

**Jornalismo, transparência pública e inteligência artificial:
framework conceitual de plataforma para análise de dados públicos**

*Journalism, public transparency, and artificial intelligence:
a conceptual framework for a platform to analyze public data.*

Maria Thaís Firmino da SILVA¹

Resumo

Tendo em vista o papel social da mídia, esta pesquisa teve como objetivo apresentar, a partir de *framework* conceitual, uma plataforma de inteligência artificial (IA) projetada para a análise de dados públicos por jornalistas e cidadãos. O estudo parte da premissa de que existe lacuna entre a disponibilidade de dados e a apropriação efetiva da informação pelos usuários. Sendo assim, por meio de pesquisa aplicada e abordagem qualitativa, foram delineados aspectos teóricos, descrição de arquitetura conceitual, principais funcionalidades, dentre outras características de um sistema inteligente voltado à extração contínua e automática de dados de fontes oficiais – por meio de *web scraping* autorizado e integração com APIs governamentais. Como resultado, a iniciativa integra interfaces acessíveis, interação em linguagem natural, possibilidade de monitoramento da utilização de recursos públicos, bem como de identificação de padrões anômalos.

Palavras-chave: Jornalismo. Transparência pública. Acesso à informação. Inteligência Artificial. Fiscalização.

Abstract

Considering the social role of the media, this research aimed to present, from a conceptual framework, an artificial intelligence (AI) platform designed for the analysis of public data by journalists and citizens. The study starts from the premise that there is a gap between the availability of data and the effective appropriation of information by users. Therefore, through applied research and a qualitative approach, theoretical aspects, a description of the conceptual architecture, main functionalities, and other characteristics of an intelligent system focused on the continuous and automatic retrieval of data from official sources – through authorized web scraping and integration with government APIs – were outlined. As a result, the initiative integrates accessible interfaces, natural language interaction, the possibility of monitoring the use of public resources, as well as the identification of anomalous patterns.

Keywords: Journalism. Public transparency. Access to information. Artificial intelligence. Inspection.

¹ Mestre em Comunicação pela Universidade Federal do Ceará (UFC). E-mail: firminothais@gmail.com

Introdução

Tendo em vista a complexidade e o elevado quantitativo de dados governamentais disponibilizados em portais da transparência e sites de entes públicos, esta pesquisa se consolidou pelo desenvolvimento conceitual de plataforma baseada em inteligência artificial (IA) para auxiliar jornalistas e cidadãos no efetivo usufruto do direito previsto na Lei de Acesso à Informação (Brasil, 2011). Isso porque a crescente disponibilização em ambiente digital de informações como dados de contratos públicos, licitações e gastos governamentais, a automação de processos pode ser entendida como fundamental.

Nesse cenário, rapidez e precisão na verificação de grande volume de dados podem ser obtidas através da utilização de sistemas inteligentes baseados em IA, favoráveis à identificação de padrões, irregularidades e obtenção de informações para o desenvolvimento de conteúdos de interesse público. Sob essa perspectiva, é factível a redução da destinação de profissionais à execução de atividades mecânicas (Larson 2021; Bostrom, 2014) que demandam carga-horária elevada para serem concluídas, sendo indispensável, no entanto, o treinamento adequado do modelo computacional e a autonomia humana na análise dos dados.

A problemática central desta pesquisa, portanto, se apresenta da seguinte forma: quais características devem ser incorporadas no desenvolvimento de uma plataforma de IA voltada para a análise de dados abertos, com o objetivo de potencializar o jornalismo pautado em transparência pública e promover a participação cidadã? Nesse sentido, o objetivo geral preconizado se consolida por desenvolver *framework* conceitual de plataforma digital voltada à análise de dados abertos da administração pública por jornalistas e cidadãos. Já os objetivos específicos são: identificar características funcionais e conceituais para o desenvolvimento de plataforma digital baseada em IA que se apresente coerente à legislação vigente (Brasil, 2011; 2018); propor a arquitetura e os componentes essenciais para a plataforma através de *framework* conceitual, com foco em usabilidade, acessibilidade, funcionalidade e integração de fontes públicas; apontar para o potencial da plataforma apresentada como instrumento de favorecimento da atividade jornalística pautada em dados abertos e fomento à participação cidadã.

