

EDITORIAL • Ciência de Dados na Administração Pública: Desafios e Oportunidades

<https://doi.org/10.36428/revistadacgu.v14i26.617>

A Ciência de Dados é uma área de pesquisa interdisciplinar para o estudo e análise de dados que está recebendo cada vez mais atenção por popularizar o uso de métodos científicos para extração de conhecimento para tomada de decisão. Matemática, Estatística, Ciência da Computação, Engenharia de Software, Aprendizado de Máquina, Análise de Dados, Mineração de Dados e Textos, Gestão de Projetos e *Big Data* são exemplos de áreas exploradas por um cientista de dados.

A administração pública é um campo que não só detém inúmeras bases de dados, mas que precisa cada vez mais desenvolver competências para transformá-los em evidências que fundamentem a melhor tomada de decisão. A título de exemplificação, só em relação aos dados abertos acessíveis por meio do Portal [Dados.gov.br](https://dados.gov.br), são disponibilizados 13.726 conjuntos de dados que abrangem 218 organizações públicas classificados em 20 grandes temas, que vão da defesa e segurança à indústria, passando pelo meio-ambiente, educação e saúde.

Para fins de contextualização dos desafios e oportunidades da Ciência de Dados na administração pública, a Revista da CGU emoldura a discussão a partir de dois grandes panoramas: a transição digital e o desenvolvimento social, propondo uma reflexão a partir do contraste do tema em pauta com a fotografia que ilustra a capa desta edição.

A foto vem do Recife e é de autoria de Wesley Marcelo D'Almeida dos Santos, sendo intitulada “Reflexo da desigualdade”. Esta foto obteve a segunda colocação no III Concurso Nacional de Fotografia da Ouvidoria-Geral da União.

A cidade do Recife notabiliza-se por ser a sede do Porto Digital, um dos principais parques tecnológicos e ambiente de inovação do Brasil, resultado de uma ação coordenada entre governo, academia e empresa para requalificação urbana de uma região antes degradada do seu centro histórico da cidade transformando-a em um polo estratégico em soluções tecnológicas aliando ainda a inclusão digital da comunidade em seu entorno.

Nesse particular, o conhecimento dos dados, principal insumo dessa transição digital que avança em passos largos para uma nova economia, intermedeia uma nova relação entre os cidadãos e o governo no provimento de políticas públicas.

O governo brasileiro empreendeu significativos esforços se preparando para essa realidade e recentemente foi reconhecido como segundo líder em governo digital no mundo pelo relatório *GovTech Maturity Index 2022* do Banco Mundial.

Posto isto, esta edição convida seus leitores à reflexão sobre em que medida esse esforço da agenda pública em direção ao governo digital pode contribuir e está efetivamente contribuindo para uma melhor prestação de serviços públicos e superação das desigualdades sociais.

Para trazer uma visão de um órgão de controle com destacada experiência avaliativa, o presente Dossiê Especial “Ciência de Dados na Administração Pública: Desafios e Oportunidades” é aberto com uma entrevista de Taka Ariga, cientista-chefe de Ciência de Dados e Diretor do Laboratório de Inovação da Agência Governamental de *Accountability* do Governo dos Estados Unidos – GAO. Taka Ariga chamou atenção para questões que precisam

ser observadas sobre decisões de políticas públicas baseadas em inteligência artificial, a importância do relacionamento com a academia e aplicações práticas da inteligência em políticas públicas, incluindo a perspectiva da previsão e mitigação de corrupção.

Em sequência à entrevista, o presente dossiê apresenta nove trabalhos inéditos, meticulosamente selecionados pela mais concorrida chamada de trabalhos da história da Revista da CGU. São cinco artigos científicos e quatro relatos técnicos, os quais são apresentados a seguir.

O artigo "*Classificação de processos judiciais segundo Objetivos de Desenvolvimento Sustentável da Agenda ONU 2030*", de Luiz Fux, Pedro Felipe Santos, Aline Braga, Pamella Edokawa e Júlio Luz de Castro, apresenta a ferramenta RAFA 2030, que viabiliza a integração entre a Agenda 2030 da ONU e o STF via aprendizagem de máquina e análise de dados fazendo a classificação de ações por ODS da Agenda 2030.

