





MINERAÇÃO DE DADOS NO CONTEXTO DOS PREGÕES ELETRÔNICOS

DATA MINING IN THE CONTEXT OF THE ELECTRONIC AUCTIONS

 Hugo Medeiros Souto¹

 Eduardo Martins de Arruda²

 Wagner Junqueira de Araújo³

¹Mestrando Profissional em Gestão nas Organizações Aprendentes (PPGOA/UFPB)

E-mail: hugomsouto@gmail.com

²Mestre Profissional em Gestão nas Organizações Aprendentes (PPGOA/UFPB)


E-mail: eduardo.arruda73@gmail.com

³ Doutor em Ciência da Informação pela UNB, Professor do PPGCI e do PPGOA/UFPB

E-mail: wagnerjunqueira.araujo@gmail.com



ACESSO ABERTO

Copyright: Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional. 

Conflito de interesses: Os autores declaram que não há conflito de interesses.

Financiamento: Não há.

Declaração de Disponibilidade dos dados:

Todos os dados relevantes estão disponíveis neste artigo.

Recebido em: 20/09/2019.

Revisado em: 01/10/2019.

Aceito em: 10/10/2019.

Como citar este artigo:

SOUTO, H. M.; ARRUDA, E. M.; ARAÚJO, W. J. Mineração de dados no contexto dos pregões eletrônicos. **Informação em Pauta**, Fortaleza, v. 4, n. especial, p. 47-64, nov. 2019. DOI: <https://doi.org/10.32810/2525-3468.ip.v4iEspecial.2019.42603.47-64>.

RESUMO

O governo tem buscado acompanhar as evoluções e tendências das tecnologias da informação e comunicação no contexto do processo licitatório. Dessa forma, o pregão eletrônico surge como um dos produtos desses esforços, caracterizando-se como uma modalidade que apresentou soluções estruturais para reduzir o excesso de burocracia na aquisição de bens e serviços comuns. Sua implantação contribuiu para o aumento da participação de fornecedores e da competitividade nos certames, pois reduziu e até eliminou, em alguns casos, a necessidade da presença física dos fornecedores. A modalidade representa hoje mais de 94% das licitações ocorridas no país, somando mais de 84% do volume de recursos gastos por meio de licitações. Apesar de seus benefícios, têm sido identificadas falhas de diversas formas, sendo uma delas a possibilidade de utilização de bots, softwares programados para se conectarem nas plataformas de pregões eletrônicos e emitirem lances de forma imediata e automatizada, obtendo vantagens desproporcionais em relação a seres humanos. Um segundo produto resultado dos esforços governamentais em busca da modernização das compras é a ampliação da transparência através da disponibilização de dados abertos das licitações e contratos, incluindo dados dos pregões eletrônicos, como parte do contexto de Governo Aberto e transformação digital. O objetivo geral desse estudo é discernir como a análise de dados pode ser útil na prevenção de fraudes em processos de pregão eletrônico, fundamentada na premissa de dados abertos do Governo Federal. Por fim, conclui-se que esse estudo apoia a importância de se discutir os efeitos do uso de bots em pregões eletrônicos e acredita-se no potencial de

maior controle social por parte do cidadão por meio do acesso e análise dos dados abertos do Governo.

Palavras-chave: Mineração de dados. Governo Eletrônico. Pregão Eletrônico. Bots. Dados abertos.

ABSTRACT

The government has sought to follow developments and trends in information and communication technologies in the context of the bidding process. Thus, electronic trading is one of the products of these efforts, being characterized as a modality that presented structural solutions to reduce the excess of bureaucracy in the acquisition of common goods and services. Its implementation contributed to increase the participation of suppliers and competitiveness in the fairs, as it reduced and even eliminated, in some cases, the need for physical presence of suppliers. The modality represents today more than 94% of the tenders that took place in the country, totaling more than 84% of the amount of resources spent

through tenders. Despite their benefits, failures have been identified in a number of ways, one of which is the possibility of using bots, software programmed to connect to electronic trading platforms and bidding immediately and automatically, obtaining disproportionate advantages over humans. A second product of government efforts to modernize procurement is increased transparency through the availability of open bidding and contract data, including e-bidding, as part of the Open Government and digital transformation context. The overall purpose of this study is to discuss how data analysis can be useful in preventing fraud in electronic trading processes, based on the Federal Government's open data premise. Finally, it is concluded that this study supports the importance of discussing the effects of the use of bots in electronic trading sessions and believes in the potential for greater social control by citizens through access and analysis of government openings.