Sendo assim, a relevância desta pesquisa, se atrela ao reconhecimento de que, mesmo diante de avanços significativos na disponibilização de dados de transparência pública, barreiras técnicas, operacionais e elucidativas podem impor distanciamento no

acesso, apropriação e uso efetivo dessas informações. Assim sendo, a resposta à problemática suscitada poderá dirimir limitações concernentes à prática de utilização de dados abertos como insumos para a investigação jornalística e participação cidadã.

Metodologia

Posto isso, esta pesquisa, que possui natureza aplicada (Marconi; Lakatos, 2004), está fundamentada em abordagem qualitativa voltada à interpretação do fenômeno a partir da análise de documentos e contextos (Gil, 2008). Nesse sentido, foram utilizadas fontes documentais, precisamente a Lei de Acesso à Informação (Brasil, 2011) e a Lei Geral de Proteção de Dados Pessoais (Brasil, 2018); e bibliográficas, voltadas a publicações com temáticas concernentes à utilização da IA na atuação jornalística e em soluções para a sociedade. Para tanto, os procedimentos metodológicos foram distribuídos nas seguintes etapas: 1) Levantamento bibliográfico e documental; 2) Sistematização e análise de dados; e 3) Desenvolvimento de *framework* conceitual. Destarte, a estruturação foi delineada com foco na problemática inicial, sendo a triangulação de fontes e métodos permissiva à identificação de requisitos funcionais e conceituais relevantes, o que se mostrou fundamental para a formulação do *framework* conceitual do sistema inteligente baseado em IA.

Jornalismo, transparência pública e inteligência artificial

A transparência pública é a base da democracia, contudo sua efetividade não está somente na disponibilização de dados, mas na apropriação de informações sobre a gestão pública pelo cidadão, o garantidor dos recursos financeiros que são administrados. Isso significa que, diferente da mera publicação, mecanismos de transparência interativa, com participação cidadã, engajamento e explicações concomitantes à divulgação dos dados, aparecem como favoráveis para a outorga de confiança por parte de indivíduos (Cucciniello & Nasi, 2014). Ou seja, além de tecnicamente disponíveis, os conteúdos devem ser compreensíveis e relevantes, ainda que, reconhecidamente, a abertura de dados governamentais não se apresente como processo neutro, mas moldado por burocracias, disputas políticas e controle do conteúdo, inclusive quanto ao momento e formato de divulgação (Ruijter et al., 2020).

No contexto brasileiro, o acesso à informação está previsto desde a Constituição da República Federativa do Brasil, que estabeleceu, além da estrutura e princípios do Estado, os direitos e deveres dos cidadãos (Brasil, 1988). Contudo, o delineamento para o usufruto dessa premissa se configurou, apenas, após duas décadas do marco constitucional (Brasil, 2011), através da Lei de Acesso à Informação (LAI), impulsionada por compromissos internacionais como a *Open Government Partnership*, que visa a promoção da transparência e da integridade pública.

Diante disso, é importante pontuar que a LAI estabelece lógica reversa para a aplicação do princípio da publicidade (Brasil, 1988), que se torna regra e o sigilo exceção (Brasil, 2011). Sob esse entendimento, a solicitação de informações por pessoa física ou jurídica assume destaque ao mesmo tempo que exige o requerente da necessidade de justificativa para o pleito. Soma-se a isso o fato de que a Lei (Brasil, 2011) preconiza procedimentos objetivos e ágeis de atendimento da demanda, bem como a necessidade de disponibilização de conteúdos pautados na acessibilidade, inclusive quanto ao uso da linguagem.

Em vista disso, a LAI (Brasil, 2011) prevê que é dever dos órgãos e entidades públicas a utilização de todos os meios e instrumentos legítimos que estiverem à disposição para a viabilização de dados de forma ativa, ou seja, sem a necessidade de solicitação prévia, e passiva, mediante requisição. Assim, se tornou obrigatória a divulgação em sítios oficiais da rede mundial de computadores, exceto para municípios com população de até dez mil habitantes, contudo, mesmo nesses casos, os entes públicos permanecem submetidos ao dever de dispor em local de fácil acesso informações de interesse coletivo ou geral por eles produzidas ou custodiadas.

