Amapá
UNIFEI	Universidade Federal de Itajubá
UNIFESP	Universidade Federal de São Paulo
UNIFESSPA	Universidade Federal do Sul e Sudeste do Pará
UNILA	Universidade Federal da Integração Latino-Americana
UNILAB	Universidade da Integração Internacional da Lusofonia Afro-Brasileira
UNIPAMPA	Fundação Universidade Federal do Pampa
UNIR	Fundação Universidade Federal de Rondônia
UniRIO	Universidade Federal do Estado do Rio de Janeiro
Univasf	Fundação Universidade Federal do Vale do São Francisco
UTFPR	Universidade Tecnológica Federal do Paraná
VALEC	Valec Engenharia, Construções E Ferrovias S.A. (Filial Rj)

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE X – Consolidação das informações referentes às 36 soluções de IA que foram detalhadas nas respostas ao questionário eletrônico

1 - Qual é o escopo pretendido da solução de IA?	Total
Acesso público	5
Compartilhado entre órgãos pré-definidos	9
Interno à instituição	22
2 - Foi realizado estudo de viabilidade financeira ou análise de custo-benefício referente a despesas com mão-de-obra, consumo energético, licenciamento ou contratação em geral?	**Total**
Sim	8
Não	28
3 - Em qual fase a solução se encontra atualmente?	**Total**
Em ambiente de experimentação com usuário final (ex.: projeto-piloto, teste A/B)	4
Em ambiente de experimentação com usuário final (ex.: teste A/B)	2
Em desenvolvimento	7
Em produção	21
Em teste/homologação pelo gestor/solicitante	2
4 - A solução envolveu quais modalidades de desenvolvimento?	**Total**
acordo de cooperação com outras instituições	9
expertise de colaboradores da própria instituição	28
mão-de-obra terceirizada	8
5 - A solução envolveu quais modalidades de contratação?	**Total**
Aquisição de licenças de software	3
Encomenda tecnológica	2
Infraestrutura de nuvem	6
Hardware	1
Não envolveu licitação ou contratação	25
6 - Foram implementados mecanismos para automação de obtenção de dados, treinamento e deploy de modelos e monitoramento?	**Total**
Sim	23
Não	13
7 - Houve algum tipo de compartilhamento com outras instituições, considerando o contexto da solução de Inteligência Artificial desenvolvida?	**Total**
Bases de dados	8
Código-fonte	12
Ferramentas	5
Infraestrutura	4
Modelo treinado	6

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

8 - Qual é o tamanho médio mensal da equipe de colaboradores envolvidos no desenvolvimento da solução?	Total
De 1 a 5	31
De 6 a 10	2
Acima de 10	3
9 - Quantos meses foram consumidos para o desenvolvimento da solução?	**Total**
Até 6 meses	18
De 6 a 12 meses	3
de 11 a 24 meses	14
Acima de 24 meses	1
10 - Qual é o grau esperado de automação da solução?	**Total**
A solução toma uma decisão pendente de autorização/homologação/modificação por um ser humano	6
A solução toma uma decisão pendente de autorização/homologação/modificação por um ser humano	15
Não se aplica - a solução não faz análise prescritiva nem direciona decisões	6
Outros	9
11 - Quais são as fontes de dados utilizadas para desenvolvimento da solução?	**Total**
Dados da própria organização	31
Dados abertos de outras organizações públicas	6
Dados obtidos de outras organizações públicas mediante acordo prévio	5
Dados provenientes de organizações do terceiro setor	0
Dados provenientes de organizações privadas	1
Dados sintéticos / data augmentation	2
11 - Qual é a classificação quanto ao sigilo dos dados utilizados no desenvolvimento do modelo de Inteligência Artificial?	**Total**
Não públicos	23
Públicos	11
Não se aplica - foram utilizados somente modelos prontos	2
12 - Foram consideradas as questões relativas à Lei Geral de Proteção de Dados no desenvolvimento da solução?	**Total**
Sim	13
Não	11
Não se aplica	12
13 - Foi feito algum tratamento de anonimização e/ou mascaramento para a análise de dados classificados?	**Total**
Sim	15
Não	10
Não se aplica	11

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

14 - Foi feita análise de risco relativo a vieses e/ou fairness em dados detreinamento?	Total
Sim	15
Não	8
Não se aplica	13
15 - Foram adotados controles para garantir a representativvidade de diferentes populações na base de treinamento?	Total
Sim	16
Não	8
Não se aplica	12
16 - A qualidade dos dados utilizados para treinamento foi avaliada?	Total
Sim	30
Não	6
17 - Foi feito planejamento e implementação para detecção automatizada de diminuição do desempenho do modelo desenvolvido em decorrência de mudanças no padrão dos dados (model drift)?	Total
Sim	7
Não	29
18 - A implementação utiliza modelos de Inteligência Artificial:	Total
Pré-treinados	10
Com treinamento/fine-tuning adicional	26
19 - Caso tenha usado modelo em nuvem, qual foi o tipo de contratação?	Total
Contratação direta de máquinas virtuais com CPUs/GPUs	4
Soluções com infraestrutura pré-configurada, com maior valor agregado e mais prontas para processar os dados (Ex.: Spark, AutoML, etc.)	2
Não se aplica	30
20 - Foram implementados controles referentes à segurança do modelo de Inteligência Artificial?	Total
Sim - em relação a ataques adversariais (ex.: exemplo construído sob medida para levar o algoritmo ao erro, ainda que de modo imperceptível por um humano)	1
Sim - em relação a envenenamento de dados (ex.: usuário fornecem dados errôneos para levar o algoritmo ao erro)	3
Sim - outras técnicas de segurança	6
Não	26
21 - Foi realizada análise de riscos referente ao uso de modelos deInteligência Artificial não interpretáveis ou não transparentes (ex.:modelos “caixa-preta”)?	Total
Sim	9
Não	20
Não se aplica	7

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

22 - Foi feita alguma auditoria interna ou externa do modelo?	Total
Sim	8
Não	28

23 - O processo de desenvolvimento da solução considerou a importância relativa dos diferentes tipos de erro (ex.: falsos positivos / falsos negativos)?	Total
Sim	26
Não	10

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

APÊNDICE XI – Exemplos de aplicações de IA no Controle

1. TCU - GEOCONTROLE

GEOCONTROLE – Detecção de objetos em imagens de satélite			
Objetivo		**Principais funcionalidades**	
Monitoramento/acompanhamento remoto de obras		✓ *Reconhecimento de minas e barragens a partir de análise das imagens de satélites* ✓ *Verificação da regularidade da exploração de minas*	
Tipo de aprendizado	**Entradas**	**Saídas**	**Origem dos dados**
Supervisionado	*Imagens de satélites* *Dados de barragens*	✓ *Coordenadas geográficas de minas detectadas nas imagens*	*Satélites Sentinel* *Agência Nacional de Águas*
Algoritmos e técnicas		**Etapa**	
Google Earth Engine ***Tensorflow*** ***Redes neurais convolucionais***		*Prova de conceito*	

Síntese apresentação geral da solução

A solução GeoControle tem por objetivo a aplicação de tecnologias de IA para detecção automática de objetos específicos a partir da análise de imagens captadas por satélites. O objetivo é promover a fiscalização remota de obras dispersas geograficamente, como, por exemplo, a construção de rodovias e ferrovias.