O artigo seguinte "*Entorno legal y adopción de blockchain como herramienta para prevenir la corrupción en contrataciones públicas: reflexiones sobre iniciativas europeas y los marcos normativos brasileños*", de Samir Santos e Lucio Alves Angelo Junior, apresenta os principais desafios e oportunidades para a adoção, pelo governo brasileiro, de blockchain como ferramenta para prevenção a corrupção em contratações públicas.

Já o artigo "*Como somos vistos? Análise da imagem organizacional pública utilizando ciência de dados*", de Carolina Silveira, Carla Marcolin e Carlos Henrique Rodrigues, apresenta por meio de uma técnica inovadora uma análise da imagem organizacional da ANEEL veiculada na imprensa. O artigo apresenta resultados interessantes, inclusive uma disparidade recorrente entre o conteúdo da manchete e o conteúdo da reportagem em si.

O autor Rubens Quaresma Santos propõe no artigo "*Estimando a Arrecadação da Dívida Ativa da União com Machine Learning: uma análise baseada nos dados de arrecadação do período de 2015 a 2021*" a aplicação de algoritmos de aprendizado de máquina para projetar a arrecadação futura em substituição da técnica da suavização exponencial hoje empregada pela Procuradoria-Geral da Fazenda Nacional. Os modelos propostos mostraram-se ancorados em parâmetros mais amplos e robustos que o cálculos atuais, viabilizando uma decisão mais bem amparada.

Em relação ao artigo intitulado "*Acesso à informação no Brasil e ciência de dados: classificação hierárquica descendente em pedidos realizados à Prefeitura de São Paulo de 2012 a 2019*", os autores Cláudio Henrique Fontenelle Santos e Ana Lúcia Romão apresentam técnicas de análise inteligente de dados para organizar e compreender os pedidos de acesso à informação pelos cidadãos, permitindo um melhor entendimento das demandas da sociedade nesse assunto.

O Relato Técnico "*A Regulação baseada em dados: a era digital na agência reguladora de São Paulo*" de Itamar Aparecido de Oliveira apresenta a experiência da Diretoria Técnica de Saneamento Básico da Agência de Serviços Públicos do Estado de São Paulo (Arsesp) em seu trajeto de modernização das atividades regulatórias por meio de estratégias de análise de dados. Os resultados indicaram que a adoção de metodologias regulatórias baseadas em análise de dados contribuíram para melhores análises regulatórias e decisões mais refinadas, fortalecendo a segurança no ambiente regulado.

Por meio do Relato Técnico "*Proposta de framework de internacionalização no contexto da evolução da Plataforma Lattes*", os autores José Francisco Salm Junior, Paulo Henrique de Assis Santana e Geraldo Sorte apresentam a proposta de intervenção denominada de *framework* CV-PTBR que objetiva estabelecer um protocolo de trabalho conjunto baseado em diretrizes entre as equipes do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq e a Fundação para a Ciência e a Tecnologia – FCT de Portugal para assimilação das tecnologias e seus elementos para o uso nos projetos de aprimoramento das rotinas de importação de dados na Plataforma Lattes e a evolução das integrações de dados na Plataforma CIÊNCVIAVITAE de Portugal.

Em sequência, o Relato Técnico "*Ciência de Dados no Controle Social em Educação*", de Gilleddson Fryttys Menezes Leite e Hugo Leonardo Tosto Cuoco, discute o uso de ferramenta de *business intelligence* para análise de dados do Sistema de Informações sobre Orçamentos Públicos em Educação (SIOPE) e do Censo Escolar, visando

a análise inteligente das despesas em educação e facilitando a visualização de dados e comparação entre diferentes entes da federação.

O dossiê ciência de dados é encerrado com o Relato Técnico "*Alice: Desafios, resultados e perspectivas da ferramenta de auditoria contínua de compras públicas governamentais com uso de inteligência artificial*", de Andre Luiz Monteiro da Rocha, Matheus Scatolino de Rezende e Tiago Chaves Oliveira, o qual apresenta o histórico da construção e os principais resultados gerados pelo Analisador de Licitações, Contratos e Editais (Alice), o sistema de identificação de indícios de irregularidades em compras públicas.

Finalizando esta edição, apresentamos três artigos científicos recepcionados pelo fluxo contínuo de submissões à Revista da CGU.