Esse cenário, delineado pela legislação, fomentou a criação de portais da transparência e páginas específicas para a prestação de contas por parte de entes públicos com, no mínimo, o registro de: competências; estrutura organizacional; endereço; telefones de unidades; horários de atendimento ao público; repasses ou transferências de recursos financeiros; despesas; informações concernentes a procedimentos licitatórios, o que inclui editais e resultados, bem como a todos os contratos celebrados; dados gerais para o acompanhamento de programas, ações, projetos e obras; além de respostas a perguntas mais frequentes da sociedade (Brasil, 2011).

No entanto, além de canais de comunicação para a transparência ativa, foi criado no Brasil o Serviço de Informação ao Cidadão (SIC) que, posteriormente, foi assimilado

pela plataforma Fala.BR, a fim de atender aos registros de manifestações de ouvidoria e solicitações de acesso à informação, ou seja, à transparência passiva. O sistema, viabilizado pelo Governo Federal para demandas do Poder Executivo, também passou a ser utilizado por órgãos dos Poderes Legislativo e Judiciário, mediante solicitação de adesão. Desse modo, a implementação da LAI exigiu da gestão pública mudanças institucionais significativas, que vão desde à criação e adaptação de sistemas eletrônicos até à capacitação de servidores públicos.

Após sete anos de promulgação da LAI, no entanto, a Lei Geral de Proteção de Dados Pessoais (LGPD) entrou em vigor e impactou significativamente a forma de disponibilização de dados abertos (Brasil, 2018), a partir da necessidade de prover equilíbrio entre o direito coletivo à informação e o direito individual à privacidade. Conseqüentemente, a obrigatoriedade da divulgação de dados de interesse público promovida pela LAI teve limites impostos pela LGPD no que tange ao tratamento de dados pessoais, inclusive no setor público.

Nesse cenário, o sigilo de informações pessoais sem comprometimento do dever do Estado de assegurar transparência impulsionou a adoção de medidas como a anonimização de dados, a análise do interesse público envolvido e a aplicação dos princípios da necessidade e proporcionalidade (Brasil, 2018). A articulação entre as leis, portanto, passou a requerer cultura institucional pautada na gestão ética e responsável da informação, a fim de associar a obrigação legal de transparência com o respeito aos direitos de privacidade e proteção de dados.

Ademais, mesmo diante das mudanças ocasionadas pela legislação, o acesso à informação ainda aparece como desafio, seja pela resistência à disponibilização de dados, falta de infraestrutura tecnológica adequada, gestão da informação deficitária, falta de padronização dos dados divulgados ou, até mesmo, por fatores políticos sobressalentes (Ruijter et al., 2020). Essa realidade aponta para o fato de que a disponibilização de dados é, apenas, uma parcela do que se perfaz como transparência pública, que deve ser pautada, inclusive, pela inteligibilidade e relevância para que, de fato, se configure como propícia ao exercício do direito à informação pelo cidadão.

Nesse contexto, o jornalismo, tendo em vista a sua função social, se consolida como elemento essencial para a promoção da transparência pública, especialmente devido à relação que estabelece com o cidadão. Contudo, em meio à complexidade e volume de dados dispostos no ambiente digital, investigações profundas tendem à dispendiosidade

se não forem associadas a ferramentas eficazes. Assim sendo, a pertinência da utilização de aprendizado de máquina (AM), visão computacional (VC) e processamento de linguagem natural (PLN) está na viabilização de escrutínio para a identificação de padrões ocultos e complexos na análise de dados abertos, sejam textuais ou imagéticos, além de detecção de anomalias e fraudes (Diakopoulos, 2019; Santos et al., 2022).