Desafio de negócio

A solução nasceu da necessidade que o TCU teve quando da fiscalização das obras da ferrovia Transnordestina. Inicialmente, houve a aquisição de imagens de satélites de alta resolução e o apoio técnico de especialistas externos em Geografia, os quais, por meio da inspeção visual nas imagens, identificaram vários achados de auditoria - objetos reconhecidos ("obras de arte" de engenharia civil) e irregularidades (depósitos de material, sinais de erosão) ao longo do percurso das obras.

Devido ao grande esforço e tempo desprendido no processo de análise das imagens, desenvolveu-se solução que buscou utilizar técnicas de geoprocessamento e visão computacional sobre imagens de satélite para acompanhar objetos de controle. Em convênio com a Universidade Federal de Minas Gerais, desenvolveu-se a referida solução.

Atualmente, a aplicação desenvolvida reconhece minas de exploração mineral, superando limitações de acesso às regiões com lavra ilegal, realizando essa identificação em larga escala, a baixo custo e podendo reatualizar a evolução das minas com alguma frequência, sem trabalho adicional.

Funcionalidades da solução

Atualmente, a solução desenvolvida encontra-se em prova de conceito, com as seguintes funcionalidades implementadas:

- Reconhecimento de minas e barragens a partir de análise das imagens de satélites dos estados do Pará e Minas Gerais;
- Desenvolvimento de modelo que acompanha em qual estado está determinada mina, por meio do cruzamento com as bases da Agência Nacional de Mineração (minas oficialmente registradas), detectando irregularidades de empresas que têm autorização apenas para pesquisa, mas já estão em fase de exploração.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Scripts reutilizáveis para baixa e processamento de imagens do Google Earth e para treinamento do modelo em nuvem.

Tipo de aprendizado

Supervisionado.

Detalhe tipo de aprendizado

O processo de aprendizado compreendeu a realização de anotações manuais (bounding boxes que identificavam as minas). Também foi necessário registrar em uma base de dados as "não minas" (situações similares a minas e que não são minas, como rios e lagos que podem ser confundidos com as barragens)

Entrada do modelo

- Conjunto de imagens por satélite de 2 km x 2 km, com 12 bandas espectrais;

- Base de barragens da Agência Nacional de Águas.

Saída do modelo

O modelo aponta locais onde foram encontradas minas, posteriormente convertidos em coordenadas geográficas.

Formato utilizado dos dados

Imagens TIFF (padrão de imagens de satélite para geotecnologias)

Tipo de armazenamento

Nuvem (Google Earth Engine + Google Drive)

Origem dos dados de treinamento

Dados públicos da Agência Nacional de Águas (ANA);

Cadastro Ambiental Rural

Imagens públicas e gratuitas dos satélites Sentinel (Agência Espacial Europeia).

Algoritmos e técnicas

Foi utilizada a API do Google Earth Engine em Python, além do Tensorflow para definir uma arquitetura própria de redes neurais convolucionais.

Utilizou-se também o ambiente do Google Colab, o que permitiu o desenvolvimento dos modelos sem a necessidade de investimentos em infraestrutura.

Quantas pessoas estiveram envolvidas (e seus papéis)?

1 colaborador em tempo integral e 2 colaboradores em tempo parcial.

Duração do projeto de desenvolvimento

Cinco meses.

Benefícios

Uma vez que solução não foi implementada em produção, o principal benefício foram os conhecimentos internalizados pela equipe do TCU referentes à utilização das tecnologias de IA no georreferenciamento.

Contatos

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

soma@tcu.gov.br

2. TCU - ALICE

ALICE – Análise de Licitações e Editais			
Objetivo		**Principais funcionalidades**	
Monitoramento de licitações de forma automatizada		✓ *Divulgação diária por e-mail de editais publicados* ✓ *Identificação automática de materialidade das licitações* ✓ *Ordenação de licitações por materialidade* ✓ *Identificação de padrões textuais que revelem indícios de irregularidades* ✓ *Geração de alertas a partir de cruzamentos de dados* ✓ *Classificação por tipo de objeto* ✓ *Distribuição temática* ✓ *Busca textual*	
Tipo de aprendizado	**Entradas**	**Saídas**	**Origem dos dados**
Supervisionado	*Parágrafos dos editais*	✓ *Materialidade das contratações* ✓ *Trechos com possíveis indícios de irregularidades* ✓ *Classificação dos objetos de contratação por tipo*	*Siasg* *Comprasnet*
Algoritmos e técnicas		**Etapa**	
Random forest ***Bernoulli Naïve Bayes*** ***Web scraping***		*Em produção*	

Síntese apresentação geral da solução

O ALICE, acrônimo para Análise de Licitações e Editais, é um sistema originalmente desenvolvido pela Controladoria Geral da União (CGU) em 2014, cujo principal objetivo é permitir o monitoramento das licitações de forma automatizada. O código foi cedido ao Tribunal de Contas da União, onde seguiu fluxo independente de desenvolvimento, evolução e implantação em relação à versão da CGU que, em uma ramificação (*fork*) de desenvolvimento paralela, realiza as manutenções necessárias às necessidades daquele órgão.

Funcionalidades da solução

Atualmente, a solução conta com as seguintes funcionalidades:

- Divulgação diária por e-mail dos editais publicados nos respectivos dias.
- Identificação automática da materialidade das licitações, uma vez que tal informação não se encontra estruturada no Comprasnet[1]. Foi construído um reconhecedor de entidades para identificar o valor estimado.
- Ranqueamento das licitações conforme materialidade.
- Identificação de padrões de texto dentro dos documentos que indiquem possíveis indícios de irregularidade, com geração de alertas para posterior análise pelo auditor.

[1] Comprasnet: https://www.gov.br/compras/pt-br/sistemas/comprasnet-siasg

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Geração de alertas a partir do cruzamento de dados baseados em metadados dos licitantes (ex.: impedimentos legais para participação em certames).
- Classificadores de objetos de contratação (ex.: contratação de obra pública ou contratação de tecnologia de informação etc.)
- Distribuição das licitações que geraram alertas às secretarias temáticas interessadas.
- Alice Nacional: inclui outras unidades da Federação.
- Busca textual.

Tipo de aprendizado

Supervisionado.

Detalhe tipo de aprendizado

Para a rotulagem dos dados necessária ao módulo extrator de materialidade, foi utilizada a base do Siasg[1] – que contém valor estimado e valor global – e a classificação é feita ao nível de parágrafos do texto.

Já para a rotulagem dos dados necessária aos módulos classificadores de tipo de objeto, foi utilizada como valor-alvo a natureza da despesa, informação também disponível no SIASG, além de bases históricas rotuladas, disponíveis na Secretaria de Fiscalização de Tecnologia da Informação (Sefti).

Entrada do modelo

Os textos dos parágrafos dos documentos a serem analisados são representados com vetorização em bag-of-words[2]. Outras variáveis utilizadas no treinamento e inferência dos modelos incluem adicionalmente heurísticas dos parágrafos, como por exemplo posição no documento.

Saída do modelo

- Materialidade das contratações (estimativa global, por item etc.).
- Trechos do texto com possíveis indícios de irregularidade.
- Classificação dos objetos em tipo (ex.: obra, TI), apresentando de forma ordenada por probabilidade de classificação.

Formato utilizado dos dados

Textual (PDF, RTF etc.), assumindo que todas as informações estão pré-digitalizadas ou em formato nativamente digital.

Tipo de armazenamento

Índice textual para busca (textos), banco de dados (metadados) e sistema de arquivos (arquivos originais gravados em disco).