Keywords: Data Mining. Electronic Government. Electronic trading floor. Bots Transparent data.

1 INTRODUÇÃO

Entre as diferentes formas de interação do Governo com a sociedade, seja em formato digital ou analógico, os processos de compras governamentais estão entre os principais, pois mantêm o suprimento de materiais e serviços necessários para o funcionamento adequado das instituições da administração pública. Diante dessas características, o interesse na ampliação e na efetividade da transparência dos dados governamentais ganha relevância perante o impacto negativo para a Administração Pública e para toda sociedade, nos casos de fraudes em processos de compras realizadas pelo Governo Federal. No entanto, acredita-se que a maior parte das fraudes não são descobertas a tempo de serem combatidas, e aquelas que vêm ao conhecimento público podem ser apenas a ponta do iceberg, baseadas em denúncias e/ou resultados de ações pontuais. Dessa forma, constitui como problemática deste artigo a seguinte questão: *é possível ter mais controle social nos pregões eletrônicos por meio da análise de dados governamentais abertos?*

O objetivo deste trabalho foi ampliar o entendimento sobre o cenário atual de possíveis ocorrências de *bots* em pregões eletrônicos e sobre como a análise de dados abertos pode ajudar no combate às fraudes.

Esse estudo traz como ideia central a possibilidade de procedimentos que podem ser replicáveis, para fins de identificação de padrões e anomalias em pregões eletrônicos, com os dados fornecidos pelo Governo Federal. Justifica-se em possíveis ganhos para administração pública e economia aos cofres públicos. Aponta como benefícios à sociedade a possibilidade de qualquer cidadão ser capaz de exercer controle social de forma autônoma e gratuita, desde que tenha acesso à internet e um software de análise de dados, viabilizando, desta forma, a contínua investigação sobre os efeitos e impactos reais do uso de *bots* e de outros fenômenos.

Trata-se de uma pesquisa de natureza descritiva, pois busca revelar a situação atual do uso de *bots* em pregões eletrônicos de uma forma objetiva. Metodologicamente, segue uma abordagem descritiva, tendo como intuito lançar ideias iniciais que podem ser mais bem exploradas em estudos futuros de cunho mais empírico e aplicado, como a averiguação de casos de uso ou testes de variáveis e análise de parâmetros sobre uma massa de dados coletados para constatação de evidências atuais do uso de *bots* nos pregões eletrônicos.

2 COMPRAS GOVERNAMENTAIS

As compras governamentais afetam a vida de todos os cidadãos através da qualidade dos serviços prestados e da infraestrutura oferecida à sociedade. São determinantes para a eficiência da alocação de recursos públicos e injetam dinheiro no mercado através de pagamentos aos fornecedores contratados. O termo “compras governamentais”, segundo Amorim (2017, p.2), abrange todas as formas de contratação possíveis à Administração Pública, tais como aquisição de bens, contratação de serviços, alienação de bens móveis e imóveis, concessões de serviços públicos, permissões de uso de bem público, entre outras.

O conceito de compra governamental está estritamente relacionado ao de licitação. Mello (2015, p.536) relata o “certame que as entidades governamentais devem promover e no qual abrem disputa entre os interessados em com elas travar determinadas relações de conteúdo patrimonial, para escolher a proposta mais

vantajosa às conveniências públicas”. O autor segue ressaltando que “estriba-se na ideia de competição, a ser travada isonomicamente entre os que preencham os atributos e aptidões necessários ao bom cumprimento das obrigações que se propõem assumir” (MELLO, 2015, p. 536). Tem-se então a isonomia, a igualdade e a impessoalidade como pilares das licitações.