O artigo "*A trajetória mexicana na consolidação do direito de acesso à informação pública: histórico e desafios*", de Luma Poletti Dutra, discute o pioneirismo do México na regulação do direito do acesso à informação na América Latina por meio de uma pesquisa qualitativa que envolveu entrevistas realizadas com pesquisadores e ativistas mexicanos.

Os autores Gabriela de Oliveira Silva, Renata Mendes de Araujo e Fernando Filgueiras no artigo "*Dados abertos: uma análise da Empresa Brasileira de Serviços Hospitalares e de sua rede de hospitais universitários federais*" analisaram o grau de alinhamento da Empresa Brasileira de Serviços Hospitalares (Ebserh) e de sua rede de hospitais universitários federais aos princípios da Política Nacional de Dados Abertos (PNDA).

No artigo "*Contribuições das denúncias e representações para o controle externo realizado pelo Tribunal de Contas do Estado de Santa Catarina*", os autores Fernanda Camila De Carli e Fabiano Maury Raupp discutiram a participação social a partir do nível de interação das espécies de controle incidentes pelo Tribunal de Contas de Santa Catarina sobre a gestão pública a partir da lente teórica do Novo Serviço Público e da *accountability* diagonal.

Por fim, agradecemos a atividade voluntária de revisão pelos pareceristas que contribuíram para esta edição por meio de avaliações construtivas para o aperfeiçoamento dos trabalhos. A todos eles, cujos nomes encontram-se divulgados ao final desta edição, nosso agradecimento.

Desejamos a todos uma boa leitura!



Daniel Matos Caldeira

Universidade de Lisboa, Portugal

matoscaldeira@gmail.com

<https://orcid.org/0000-0001-9823-8356>

Doutorando em Administração Pública pela Universidade de Lisboa. Auditor Federal de Finanças e Controle da Controladoria-Geral da União.



Dr. Alex Lopes Pereira (Enap)

<https://orcid.org/0000-0002-6405-429X>

Doutor em Engenharia Eletrônica e de Computação pelo Instituto Tecnológico de Aeronáutica – ITA. É integrante da carreira de Analista de Planejamento e Orçamento. Atualmente, trabalha no Ministério da Economia, professor de Ciência de Dados e empreendedor.

**Dr. Ricardo Marcacini (ICMC/USP)**

Universidade de São Paulo (USP), Brasil

<https://orcid.org/0000-0002-2309-3487>

Ricardo Marcacini tem doutorado em Ciência da Computação e Matemática Computacional pelo ICMC/USP, e atualmente é docente do Departamento de Ciências de Computação (SCC) no ICMC/USP. É orientador do programa de mestrado e doutorado de Ciência de Computação e Matemática Computacional do ICMC/USP, na qual também ministra disciplinas do tema de Aprendizado de Máquina e Inteligência Artificial. Atua no Laboratório de Inteligência Computacional (LABIC), com pesquisas nas áreas de Inteligência Artificial, Mineração de Textos, Mineração de Opiniões e Sensoriamento da Web.

**Dra. Solange Oliveira Rezende**

Universidade de São Paulo (USP), Brasil

<https://orcid.org/0000-0002-5233-7639>

Solange Rezende é professora no Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação (ICMC) da Universidade de São Paulo (USP), desde 1991. Tem experiência na área de Ciência da Computação, com ênfase em Inteligência Artificial, atuando principalmente nos temas relacionados com Mineração de Dados e Textos, Análise de Sentimentos e Sistemas de Recomendação. O foco das pesquisas realizadas é como extrair Conhecimento de Dados e Informação. Possui graduação em Licenciatura em Ciências Habilitação em Matemática pela Universidade Federal de Uberlândia (UFU) (1986), mestrado em Ciências de Computação e Matemática Computacional pela Universidade de São Paulo (1990) e doutorado em Engenharia Mecânica - São Carlos pela Universidade de São Paulo (1993). Realizou pós-doutorado na Universidade de Minnesota, USA (1995-1996). É orientadora no Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional; no Mestrado Profissional em Matemática, Estatística e Computação Aplicadas à Indústria e Vice Coordenadora do MBA em Inteligência Artificial e Big Data do ICMC/USP. Atua como pesquisadora no Laboratório de Inteligência Computacional (labic.icmc.usp.br) do Departamento de Ciências de Computação e no C4AI. Informações específicas sobre publicações e orientações podem ser acessadas em: <http://lattes.cnpq.br/8526960535874806>.