Origem dos dados de treinamento

- Sistema Integrado de Administração de Serviços Gerais (Siasg)
- Comprasnet

Algoritmos e técnicas

- *Random forest* para o cálculo de materialidade

[1] Siasg: https://www.gov.br/compras/pt-br/acesso-a-informacao/institucional/siasg

[2] Bag-of-words model: https://en.wikipedia.org/wiki/Bag-of-words_model

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Bernoulli Naive Bayes para classificação de tipos de objeto de contratação.
- Técnica de *Web scraping*[1] para coletar os dados automaticamente a partir do sítio do Comprasnet, bem como de documentos (ex.: editais). Essa etapa é realizada no dia da publicação de cada edital, sendo agendada para executar diariamente. A carga de informações precisou ser feita por meio de *web scraping* porque o Comprasnet não possui interface de programação que disponibilize os dados de maneira tempestiva.

Benefícios

- Expectativa de controle – gestores começaram a tomar cuidados adicionais para que editais não fossem enquadrados nas regras de identificação de indícios empregadas pelo Alice.
- Emprego de mecanismos tecnológicos que classificam objetos de controle para facilitar o encaminhamento e distribuição dos trabalhos entre os auditores.
- Ganhos em produtividade e satisfação dos usuários.

Fontes utilizadas para construir a solução

A versão inicialmente desenvolvida pela CGU foi utilizada como ponto de partida.

Contatos

soma@tcu.gov.br

3. TCU – AGATA

AGATA – Aplicação Geradora de Análise Textual com Aprendizado			
Objetivo		**Principais funcionalidades**	
Prover ao usuário final uma ferramenta de rotulagem de informações		✓ *Aprendizado contínuo a partir do feedback do usuário* ✓ *Integração com o ALICE*	
Tipo de aprendizado	**Entradas**	**Saídas**	**Origem dos dados**
Supervisionado *Não supervisionado*	*Fornecidas pelo usuário*	*Insumos para treinamento de novos modelos pelo ALICE*	*SIASG* *Comprasnet*
Algoritmos e técnicas		**Etapa**	
Regressão logística *Clusterização hierárquica aglomerativa* *Aprendizado ativo*		*Em produção*	

Síntese apresentação geral da solução

O AGATA é um módulo à parte do ALICE, cujo objetivo é prover uma ferramenta de rotulagem de informações. Quando uma pesquisa é feita na base de informações coletadas pelo ALICE, o usuário classifica quais resultados da busca são relevantes, realimentando o motor de inferência da ferramenta, empregando, portanto, a técnica de aprendizado ativo.

Funcionalidades da solução

- O sistema aprende com a classificação fornecida pelo usuário. Como exemplo, o próprio usuário destaca os trechos dos textos das licitações que podem ser classificados como indícios de determinados tipos de irregularidades.

[1] Web scraping: https://en.wikipedia.org/wiki/Web_scraping

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Integração com o ALICE. A base rotulada pelo usuário é enviada ao ALICE, que treina um novo modelo, que passará então a ser aplicado às licitações em suas futuras análises de possíveis indícios de irregularidades.
- Envio de e-mail ao usuário para que ele valide os resultados das análises, proporcionando assim mais um ciclo de *feedback*.

Tipo de aprendizado

Supervisionado e não supervisionado.

Entrada do modelo

Os mesmos documentos utilizados pelo Alice (ex.: editais de licitação). Os rótulos-alvo para a aprendizagem supervisionada são fornecidos pelo usuário.

Algoritmos e técnicas

- Regressão logística – algoritmo com baixo tempo de resposta, pré-requisito essencial para melhor interação com o usuário.
- Agrupamento (*clustering*) hierárquico aglomerativo - para agrupar trechos muito semelhantes de textos.
- Aprendizado ativo, para incorporar a rotulagem realizada pelo usuário final ao treinamento dos modelos. A vantagem do emprego desta técnica no caso concreto é a superação dos limites de criação de expressões regulares, originalmente empregadas no ALICE. Assim, usuários especialistas de negócio, sem conhecimento de tecnologia de informação ou aprendizagem de máquina tornam-se capazes de criar seus próprios classificadores.

Benefícios

Permite escalar a construção de regras para identificação de possíveis irregularidades, tornando esse recurso disponível para usuários finais e consequentemente liberando a equipe de desenvolvimento da tarefa de construção de regras.

Fontes utilizadas para construir a solução

Literatura científica.

Dificuldades encontradas

Possibilidade de incluir vieses, devido à interação com o usuário final, corresponsável por prover os dados de treinamento para os modelos.

Contatos

soma@tcu.gov.br

4. TCU - COVIDATA

CoviData			
Objetivo		**Principais funcionalidades**	
Disponibilizar escores de risco para fornecedores que participaram de contratações relacionadas à aquisição de insumos para combate à pandemia de Covid-19		*Estimativa de risco de contratação de fornecedores envolvidos em certames voltados ao combate da pandemia*	
Tipo de aprendizado	**Entradas do modelo de IA**	**Saídas**	**Origem dos dados de treinamento**

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Supervisionado	✓ *Dados estruturados em bancos de dados relacionais* ✓ *Dados textuais em notícias da mídia*	*Notas de risco para cada fornecedor*	✓ *Bases de dados do TCU obtidas previamente por acordos de cooperação* ✓ *Google News para obtenção de notícias da mídia*
Algoritmos e técnicas		**Etapa**	
Algoritmos estatísticos de aprendizagem de máquina *Transferência de aprendizagem com algoritmo BERT*		*Projeto-piloto*	

Síntese apresentação geral da solução

O CoviData é um painel que disponibiliza escores de risco para fornecedores que participaram de contratações relacionadas à aquisição de insumos para combate à pandemia de Covid-19, a exemplo de respiradores, EPIs, construção de hospitais de campanha etc.

Funcionalidades da solução

Atualmente, a solução tem como principal funcionalidade a estimativa de risco de contratação de cada fornecedor envolvido em certames voltados ao combate da pandemia. Tal estimativa, contudo, utiliza como base uma grande quantidade de variáveis e heurísticas, cujos pesos podem inclusive ser definidos pelo usuário final. Dentre as heurísticas, estão regras que envolvem cruzamentos de dados entre várias bases com informações de contratações públicas, além de um módulo responsável por processar notícias da mídia relacionadas a contratações públicas e identificar entidades mencionadas nessas notícias (empresas, instituições públicas, nomes de pessoas e lugares).

Tipo de aprendizado

Supervisionado.

Detalhe tipo de aprendizado

Para a rotulagem da maior parte das informações que servem de insumo ao cálculo de risco de contratação, foram utilizadas bases históricas de contratações da administração pública federal previamente internalizadas pelo TCU, que permitem, por exemplo, identificar relacionamentos entre fornecedores, fornecedores declarados inidôneos etc.

Já a rotulagem dos dados necessários à construção do modelo de reconhecimento de entidades mencionadas em textos de notícias foi feita por meio de anotação manual de uma base com cerca de 200 notícias.

Entrada do modelo

Dados estruturados em bancos de dados relacionais e dados textuais contidos em notícias da mídia, sendo estes últimos representados por meio de vetores BERT (Devlin, Chang, Lee, & Toutanova, 2018).

Saída do modelo

A saída fornecida pelo painel CoviData inclui notas de risco para cada fornecedor.