Dessa forma, a automação baseada em IA tem se mostrado promissora para que jornalistas analisem informações dispersas, documentos públicos, transações comerciais, relatórios financeiros, dentre outros conteúdos que, sem essa tecnologia, seria necessário tempo e recursos humanos sobressaltados ou, até mesmo, se tornaria inviável de execução (Diakopoulos, 2019). Por esse motivo, em diferentes países, inclusive no Brasil, Peru e Argentina, a utilização de algoritmos de IA tem assumido função de aliada nas redações, inclusive para a investigação de escândalos de corrupção (Dalben, 2022; Santos et al., 2022), o que aponta para a pertinência do uso de sistemas inteligentes na investigação jornalística, sem que a autonomia humana seja suplantada.

Essa ressalva é indispensável, porque, embora a IA se perfaça pelo alto poder computacional, a complexidade do raciocínio humano não é replicada em seu âmbito (Larson, 2021), o que evidencia a relevância da proeminência de indivíduos em processos com automação por sistemas inteligentes. Soma-se a isso o fato de que a compreensão e o controle crítico de ferramentas baseadas em IA se apresentam de forma opaca ao usuário e, por esse motivo, a capacidade de manutenção de autonomia, preliminarmente, se mostra como fragilizada (Russell, 2019).

Sob essa ótica, a IA deve ser entendida como fenômeno técnico que necessita de análise a partir de diferentes esferas, porém não deve ser desconsiderada a sua capacidade instrumental de eficiência, inovação e resolução de problemas complexos na sociedade contemporânea. Isso significa que a inovação deve ser equilibrada com a consideração de seus riscos, e que a análise crítica não deve estar apartada da relação homem-tecnologia, a fim de que a utilização da IA para favorecer a eficiência, velocidade e análise preditiva, não se consolide em visão unilateral ofuscada pela conveniência oferecida pelos sistemas.

Por esse motivo, o *framework* conceitual apresentado a seguir considera aspectos legítimos para a representação das principais características da plataforma de análise de dados públicos concernente a este escopo. Esta etapa, portanto, se mostra como fundamental para a comunicação da estrutura de forma prévia ao desenvolvimento do

projeto em sua totalidade, trata-se de base para refinamento de versão subsequente (Norman, 2013).

Framework conceitual da plataforma

Preliminarmente, é essencial mencionar que este *framework* conceitual, destinado ao desenvolvimento de plataforma digital voltada à análise de dados abertos por jornalistas e cidadãos, tem como objetivo primaz potencializar o acesso à informação (Brasil, 2011) ancorado à compreensão e à utilização de grandes volumes de dados de transparência pública. A premissa, portanto, é a existência de lacuna entre a disponibilização de dados e a efetiva apropriação de informações por parte de usuários. Nesse sentido, a proposta se utiliza de recursos de IA para automatizar processos de filtragem, categorização, visualização e interação, a fim de favorecer o uso dos dados (Bostrom, 2014).

Assim sendo, a arquitetura da plataforma é composta por quatro módulos, ancorados em usabilidade, acessibilidade, funcionalidade e integração de fontes públicas, que estão conectados entre si. São eles: 1) Coleta de dados públicos; 2) Processamento e análise de dados após a coleta; 3) Visualização de dados; e 4) Interface de usuário. Em primeiro lugar, a plataforma está pautada na extração contínua e automática de dados em diferentes fontes oficiais do Poder Executivo Federal — como APIs, portais de dados abertos e documentos públicos disponibilizados de forma *online* —, sendo a coleta viabilizada por *web scraping* autorizado e integração com APIs de instituições governamentais e outras fontes relevantes, conforme previsto pela LAI (Brasil, 2011) e considerado nos pressupostos da LGPD (Brasil, 2018).

Já o processamento e análise de dados se estrutura pela IA tradicional, adequada a aplicações que exigem explicações, transparência e controle lógico, em associação com a IA generativa, viabilizada por algoritmos de aprendizado supervisionado e análise semântica voltada à classificação e contextualização de informações. Assim, o modelo computacional se apresenta como fundamental para a classificação automatizada, identificação de padrões, mineração e extração de informações relevantes para a correlação entre dados e temas de interesse público (Bostrom, 2014). Para tanto, técnicas de PLN aparecem como fundamentais para a compreensão e acessibilidade de dados por jornalistas e cidadãos.