Por trás da necessidade da realização de processo licitatório no âmbito das compras públicas, não há apenas a preocupação em garantir a integridade das disputas e a eficácia dos resultados, mas também a concretização de um objetivo ético e moral para com os cidadãos. Com isso, fomenta-se o desenvolvimento sustentável do país e torna-se a licitação um meio para alcance de políticas públicas de incentivo a economias locais, sustentabilidade ambiental, inclusão social, regularização fiscal, entre outras.

Justen Filho (2012) ressalta outros aspectos da relevância do contrato administrativo, sendo este o instrumento de formalização dos procedimentos de compras governamentais. O autor traz a noção de que há íntima relação entre a democratização do Estado e a ampliação da utilização do contrato administrativo, por meio da observância de certos procedimentos e atuação dentro de limites específicos, em detrimento à requisição compulsória de bens e serviços por um Estado autoritário.

Vinculada ao Ministério do Planejamento, Orçamento e Gestão – MP, a plataforma digital Comprasnet (Disponível em: <https://www.comprasgovernamentais.gov.br/>) constitui-se como módulo do Sistema Integrado de Administração de Serviços Gerais (SIASG) do Governo Federal, onde são operacionalizados e divulgados os pregões, demais licitações e divulgadas as contratações diretas. Além disso, há esclarecimentos e divulgação dos casos de dispensa e inexigibilidade de licitação previstos em lei. Para os potenciais fornecedores de serviços e produtos, a plataforma representa um conjunto de facilidades que desburocratiza a participação em processos licitatórios.

Vale destacar que a plataforma não se restringe apenas ao uso dos fornecedores credenciados. Ela pode ser acessada por qualquer pessoa que tenha interesse em acompanhar os avisos de licitação, as contratações realizadas, a execução de processos de aquisição pela modalidade de pregão e outras informações relativas às contratações realizadas pela Administração Federal. Além disso, é possível encontrar na plataforma do SIASG toda legislação que regula o processo de serviços gerais e de contratação, publicações relativas ao assunto, manuais relativos aos serviços oferecidos pelo Portal, além de suporte técnico e ajuda no uso da ferramenta.

2.1 Pregões eletrônicos

O pregão é uma modalidade de licitação recente, criada quase uma década após a publicação da Lei de Licitações, através da Lei nº 10.520, de 17 de julho de 2002, conhecida como Lei do Pregão. Surgiu como uma alternativa de processo licitatório mais eficiente na aquisição de serviços e bens. O art. 1º da Lei do Pregão diz que são bens e serviços “cujos padrões de desempenho e qualidade admitam definição objetiva no edital, por meio de especificações usuais de mercado” (BRASIL, 2002).

As despesas licitadas através de pregão eletrônico corresponderam a mais de 94% das licitações executadas pela esfera federal entre 2012 e 2017 (PAINEL DE COMPRAS, 2018), representando mais de 84% dos gastos totais com licitações no mesmo período. Seu crescimento foi possibilitado principalmente pelos benefícios e facilidades advindos da operacionalização através da internet. Esses quantitativos se mostram ainda mais relevantes quando comparamos o volume da realização de pregões com as demais licitações, representando 94,1% do total, configurando um percentual de 85,8% de todo o gasto com licitações pelo Governo Federal no período indicado.

Vale destacar que o procedimento do pregão eletrônico, regulamentado pelo Decreto nº 5.450 de 2005, é composto por duas fases, interna e externa. A fase interna corresponde à fase preparatória e abrange a identificação da demanda, elaboração de todos os documentos necessários ao detalhamento da contratação e das regras do pregão propriamente dito. A fase externa tem como núcleo o que ocorre na chamada sessão pública, liderada pelo pregoeiro e realizada em conjunto com uma equipe de apoio, todos devidamente nomeados por alguma autoridade regimentalmente competente. É então iniciado um processo de várias etapas, conforme descrito no Quadro 1, simultaneamente ou não, pois os itens, em regra, são independentes entre si.