Formato utilizado dos dados

Estruturado (banco de dados relacional) e textual, a depender da análise a ser implementada pelo sistema.

Tipo de armazenamento

Banco de dados (metadados) e sistema de arquivos (notícias da mídia).

Origem dos dados de treinamento

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Bases de dados do TCU obtidas previamente por meio de acordos de cooperação.
- Google News[1], para obtenção das notícias da mídia.

Algoritmos e técnicas

Foram empregados algoritmos clássicos de aprendizagem de máquina para o cálculo de risco.

O submódulo reconhecedor de entidades em notícias da mídia utiliza o algoritmo BERT (Devlin, Chang, Lee, & Toutanova, 2018), a partir de um modelo pré-treinado para a língua portuguesa (Souza, Nogueira, & Lotufo), que por sua vez recebeu uma etapa complementar de treinamento (*fine-tunning*) com a base de notícias. Trata-se do emprego da técnica de transferência de aprendizagem, que tem a vantagem de exigir menor quantidade de dados (ex.: menos de 200 notícias) e ainda assim obter resultados satisfatórios.

Benefícios

- Possibilidade de analisar riscos das contratações públicas voltadas ao combate à pandemia para subsidiar a atuação tanto das unidades técnicas do TCU responsáveis pelo acompanhamento das ações governamentais como de órgãos parceiros do TCU nos estados.
- Possibilidade de personalização de pesos de componentes e subcomponentes resultantes dos cruzamentos de dados de diversos sinais de alerta que formam o escore de risco de CNPJs.

Fontes utilizadas para construir a solução

Modelo de linguagem pré-treinado (Souza, Nogueira, & Lotufo) para o submódulo reconhecedor de entidades em notícias da mídia.

Contatos

soma@tcu.gov.br

sti@tcu.gov.br

5. TCU - PLACC

PLACC - Plataforma de Análises Cognitivas para o Controle			
Objetivo		**Principais funcionalidades**	
Extração de informações de relacionamento entre pessoas físicas e jurídicas em processos judiciais		✓ Detecção de pessoas físicas e jurídicas ✓ Enriquecimento de informações a partir de cruzamento de dados com a base de CPFs e CNPJs da Receita Federal ✓ Extração de relacionamentos entre pessoas físicas e jurídicas ✓ Análise de relacionamentos em grafo	
Tipo de aprendizado	**Entradas**	**Saídas**	**Origem dos dados**
Supervisionado	Textos das peças de processo	✓ Entidades nomeadas detectadas e *Part-of-speech tagging* de cada palavra	Processos da Justiça Federal e Planilhas de empreendimentos
Algoritmos e técnicas		**Etapa**	
Detecção de Entidades Nomeadas **Long short-term memory (LSTM)**		Em produção	

[1] Google News: https://en.wikipedia.org/wiki/Google_News

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Conditional Random Field (CRF)
Part-Of-Speech Tagging (POS-Tagging)
Regular Expression (Regex)

Síntese apresentação geral da solução

A Plataforma de Análises Cognitivas para o Controle (PLACC) é uma solução Web que faz busca em um grafo composto por pessoas físicas, jurídicas, contratos e projetos como vértices e suas relações como arestas. Essa pesquisa permite fazer análises centradas em entidades (vértices) e seu círculo de relação próximo, bem como filtrar por relações específicas dentro deste grupo informando palavras de interesse como propina, repasse de dinheiro, operadores e afins. Através da análise do grafo, é possível traçar fluxos de repasses de dinheiro entre vários atores e empresas descobrindo origem e destino e seus intermediários que tornaram o fluxo possível.

Funcionalidades da solução

Atualmente, a solução conta com as seguintes funcionalidades:

- Detecção de entidades nomeadas (pessoas físicas e jurídicas).
- Extração de informações que denotem alguma ação que envolveu essas pessoas.
- Geração de grafo para análise das ligações entre as pessoas.

Tipo de aprendizado

Supervisionado para a detecção de entidades nomeadas e *Part-of-speech tagging*.

Detalhe tipo de aprendizado

Para a predição de entidades nomeadas, foram usados rótulos sobre peças de processo judicial e do Tribunal de Contas da União.

Para *Part-of-speech tagging* foi usado um modelo pré-treinado em notícias em português.

Entrada do modelo

Os dois modelos usaram como entrada os textos das peças dos processos judiciais e alguns metadados semiestruturados disponíveis no site da justiça federal.

Saída do modelo

O detector de entidades nomeadas tem como saída as entidades detectadas

O gerador de *Part-of-speech tagging* tem como saída a classe gramatical da palavra, como verbo, substantivo, adjetivo, advérbio, preposição, entre outros.

Formato utilizado dos dados

Documentos em PDF, HTML, XLSX.

Tipo de armazenamento

Sistema de arquivos e banco de grafo.

Origem dos dados de treinamento

Processos da Justiça Federal e planilhas de empreendimentos para contratos e projetos.

Algoritmos e técnicas

- Para detecção de entidades nomeadas:
  - *Long short-term memory* (LSTM).
  - *Conditional Random Field (CRF).*
- Para *Part-of-speech tagging*:

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

- Modelo pré-treinado do *Spacy* para português em notícias (pt_core_news_sm).

Benefícios

- Ganho de eficiência: Redução do tempo de recuperação de informações relacionadas a pessoas.
- Pesquisa está disponível a qualquer auditor sem conhecimento profundo em Tecnologia da Informação.
- Descoberta de vínculos entre pessoas físicas e jurídicas obtidos a partir de diferentes peças e visualizados de maneira unificada e simples.

Contatos

sti@tcu.gov.br – STI - Secretaria de Soluções de Tecnologia da Informação

soma@tcu.gov.br – Soma - Secretaria de Orientação, Métodos, Informações e Inteligência para o Controle Externo e o Combate à Corrupção

6. TCU - Seleção de Atos de Pessoal para Análise por Critério de Risco por meio de Inteligência Artificial

13. A Secretaria de Fiscalização de Integridade de Atos e Pagamentos de Pessoal e de Benefícios Sociais (Sefip) é a unidade do TCU responsável pela análise dos atos de pessoal para fins de registro (admissões, aposentadorias, reformas e pensões civis e militares). Atualmente, os atos de pessoal sujeitos a registro são cadastrados pelos gestores de pessoal dos diversos órgãos da APF no sistema e-Pessoal, onde passam por uma série de críticas automáticas. Quando o sistema não encontra nenhum problema significativo, os atos passam por um rito sumário na Sefip e seguem em lote para julgamento pelo plenário, quase sempre pela legalidade. Quando o sistema identifica algum problema potencial, o ato deve necessariamente passar por análise de um auditor.

14. Atualmente, o número de atos do e-Pessoal com algum problema potencial que estão na situação de "Aguardando Análise do Auditor" é superior a 130 mil. Com o atual quantitativo de auditores da Sefip, seriam necessários mais de 10 anos para que todo este estoque de atos fosse tratado, isso sem considerar a chegada de novos atos. Assim sendo, é fundamental que exista um sistema de seleção de atos que priorize a análise dos atos com maior potencial de estarem gerando danos ao erário, de forma que sejam julgados pela ilegalidade o quanto antes. Para este fim, foi proposta a estratégia de priorização de atos com base em risco.