No que tange à visualização de dados, a utilização de *dashboards* interativos — com possibilidade de customização através da aplicação de filtros e ordenação de informações de forma intuitiva — com gráficos, mapas e tabelas exportáveis, se alinham à portabilidade dos resultados para o ambiente *offline*. Além disso, o recurso também atua como interface inicial que possibilita acesso ao mecanismo de interação conversacional baseado em linguagem natural.

Nesse ponto, é importante mencionar que esse elemento de interação — alicerçado na arquitetura *transformer* (Vaswani et al., 2017) aplicada ao formato de *chatbot* — trata-se de seção voltada ao esclarecimento e suporte, por meio de perguntas e conversas intuitivas, adaptadas ao nível de conhecimento de cada usuário, para a interpretação de padrões e resultados de dados públicos apresentados nos *dashboards* e evocados pelo usuário. A inclusão dessa funcionalidade encontra respaldo no fato de que 54% dos brasileiros utilizaram a IA generativa em 2024, enquanto a média global foi de 48% (Forbes Tech, 2025), o que torna a função um diferencial da plataforma para a promoção da participação cidadã ativa.

Por fim, a interface de usuário, projetada para navegação híbrida - com combinação de filtros manuais e comandos em linguagem natural -, tem como foco acessibilidade e usabilidade (Norman, 2013), e contempla: design responsivo para acesso em dispositivos móveis e desktop; viabilidade para leitores de tela, navegação por teclado e contraste visual; personalização de perfil cidadão (simplificado para interação com os dados) e jornalista (com funcionalidades avançadas de análise e customização de relatórios); além da interação conversacional, que contribui para a acessibilidade cognitiva de usuários com habilidades digitais restritas.

Sendo assim, a plataforma visa a redução da barreira de acesso à informação e o uso consciente e efetivo de dados públicos, através do fomento à transparência, suporte à atividade jornalística e aproximação dos cidadãos à cultura de dados. É importante mencionar, contudo, que desafios como padronização e atualização contínua das fontes de dados; curadoria constante de respostas geradas pela IA; e necessidade de compromisso irrevogável com a instrução digital dos usuários podem ser previstos.

Diante disso, o entendimento é de que o usufruto da conveniência propiciada pela tecnologia deve ser acompanhado de responsabilidade na esfera de desenvolvimento e de utilização. Isso porque a IA amplia horizontes da ação racional humana e, com isso, se estabelece como nova forma de articulação entre conhecimento técnico e prática social.

Considerações finais

Em suma, este estudo logrou êxito em seu objetivo, visto que foi viabilizado *framework* conceitual de plataforma digital voltada à análise de dados abertos da administração pública por jornalistas e cidadãos. Assim como foi possível identificar características funcionais e conceituais para o desenvolvimento de forma coerente à legislação vigente. Também foi apresentada arquitetura conceitual e componentes essenciais da plataforma, com foco em usabilidade, acessibilidade, funcionalidade e integração de fontes públicas. Foi apontado, ainda, o potencial da iniciativa para o favorecimento da atividade jornalística e fomento à participação cidadã, com alinhamento entre práticas automatizadas de fiscalização da gestão pública e manutenção da autonomia humana no processo interpretativo.

É possível afirmar, portanto, que o *framework* apresentado oferece contribuições relevantes na articulação entre jornalismo, transparência pública, inteligência artificial e acessibilidade a dados abertos governamentais. Trata-se, assim, de abordagem que não se limita à disponibilização de informações ou à viabilização de solução tecnológica isolada, mas se perfaz como plataforma delineada pela conexão entre dados de interesse público e usuários, a partir do entendimento de que a efetivação do direito à informação, para além do acesso, depende, também, da inteligibilidade, contextualização e possibilidade de apropriação das informações.