Quadro 1 - Fase Externa do Pregão Eletrônico

Módulo/Fase	Procedimentos e Ações Possíveis
Operar	<p>Início: A partir do horário previsto no edital, tem início a Sessão Pública, quando é disponibilizado chat de comunicação com os fornecedores para mensagem de abertura da Sessão Pública, que permanece disponível ao Pregoeiro por toda a sessão;</p> <p>Abrir item para lances: é realizada a análise, classificação e desclassificação das propostas para cada item. Em seguida o item é aberto para lances.</p> <p>Suspender Item: opção disponível na etapa de lances e poderá ser utilizada em situações de dúvida sobre um determinado item. A suspensão do item impedirá o envio de lances enquanto perdurar tal situação;</p> <p>Excluir Lance: tal procedimento não impedirá a continuidade do envio de lances pelos fornecedores;</p>

Módulo/Fase	Procedimentos e Ações Possíveis
	Encerrar o Item para Lances: um item encerrado não pode ser reaberto. Ao clicar em encerrar item, o pregoeiro deve informar o tempo de iminência, entre 1 e 60 minutos, em formato numérico. Após esse tempo, os itens entram no horário de encerramento aleatório do sistema no período de 1 a 30 minutos. Após o encerramento aleatório, os itens passam a constar na situação de “encerrado”, devendo o pregoeiro ir para a fase de aceitação.
Aceitar	Nesta fase, o pregoeiro analisa as propostas de melhor lance, relativamente ao objeto e o valor, decidindo por sua aceitação ou recusa.
Habilitar	Destina-se a fornecedores que tiveram suas propostas aceitas para o item.
Fechar/Reabrir Prazo de Intenção de Recurso	Nesta funcionalidade o pregoeiro informa o prazo para os fornecedores manifestarem a intenção de recurso.
Juízo de Admissibilidade	Nesta fase o pregoeiro deve acatar ou recusar a intenção de recurso do licitante. Havendo intenção de recurso aceita, o sistema disponibiliza tela própria para que o pregoeiro informe os prazos para registro da razão, da contrarrazão e da decisão do recurso.
Encerrar Sessão	O sistema eletrônico gera a Ata da Sessão Pública do pregão, com indicação do lance vencedor e demais informações relativas à Sessão Pública.
Decidir Recurso	Caso tenha sido aceito alguma intenção de recurso, findos os prazos legais para registro das razões do recurso e das contrarrazões dos licitantes, o pregoeiro informa à autoridade competente da sua decisão, que acatará ou não, emitindo sua decisão final.
Adjudicar	Decididos os recursos e constatada a regularidade dos atos praticados, a autoridade competente adjudicará o procedimento licitatório.
Homologar	É um procedimento da autoridade competente que equivale à decisão final e/ou à ratificação dos atos do pregoeiro.

Fonte: adaptado de Brasil, 2005b.

Além de facilitar a participação dos fornecedores, o pregão eletrônico gera menos desconfiança ao mitigar algumas das práticas tradicionais de conluio e fraudes em licitações presenciais. Entre os benefícios dos pregões eletrônicos está o aumento da competitividade nos certames e conseqüentemente a redução dos preços praticados nas compras governamentais e dos riscos de fraudes relacionadas a combinações entre fornecedores e licitantes.

No entanto, surgiram outras formas de desequilibrar a disputa, relacionadas à tecnologia utilizada na plataforma Comprasnet. Uma das questões tecnológicas discutidas atualmente é o uso de *bots* pelos fornecedores na fase de lances, a etapa competitiva de um pregão eletrônico.

Quadro 2 – Formas de fraude em leilões eletrônicos

Forma de fraude	Descrição
Deturpação de um produto para venda	Utiliza descrição falsa de características ou condições do produto
Não entrega de mercadorias ou serviços vendidos	Burla os mecanismos de proteção das plataformas intermediárias para receber o pagamento sem que haja produto ou serviço real a ser comercializado
Vendendo bens do mercado negro	Comercialização de bens ilícitos ou com venda proibida

Triangulação	Os fraudadores compram itens usando um cartão de crédito roubado, vendendo os itens a compradores não iniciados, retendo assim o dinheiro e transferindo o risco de apreensão para o destinatário final
Cobrança de taxas	Cobrando dinheiro extra após o término de um leilão
Lances múltiplos	Os compradores aumentam os preços usando pseudônimos, o que frustra os concorrentes e, no último momento, os lances altos são retirados para garantir um lance baixo
Figurante (<i>shill bidding</i>)	Vendedores ou seus associados fazem lances em seus próprios leilões para fins fraudulentos

Fonte: Adaptado de Dong; Shatz e Xu, 2009 [tradução livre].