15. Esta estratégia baseada em risco é formada por dois eixos: benefício de controle e probabilidade de ilegalidade. O risco é dado pela combinação do benefício monetário esperado (ou materialidade) com a probabilidade de ilegalidade. O benefício estimado do ato é calculado automaticamente por meio de uma série de regras especificadas pelos especialistas da Sefip que consideram os benefícios identificados pela fiscalização contínua da folha de pagamentos, os valores pagos mensalmente ao servidor ou aposentado e a ficha financeira do ato. Já a probabilidade de ilegalidade é calculada por um modelo de inteligência artificial que considera o histórico de decisões passadas do Tribunal sobre os atos como base de aprendizagem para estimativa das probabilidades das decisões futuras.

16. O algoritmo de IA utilizado é o XGBoost e as variáveis utilizadas no modelo foram selecionadas por um rigoroso processo de análise exploratória de dados. Este modelo foi avaliado automaticamente e atingiu um AUC de 97%. Também foi realizado um teste de controle, comparando o resultado do modelo com o obtido por um especialista humano. Dados os atos aguardando análise, tanto o modelo quanto o especialista selecionaram 50 atos cada. Os 100 atos selecionados dessa forma foram submetidos a um auditor humano, que após análise detalhada concluiu pela ilegalidade ou legalidade de

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

cada um dos 100 atos. O teste consistiu em comparar a taxa de atos ilegais identificada em cada um dos grupos de 50 atos. A taxa de ilegalidade dos atos selecionados pelo especialista da Sefip foi de 90% enquanto a taxa de ilegalidade dos atos selecionados pelo modelo atingiu 95%.

17. Os resultados atingidos pelo modelo foram considerados bastante interessantes. As probabilidades preditas por meio de inteligência artificial foram utilizadas para compor um painel que também traz os valores de benefício monetário e permite a seleção de atos com base em risco para posterior encaminhamento para análise pelos auditores.

Outros Órgãos de Controle

1. CGU - Malha Fina de Convênios

18. A solução denominada **Malha Fina de Convênios** utiliza-se de inteligência artificial para analisar as prestações de contas de convênios para transferência de recursos da União para entes subnacionais ou entidades sem fins lucrativos. Dentro desse contexto, a CGU identificou um desequilíbrio entre o estoque de convênios celebrados e a capacidade dos órgãos concedentes de analisar com celeridade as respectivas prestações de contas, motivando o desenvolvimento de uma solução automatizada para apoiar esse processo.

19. Assim, o modelo preditivo que viria a denominar-se "Malha Fina de Convênios", fruto de trabalho da CGU em acordo de cooperação com o Ministério da Economia, prevê o resultado da análise, atribuindo uma nota que varia de 0 a 1: quanto mais próxima de zero, maior a recomendação para aprovação automática da prestação de contas, e quanto mais próxima de um, maior indicação de chance de rejeição, demandando, portanto, uma análise manual cuidadosa pelo auditor antes de decidir pela aprovação ou rejeição das contas. Cabe ao órgão usuário definir uma nota de corte abaixo da qual as contas serão automaticamente aprovadas, sendo esse valor diretamente proporcional ao apetite ao risco de aprovar contas potencialmente reprováveis, em troca de maior celeridade nas análises, uma vez que a análise das prestações de contas é um processo trabalhoso e que requer servidores públicos especializados.

20. A equipe responsável pelo sistema tem realizado análises de algoritmos e do tamanho ideal da base de treinamento para que o sistema forneça melhor acurácia preditiva. Exemplos de resultados obtidos indicaram que o algoritmo *random forest* fornece os melhores resultados, e que quanto maior o tamanho da base de treinamento, melhor será o desempenho obtido pelo modelo, tendo esse tamanho, entretanto, um ponto de saturação, a partir do qual o ganho em desempenho não é significativo.

21. Mais detalhes sobre o Malha Fina de Convênios podem ser encontrados em (Carvalho & Fajardo, 2020).

2. Tribunal de Contas do Estado da Paraíba (TCE-PB)

22. O Tribunal de Contas do Estado da Paraíba, em entrevista, reportou como principais usos de IA em seus desenvolvimentos um chatbot para interação com o público e um projeto em parceria com a Universidade Federal da Paraíba e a Universidade Federal de Campina Grande. Este último tem como foco o problema de fiscalização do percentual constitucional de gastos em saúde e educação.

23. Atualmente essa fiscalização é feita por amostragem pela auditoria do histórico de empenhos do governo estadual. No projeto que está em desenvolvimento, geraram um banco de dados para treinamento a partir das decisões dos auditores, com emprego de rede neural treinada pelas universidades, obtendo alto grau de acurácia nos resultados fornecidos pelo modelo (97% segundo informado pela equipe). O sistema ainda não está em produção, porém se encontra em estágio avançado de desenvolvimento.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

24. Paralelamente, o TCE-PB cogita implantar uma estrutura permanente para realização de *hackatons*, embora admitam que, apesar de algumas soluções terem surgido a partir de *hackatons*, atualmente o foco está na parceria com universidades, inclusive através de programas de mestrado profissional. Isso porque laboratórios de universidades acabam funcionando como *hackatons* por meio dos seus alunos, segundo visão fornecida pela equipe entrevistada.

25. Por outro lado, foi reportado que, uma vez concluído um projeto por uma universidade, em geral fazendo uso de tecnologias estado-da-arte, a manutenção do que foi desenvolvido tende a ser difícil. Foi informado que, atualmente, o estado da Paraíba possui forte capital intelectual em IA e tecnologias disruptivas em geral, com forte tendência em iniciativas empreendedoras com foco em gestão pública, tanto na capital como no interior do estado. Daí existir também o interesse em estimular o desenvolvimento, por meio de *hackatons*, de soluções úteis para a gestão pública a serem desenvolvidas pela iniciativa privada, não apenas com foco em fiscalização.

26. Por fim, a equipe reportou iniciativas, ainda em fases iniciais, para prospecção do uso de inteligência artificial na análise de notas fiscais, na análise de portais de transparência, e no processamento de imagens com emprego de geotecnologias a partir de fotos tiradas por drones de imóveis cujos valores de Imposto Predial e Territorial Urbano (IPTU) possam estar desatualizados.

3. Tribunal do Contas do Estado do Rio de Janeiro (TCE-RJ)

27. O Tribunal de Contas do Estado do Rio de Janeiro, em entrevista pela equipe deste levantamento, apresentou uma visão geral das iniciativas em andamento nas áreas de inteligência artificial e ciência de dados, cujo resumo será descrito nesta seção.

28. Em 2012, o TCE-RJ iniciou um movimento de inovação tecnológica, quando a área de Tecnologia da Informação começou a cogitar o uso de IA para apoiar as auditorias. Havia, por exemplo, a necessidade de identificar indícios de irregularidades em editais de licitações. Foi criado então um núcleo para a área de *business intelligence* (BI). Os trabalhos realizados pelo núcleo têm incluído análise exploratória de dados, portal de BI etc. O portal de BI, por exemplo, gera insumos para os trabalhos de auditoria. O núcleo de BI atualmente possui 5 servidores, incluindo titulares de mestrado e pós-graduação nos temas relacionados a análise de dados.