Nesse sentido, a arquitetura modular apresentada, que integra automação, análise semântica e interação como estratégias para a redução de barreiras técnicas, cognitivas e operacionais, se apresenta como uma das principais contribuições do estudo. Isso porque a incorporação de *dashboards* interativos associados a mecanismos conversacionais baseados em IA amplia o potencial de uso tanto por jornalistas, em atividades investigativas complexas, quanto por cidadãos com diferentes níveis de aptidão digital. Além disso, ao considerar desde a concepção do *framework* os limites e exigências concernentes à Lei de Acesso à Informação e à Lei Geral de Proteção de Dados Pessoais, a iniciativa se ancora em pressupostos éticos, legais e institucionais que reforçam sua viabilidade conceitual e relevância social.

Por fim, o resultado deste estudo poderá contribuir com a ampliação do acesso à informação; a automatização e aprofundamento de dados que fundamentam a análise e produção jornalística; o entendimento e interpretação de informações administrativas

complexas pelos cidadãos; o escrutínio da coerência de dados informados pela gestão pública; a inovação e criação de novos modelos de produção de conteúdo; dentre outras possibilidades. Além disso, pesquisas futuras poderão satisfazer delineamentos favoráveis à prototipação e realização de testes que apontem para a utilização da plataforma em contextos reais.

Referências

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. **Regula o acesso a informações previsto no inciso XXXIII do art. 5º da Constituição Federal**; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991. Diário Oficial da União: seção 1, Brasília, DF, p. 1, 18 nov. 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 2 jun. 2025.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965**, de 23 de abril de 2014 (Marco Civil da Internet). Diário Oficial da União: seção 1, Brasília, DF, p. 1, 15 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 2 jun. 2025.

BOSTROM, Nick. **Superintelligence: paths, dangers, strategies**. Oxford: Oxford University Press, 2014.

CUCCINIELLO, Maria; NASI, Greta. **Transparency for trust in government: How effective is formal transparency?** *International Journal of Public Administration*, v. 37, n. 13, p. 911-921, 2014. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01900692.2014.949754>. Acesso em: 2 jun. 2025.

DALBEN, Silvia. Uso de inteligência artificial nas redações jornalísticas na guerra contra a corrupção na América Latina. *In: CANAVILHAS, J.; RODRIGUES, C.; MORAIS, R.; GIACOMELLI, F. Mobilidade e Inteligência Artificial. Os Novos Caminhos do Jornalismo*. Covilhã: Livros LabCom, p. 421-444, 2022.

DIAKOPOULOS, Nicholas. **Automating the news: how algorithms are rewriting the media**. Harvard University Press, 2019.

FORBES TECH. **54 % dos brasileiros utilizaram ferramentas de IA em 2024**. *Forbes Brasil*, 14 jan. 2025. Disponível em: <https://forbes.com.br/forbes-tech/2025/01/54-dos-brasileiros-utilizaram-ferramentas-de-ia-em-2024/>. Acesso em: 2 jun. 2025.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. Editora Atlas SA, 2008.

MARCONI, M. de A.; LAKATOS, Eva Maria. **Metodologia científica**. São Paulo: Atlas, 2004.

LAROSE, Daniel T. **Data mining and predictive analytics**. John Wiley & Sons, 2015.

LARSON, Erik J. **The myth of artificial intelligence: why computers can't think the way we do**. Cambridge: Harvard University Press, 2021.

NORMAN, Donald A. **The design of everyday things**. MIT Press, 2013.

RUIJER, Erna et al. The politics of open government data: Understanding organizational responses to pressure for more transparency. **The American review of public administration**, v. 50, n. 3, p. 260-274, 2020. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/0275074019888065>. Acesso em: 10 jun. 2025.

RUSSELL, Stuart. **Human compatible: Artificial Intelligence and the Problem of Control**. New York: Viking, 2019.

SANTOS, M. F. et al. Inteligência Artificial na mídia: visões atuais e projeções futuras. In: CANAVILHAS, J.; RODRIGUES, C.; MORAIS, R.; GIACOMELLI, F. **Mobilidade e inteligência artificial: os novos caminhos do jornalismo**. Covilhã: Livros LabCom, p. 445-474, 2022.

VASWANI, Ashish et al. **Attention is all you need**. *Advances in neural information processing systems*, v. 30, 2017. Disponível em: <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>. Acesso em: 10 jun. 2025.