Aplicada à realidade das licitações brasileiras, mais especificamente à dos pregões eletrônicos, as duas últimas formas listadas Quadro 2, pois são as mais recorrentes e podem ser potencializadas pelo uso de *bots* como será melhor explicado a seguir.

2.2 Uso de bots em pregões eletrônicos

Os *bots*, no contexto da Internet, são software de computador que são programados para um conjunto de tarefas automatizadas (MONAHAN, 2018). Na literatura nacional, o termo “robô eletrônico” é amplamente utilizado como um sinônimo de *bot*. Será adotado o termo original em inglês por trazer consigo significado mais específico, podendo o termo “robô eletrônico” ser utilizado, eventualmente, em contextos diferentes do aqui estudado.

No pregão eletrônico, os *bots* atuam geralmente gerenciando a emissão de lances simultaneamente nos itens que estiverem disputando, de acordo com a programação definida e como se fossem os próprios representantes das empresas. A questão principal relacionada à sua utilização é que eles podem tornar a disputa desequilibrada, favorecendo quem os utiliza, em detrimento da capacidade humana limitada de resposta durante a disputa de lances. Ao mesmo tempo, não há norma que a proíba explicitamente. O que há são, em sua maioria, algumas fundamentações jurídicas que se valem de princípios que versam sobre o equilíbrio entre as chances de participação em licitações e pregões, porém ainda não trata a questão dos *bots* com clareza.

De forma a se proteger dos malefícios do uso de *bots* em pregões, a plataforma Comprasnet sofreu implementação de dispositivos que buscam inibir os efeitos deste tipo de tecnologia ao longo da última década, mas não foram suficientes para impedir totalmente sua prática. Continuam, assim, sendo cada vez mais aperfeiçoados e

utilizados em concorrências públicas. Uma das barreiras no desenvolvimento de mecanismos eficazes que impedem a utilização de *bots* é a dificuldades de detecção. Os lances emitidos por pessoas naturais e os emitidos por *bots* são indistinguíveis do ponto de vista de seu registro no sistema, ou seja, não há nada que indique que um lance foi emitido por um humano ou um software robô, tornando a investigação por indícios uma das únicas formas possíveis.

Nesse sentido, alguns desses indícios baseados em comportamentos anômalos por parte dos licitantes se caracterizam pela:

- a) alta quantidade de participação em pregões eletrônicos;
- b) alta frequência de lances emitidos por item que participa;
- c) alta frequência de participação simultânea em vários itens de um mesmo pregão;
- d) habitualidade no registro de lances com intervalos similares de tempo e/ou valor em relação a novos menores lances de cada item;
- e) alto índice de sucesso nas disputas em relação ao número de itens que participa.

Conforme Dong, Shatz e Xu (2009), algumas estratégias mais conhecidas do uso de *bots* em leilões e concorrências eletrônicas são:

Quadro 3 - Cinco estratégias comuns de lances

Tipo	Comportamento
Skeptico (Cético)	Dá vários lances, mas aumentando o valor o mínimo possível.
Proxy bidding (Licitante por procuração)	Dá quantos lances forem necessários até atingir o máximo previamente especificado.
Sniping (Atirador de precisão)	Dá lance nos últimos segundos, não dando tempo para outros cobrirem.
Unmasking (Desmascarador)	Dá vários lances em um curto espaço de tempo com o objetivo de expor o lance máximo ou os participantes mais competitivos.
Evaluator (Avaliador)	Dá apenas um lance no início com um valor alto.

Fonte: Dong, Shatz e Xu, 2009 [tradução livre].