29. Outras iniciativas do órgão em ciência de dados desde então incluem:

- Projeto IRIS: repositório de bibliotecas para análises diversas (ex.: componente responsável por inferir se um determinado endereço se refere a uma empresa de fachada, por meio de técnicas de visão computacional).
- A universidade corporativa do órgão possui uma trilha de formação em análise de dados.
- Implantação de uma ferramenta corporativa, baseada no JupyterHub[1], com recursos para uso de BI pelo usuário final e integração com ambiente de aprendizagem de máquina e programação, disponibilizando um ambiente baseada em navegador Web no qual o usuário possa desenvolver seus próprios códigos.
- Soluções diversas para classificação textual.
- Comunidade de prática informal, na qual as pessoas demonstram interesse e conversam sobre o assunto.
- Extração de textos por meio de OCR – este foi apontado como um dos poucos usos efetivos de inteligência artificial. Trata-se de uma iniciativa em andamento, com uso de bibliotecas e

[1] JupyterHub: https://jupyter.org/hub

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

frameworks de mercado diversos. O objetivo é ter ao final um serviço que processe uma fila de documentos a serem lidos.

30. Com relação ao uso de ferramentas técnicas de programação pelos auditores do órgão, a equipe está confiante no patrocínio disponível no planejamento estratégico para 2020-2021, que inclui indicadores para criação de novos projetos na área de ciência de dados. Além disso, uma das disciplinas da trilha de capacitação em análise de dados trata especificamente da Lei Geral de Proteção de Dados – LGPD.

31. Sobre a governança de projetos descentralizados produzidos pela área de controle externo com uso de análise de dados, foi determinado que novos desenvolvimentos precisam ser feitos com foco no ambiente de produção, uma vez que ainda que o desenvolvimento tenha seu início de forma descentralizada, eventualmente será assumido pela área de TI.

4. TCE-RS

32. O Tribunal de Contas do Estado do Rio Grande do Sul já vem há algum tempo focando no uso de Tecnologia da Informação em geral como ferramenta de apoio ao controle externo. Exemplos incluem o uso de ferramentas automatizadas para trazer aos relatórios de fiscalização informações que o TCE-RS recebe dos seus jurisdicionados através dos seus sistemas, bem como outras fontes de informações.

33. Atualmente, o órgão possui em andamento projetos de desenvolvimento de ferramentas de apoio a auditorias, juntamente com uma empresa terceirizada especialista em ciência de dados. Uma das ferramentas – LAIS – funciona como um banco de preços para fins de comparação de preços de produtos adquiridos por jurisdicionados do TCE-RS, com recurso de geração de alertas quanto a indícios de sobrepreço. Para tal, a equipe utilizou como fonte de dados os itens das licitações cadastradas em um sistema do próprio TCE-RS, que armazena as informações estruturadas de contratações do estado e dos municípios, fornecidas pelos próprios jurisdicionados. Tais informações são remetidas de forma contínua, com prazo de envio de 5 (cinco) dias a cada evento envolvido na elaboração de um edital. O sistema também tem acesso a dados não estruturados como documentos e planilhas.

34. Em relação ao uso de notas fiscais, a equipe reportou como vantagem o fato de que o cadastro de licitações reduz as diferenças na descrição dos itens, que tendem a ser mais genéricas. As descrições por sua vez são pré-processadas e vetorizadas com o algoritmo Word2Vec (Mikolov, Chen, Corrado, & Dean, 2013). A partir daí, tais descrições são comparadas entre si por meio de similaridade de cossenos[1], sendo adicionalmente levadas em consideração as unidades de medida para o cálculo de similaridade.

35. Diariamente, os auditores fazem uso do sistema para detecção de indícios de sobrepreço, ao passo que o sistema utiliza também outras heurísticas baseadas em dados estruturados para gerar notas de risco para processos licitatórios. Segundo a equipe, o feedback fornecido pelos usuários tem sido positivo em termos de acurácia e rapidez dos resultados, tendo crescido no sistema de benefícios do controle do órgão as referências à LAIS, sendo atualmente o sistema com melhor avaliação pelos auditores.

36. Outras soluções desenvolvidas pelo órgão incluem sistemas para melhoria da qualidade das informações recebidas, identificação de anomalias em despesas e auditoria de atos de pessoal, além de APIs para agrupamento e cálculo de similaridade entre documentos.

[1] Cosine similalirity: https://en.wikipedia.org/wiki/Cosine_similarity

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

37. Dentre as dificuldades reportadas pela equipe, estão a ausência de padronização na forma como as denúncias devem chegar ao órgão, com forte dependência em relação à pessoa que recebe a denúncia e o encaminhamento que ela costuma dar.

38. Foi reportado ainda que a terceirização no desenvolvimento de soluções com uso de ciência de dados é necessária tendo em vista que os sistemas transacionais desenvolvidos por equipe própria requerem manutenção constante. Porém, segundo a equipe, existem planos de que o pessoal próprio venha a se envolver no desenvolvimento de soluções com emprego de IA. O modelo de contratação adotado, segundo os entrevistados, segue um esquema de mentoria, que no caso concreto se caracteriza por: documentação de todo o histórico de desenvolvimento, incluindo as vias sem saída; desenvolvimento em conjunto com a equipe do TCE-RS, incluindo detalhes sobre quais algoritmos utilizar; gravação das apresentações dos desenvolvimentos implementados; manutenção executada pelo próprio TCE-RS; processo ágil de desenvolvimento e teste da documentação disponibilizada.

39. A métrica utilizada no processo de contratação é o número de horas gastas em cada ordem de serviço. As entregas são sempre APIs, disponibilizadas por meio de conteinerização. Foi reportado adicionalmente que o contrato custou menos do que o estimado, pois foram utilizadas menos horas do que o previsto. Com relação a rotatividade de pessoal, foi dito que embora ocorra um certo grau de rotatividade que por vezes gera perda de históricos de resolução de problemas, o desempenho do serviço entregue não foi comprometido. Uma cláusula contratual impõe prazo de transição entre a saída de um desenvolvedor e a entrada de um novo em seu lugar. Em linhas gerais, a equipe reportou que a seleção da empresa desenvolvedora se mostrou bem-sucedida.

40. O TCE-RS tem um projeto estratégico interno para estimular a formação nas áreas de ciência de dados e inteligência artificial. Em relação à adesão por parte de auditores que não são da área de TI ao uso de ferramentas baseadas em IA, foi reportado que o perfil dos interessados em geral inclui aqueles que já têm conhecimentos em estatística ou economia, indicando que, embora as pessoas gostem da ideia do uso de robôs, dificilmente assumem o custo inicial de aprender a programar essas aplicações, possivelmente por falta de tempo ou de perfil adequado.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Referências

Stanford Vision Lab, Stanford University, Princeton University. (2021). *ImageNet*. Retrieved 07 15, 2021, from ImageNet: https://image-net.org/

Angelov, D. (2020). *Top2Vec: Distributed Representations of Topics.* Retrieved 07 28, 2021, from Arxiv: https://github.com/ddangelov/Top2Vec

Antonitsch, A., Figueira, A., Amaral, D., Fonseca, E., Vieira, R., & Collovini, S. (2016). Summ-it++: an Enriched Version of the Summ-it Corpus. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)* (pp. 2047-2051). Paris: European Language Resources Association (ELRA).

Attree, S. (2019). Gendered Ambiguous Pronouns Shared Task: Boosting Model Confidence by Evidence Pooling. *Proceedings of the First Workshop on Gender Bias in Natural Language Processing* (pp. 134-146). Florença: Association for Computational Linguistics. doi:10.18653/v1/W19-3820

Baldini Soares, L., FitzGerald, N., Ling, J., & Kwiatkowski, T. (2019). Matching the Blanks: Distributional Similarity for Relation Learning. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 2895-2905). Florença: Association for Computational Linguistics. doi:10.18653/v1/P19-1279

Bianchi, F., Terragni, S., & Hovy, D. (2021). Pre-training is a Hot Topic: Contextualized Document Embeddings Improve Topic Coherence. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)* (pp. 759-766). Online: Association for Computational Linguistics. doi:10.18653/v1/2021.acl-short.96

Bianchi, F., Terragni, S., Hovy, D., Nozza, D., & Fersini, E. (2021). Cross-lingual Contextualized Topic Models with Zero-shot Learning. *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume* (pp. 1676-1683). Online: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2021.eacl-main.143

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., . . . Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems* (pp. 1877--1901). H. Larochelle and M. Ranzato and R. Hadsell and M. F. Balcan and H. Lin.

Brum, H., & das Graças Volpe Nunes, M. (2018). Building a Sentiment Corpus of Tweets in Brazilian Portuguese. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).* Miyazaki, Japão: {European Language Resources Association (ELRA).

Bui, T.-C., Le, V.-D., To, H., & Cha, S. (2021). Generative Pre-training for Paraphrase Generation by Representing and Predicting Spans in Exemplars. *2021 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 83-90.

C4AI. (2020). *NLP2 - Recursos para Levar o Processamento de Linguagem Natural em Português para o Estado-da-Arte*. Retrieved 07 28, 2021, from C4AI: http://c4ai.inova.usp.br/pt/nlp2-pt/

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Câmara dos Deputados. (2021). *Estratégia de Transformação Digital da Câmara dos Deputados 2021-2024.* Retrieved 08 20, 2021, from Câmara dos Deputados: https://www2.camara.leg.br/a-camara/estruturaadm/gestao-na-camara-dos-deputados/gestao-estrategica-na-camara-dos-deputados/gestao-estrategica-de-tic/Estrategia_TDCD_Oficialcompactado4.pdf

CAPES. (2016). *Plataforma Sucupira*. Retrieved 07 13, 2021, from https://sucupira.capes.gov.br/sucupira/public/index.xhtml

Carvalho, S. T., & Fajardo, B. d. (2020). Uso de Inteligência Artificial na Auditoria das Prestações de Contas das Transferências. *XLIV ENCONTRO DA ANPAD - EnANPAD 2020.* on-line: ANPAD.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*.

Dou, Z.-Y., Liu, P., Hayashi, H., Jiang, Z., & Neubig, G. (2021). GSum: A General Framework for Guided Neural Abstractive Summarization. *Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).*

Fonseca, E., Sesti, V., Collovini, S., Vieira, R., Leal, A. L., & Quaresma, P. (1998). Collective Elaboration of a Coreference Annotated Corpus for Portuguese Texts. *Proceedings of II workshop on Evaluation of Human Language Technologies for Iberian Languages*, (pp. 68--82). Murcia, Espanha.

Freitas, C., Santos, D., Mota, C., Gonçalo Oliveira, H., & Carvalho, P. (2009). Relation detection between named entities: report of a shared task. *Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions ({SEW}-2009)* (pp. 129-137). Boulder, Colorado: Association for Computational Linguistics.

Freitas, L. A., & Vieira, R. (2015). Exploring Resources for Sentiment Analysis in Portuguese Language. *Brazilian Conference on Intelligent Systems, BRACIS*, (pp. 152-156).

Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). Synthetic data augmentation using GAN for improved liver lesion classification. *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018*, (pp. 289-293). doi:10.1109/ISBI.2018.8363576

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., . . . Bengio, Y. (2014). Generative Adversarial Networks. *Advances in Neural Information Processing Systems, 3*. doi:10.1145/3422622

Grootendorst, M. (2020). *BERTopic: Leveraging BERT and c-TF-IDF to create easily interpretable topics.* doi:10.5281/zenodo.4381785

Gulati, A., Qin, J., Chiu, C.-C., Parmar, N., Zhang, Y., Yu, J., . . . Pang, R. (2020). *Conformer: Convolution-augmented Transformer for Speech Recognition.* Retrieved from Arxiv: https://arxiv.org/abs/2005.08100

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (pp. 770-778).

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Heinzerling, B., & Strube, M. (2019). Sequence Tagging with Contextual and Non-Contextual Subword Representations: A Multilingual Evaluation. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 273-291). Forença: Association for Computational Linguistics.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-term Memory. *Neural computation, 9*, pp. 1735-80. doi:10.1162/neco.1997.9.8.1735

Howard, J., & Ruder, S. (2018). Universal Language Model Fine-tuning for Text Classification. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 328-339). Melbourne: Association for Computational Linguistics.

Hsu, W.-N., Bolte, B., Tsai, Y.-H. H., Lakhotia, K., Salakhutdinov, R., & Mohamed, A. (2021). *HuBERT: Self-Supervised Speech Representation Learning by Masked Prediction of Hidden Units.* Retrieved from Arxiv: https://arxiv.org/abs/2106.07447

Hu, R., Singh, A., Darrell, T., & Rohrbach, M. (2020). Iterative Answer Prediction With Pointer-Augmented Multimodal Transformers for TextVQ. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR*, pp. 9989-9999.

ICO; The Alan Turing Institue. (2020). *Explaining decisions made with AI*. Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-ai/

Jurafsky, D., & Martin, H. J. (2020, 12 30). *Speech and Language Processing (3rd ed. draft)*. Retrieved from Speech and Language Processing: https://web.stanford.edu/~jurafsky/slp3/

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and Improving the Image Quality of StyleGAN. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (pp. 8107-8116).

Kim, B., Wattenberg, M., Gilmer, J., Cai, C., Wexler, J., Viegas, F., & Sayres, R. (2018). Interpretability Beyond Feature Attribution: Quantitative Testing with Concept Activation Vectors ({TCAV}). *Proceedings of the 35th International Conference on Machine Learning* (pp. 2668-2677). Dy, Jennifer and Krause, Andreas.

Kirstain, Y., Ram, O., & Levy, O. (2021). Coreference Resolution without Span Representations. Retrieved from https://arxiv.org/abs/2101.00434

Lai, T., Hung Tran, Q., Bui, T., & Kihara, D. (2020). A Simple But Effective Bert Model for Dialog State Tracking on Resource-Limited Systems. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (pp. 8034-8038). doi:10.1109/ICASSP40776.2020.9053975

Li, R., Lin, C., Collinson, M., Li, X., & Chen, G. (2019). A Dual-Attention Hierarchical Recurrent Neural Network for Dialogue Act Classification. *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)* (pp. 383-392). Hong Kong: Association for Computational Linguistic. doi:10.18653/v1/K19-1036

Liu, L., Liu, X., Gao, J., Chen, W., & Han, J. (2020). Understanding the Difficulty of Training Transformers. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020.*

Liu, Q., Chen, Y., Chen, B., Lou, J.-G., Chen, Z., Zhou, B., & Zhang, D. (2020). You Impress Me: Dialogue Generation via Mutual Persona Perception. *Proceedings of the 58th*

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Annual Meeting of the Association for Computational Linguistics (pp. 1417-1427). Association for Computational Linguistics.

Liu, X., Duh, K., Liu, L., & Gao, J. (2020). *Very Deep Transformers for Neural Machine Translation.* Retrieved from Arxiv: https://arxiv.org/abs/2008.07772

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 4768–477). Long Beach: Curran Associates Inc.

Manning, C. D., Raghavan, P., & Schütze, H. (2009). *Introduction to Information Retrieval.* Cambridge University Press. Retrieved from https://nlp.stanford.edu/IR-book/

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, pp. 115–133.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781.

Molnar, C. (2021). *Interpretable Machine Learning.* Livro licenciado sob a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence.

Mota, C., & Santos, D. (2008). *Desafios na avaliação conjunta do reconhecimento de entidades mencionadas: O Segundo HAREM.* Retrieved 07 28, 2021, from Linguateca: http://www.linguateca.pt/LivroSegundoHAREM/

Nayak, T., & Ng, H. (2020). Effective Modeling of Encoder-Decoder Architecture for Joint Entity and Relation Extraction. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*, pp. 8528-8535. doi:10.1609/aaai.v34i05.6374

Nori, H., Jenkins, S., Koch, P., & Caruana, R. (2019). *InterpretML: A Unified Framework for Machine Learning Interpretability.* Retrieved from arXiv:1909.09223: https://arxiv.org/abs/1909.09223

Paranjape, A., See, A., Kenealy, K., Li, H., Hardy, A., Qi, P., . . . Manning, C. D. (2020). Neural Generation Meets Real People: Towards Emotionally Engaging Mixed-Initiative Conversations. *3rd Proceedings of Alexa Prize.*

Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep Contextualized Word Representations. *Proceedings of the 2018 Conference of the North {A}merican Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)* (pp. 2227-2237). New Orleans: Association for Computational Linguistics.

Portuguese SQuAD v1.1 Dataset. (2019). Retrieved 07 23, 2021, from Metatext: https://metatext.io/datasets/portuguese-squad-v1.1

Radford, A., & Narasimhan, K. (2018). *Improving Language Understanding by Generative Pre-Training.*

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). *Language Models are Unsupervised Multitask Learners.* Retrieved from OPENAI: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf

Redmon, J., & Farhadi, A. (2018). *YOLOv3: An Incremental Improvement.* Retrieved from arXiv:1804.02767: https://arxiv.org/abs/1804.02767

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Reimers, N. (2021). *Clustering*. Retrieved 07 26, 2021, from SentenceTransformers: https://www.sbert.net/examples/applications/clustering/README.html

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). New York: Association for Computing Machinery.

Rodrigues, R. (2021, 05 17). *GPT-3 é a Inteligência Artificial mais avançada já criada pelo homem.* Retrieved 07 19, 2021, from Showmetech Corporate: https://corporate.showmetech.com.br/gpt-3-inteligencia-artificial-mais-avancada/

Ruder, S. (2019). *Text classification*. Retrieved 07 19, 2021, from NLP-progress: https://nlpprogress.com/english/text_classification.html

Ruder, S. (2021). *NLP-progress*. Retrieved 07 12, 2021, from https://nlpprogress.com/

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning Representations by Back-propagating Errors. *Nature*, 533-536.

Sanabria, R., Caglayan, O., Palaskar, S., Elliott, D., Barrault, L., Specia, L., & Metze, F. (2018). How2: A large-scale dataset for multimodal language understanding. *Proceedings Visually Grounded Interaction and Language (ViGIL).* Montreal: Neural Information Processing Society (NeurIPS).

Shahzad, M., Amin, A., Esteves, D., & Ngonga Ngomo, A.-C. (2021). InferNER: an attentive model leveraging the sentence-level information for Named Entity Recognition in Microblogs. *The International FLAIRS Conference Proceedings*, *34.* doi:10.32473/flairs.v34i1.128538

Soares, F., Yamashita, G. H., & Anzanello, M. J. (2018). A Parallel Corpus of Theses and Dissertations Abstracts. *International Conference on Computational Processing of the Portuguese Language* (pp. 345--352). Springer.

Souza, F., Nogueira, R., & Lotufo, R. (n.d.). BERTimbau: pretrained BERT models for Brazilian Portuguese. *9th Brazilian Conference on Intelligent Systems, BRACIS.* Rio Grande do Sul.

Sutton, R. S., & Barto, G. A. (2018). *Reinforcement Learning: An Introduction* (2a. ed.). MIT Press. Retrieved from http://incompleteideas.net/book/the-book-2nd.html

Sweenor, D., Hillion, S., Rope, D., Kannabiran, D., Hill, T., & O'Connel, M. (2020). *ML Ops: Operationalizing Data Science.* Sebastopol, CA: O'Reilly.

Tiedemann, J. (2012). Parallel Data, Tools and Interfaces in OPUS. *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12).* Istanbul: European Language Resources Association (ELRA).

Treveil, M., Omont, N., Stenac, C., Lefevre, K., Phan, D., Zentici, J., . . . Heidmann, L. (2020). *Introducing MLOps.* O'Reilly.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems.* I. Guyon and U. V. Luxburg and S. Bengio and H. Wallach and R. Fergus and S. Vishwanathan and R. Garnett.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Vincent, J. (2016, 03 24). Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. Retrieved 08 13, 2021, from https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., & Bowman, S. (2018). GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding. *Proceedings of the 2018 {EMNLP} Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP* (pp. 353-355). Bruxelas: Association for Computational Linguistics. doi:10.18653/v1/W18-5446

Wang, X., Jiang, Y., Bach, N., Wang, T., Huang, Z., Huang, F., & Tu, K. (2021). Automated Concatenation of Embeddings for Structured Prediction. *The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021).* Association for Computational Linguistic.

Wexler, J., Pushkarna, M., Bolukbasi, T., Wattenberg, M., Viegas, F., & Wilson, J. (2019). The What-If Tool: Interactive Probing of Machine Learning Models. *IEEE Transactions on Visualization and Computer Graphics*.

Yamada, I., Asai, A., Shindo, H., Takeda, H., & Matsumoto, Y. (2020). LUKE: Deep Contextualized Entity Representations with Entity-aware Self-attention. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 6442-6454). Online: Association for Computational Linguistics.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding. In H. Wallach, H. Larochelle, A. Beygelzimer, F. Alche-Buc, E. Fox, & R. Garnett (Ed.), *Advances in Neural Information Processing Systems. 32.* Curran Associates, Inc.

Zhong, M., Liu, P., Chen, Y., Wang, D., Qiu, X., & Huang, X. (2020). Extractive Summarization as Text Matching. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 6197--620). Online: Association for Computational Linguistics.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 69302758.

Para verificar as assinaturas, acesse www.tcu.gov.br/autenticidade, informando o código 71234330.

PLATAFORMA DE SERVIÇOS DIGITAIS CONECTA-TCU

TERMO DE CIÊNCIA DE COMUNICAÇÃO

(Documento gerado automaticamente pela Plataforma Conecta-TCU)

Comunicação: Ofício 027.378/2022-SEPROC

Assunto: NOTIFICACAO

Processo: 006.662/2021-8

Órgão/entidade: Tribunal Regional Do Trabalho Da 7ª Região

Destinatário: TRIBUNAL REGIONAL DO TRABALHO DA 7ª REGIÃO/CE

Informo ter tomado ciência, nesta data, da comunicação acima indicada dirigida à/ao TRIBUNAL REGIONAL DO TRABALHO DA 7ª REGIÃO/CE pelo Tribunal de Contas da União, por meio da plataforma Conecta-TCU.

Data da ciência: 17/06/2022

(Assinado eletronicamente)

NEIARA SAO THIAGO CYSNE FROTA

Usuário habilitado a receber e a acessar comunicações pela plataforma Conecta-TCU.