Por fim, fica evidenciado que os *bots* procuram replicar o comportamento humano a todo custo, porém, podem ser identificados pelo tempo de resposta e pela capacidade de carga de trabalho que são incompatíveis com o ser humano. Encontrar outras formas de identificar e restringir tais práticas é um desafio. Neste estudo, um dos caminhos trilhado foi trabalhar com a análise dos dados produzidos durante os pregões. O uso dos dados governamentais abertos auxilia no controle de fraudes. Apesar de este

estudo não produzir instrumentos para bloquear os *bots* durante o pregão, permitirá que as empresas que fazem uso desta tecnologia sejam identificadas. Para isso, o primeiro passo é entender o que são os dados abertos e como acessá-los.

3 DADOS GOVERNAMENTAIS ABERTOS

O programa de desenvolvimento e implementação de estratégias de governo digital, da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), apresenta 12 princípios fundamentais para o êxito, dentre os quais está a criação de uma cultura orientada a dados no setor público (OCDE, 2018). Aponta que os governos devem desenvolver uma cultura de análise e uso de dados públicos que ajude a prever novas necessidades e tendências e a entender como melhorar os processos e dinâmicas existentes.

No Brasil, isso tem sido reforçado por meio de políticas de dados abertos no âmbito do *e-government (e-gov)*, ou governo eletrônico. Segundo definição da *Open Knowledge Internacional*, organização global sem fins lucrativos focada em fomentar a cultura de dados abertos, dados são considerados abertos quando qualquer pessoa pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo, exigências que visem preservar sua proveniência e sua abertura (OPEN KNOWLEDGE INTERNACIONAL, 2019). Na prática, esses critérios têm sido satisfeitos pelas entidades governamentais através da criação de portais que disponibilizam conjuntos de dados estruturados (linhas de dados dispostos em colunas pré-determinadas) e organizados sob uma licença aberta.

O Governo Federal, através do agora denominado Ministério da Economia, vem expandindo e aprimorando suas estratégias de governança digital, incluindo ações voltadas à ampliação da transparência e à abertura dos dados. Publicou em 2016 a Estratégia do Governança Digital, o EGD (BRASIL, 2016), para o período de 2016 a 2019, alinhado com a Política de Governança Digital (instituída pelo Decreto nº 8.638, de 15 de janeiro de 2016) e com o Plano Plurianual (PPA) 2016-2019. O documento é composto por três eixos, dez objetivos e 51 iniciativas que visam integrar a atuação governamental na vida das pessoas através da tecnologia. Um dos desafios a serem enfrentados é o de “mitigar as vulnerabilidades de segurança nos sistemas de informação governamentais” (BRASIL, 2016, p. 12).

A transparência, a segurança e a integridade das informações que se encontram nos sistemas do governo são aspectos fundamentais para que todos os esforços em governança digital se transformem em confiabilidade para a sociedade. Destacam-se tanto que uma das iniciativas da EGD 2016-2019 (a IE.03.06) é “promover cooperação nacional e internacional com setor produtivo e academia, visando à troca de experiências e o fortalecimento dos temas de Segurança da Informação e Comunicação e de Segurança Cibernética” (BRASIL, 2016, p. 29), indicando ser um dever de todos, como cidadãos, contribuir para a promoção da consolidação do Governo Eletrônico.

Segundo a CGU (2018), na divulgação do 4º Plano de Ação Nacional em Governo Aberto, este é um conceito que:

se refere a uma nova visão da Administração Pública, que promove projetos e ações voltados ao aumento da transparência, à luta contra a corrupção, ao incentivo à participação social e ao desenvolvimento de novas tecnologias que tornem os governos mais responsáveis por suas ações e preparados para atender às necessidades dos cidadãos. (CGU, 2018, p. 6).

Como efeito, busca trazer ferramentas que possam ser utilizadas pelo cidadão em uma perspectiva de participação ativa, trazendo-o ao foco das políticas públicas. Considerando o cenário histórico político nacional, representa uma profunda quebra de paradigmas. Neste novo cenário, alguns princípios ganham ênfase, tais como a *Accountability* aplicada ao conceito de Governo Aberto:

Accountability (prestação de contas e responsabilização) – um governo responsivo estabelece normas, regulamentos e mecanismos que obrigam atores governamentais a justificar ações, agir de acordo com as críticas ou exigências colocadas e aceitar a responsabilidade no cumprimento de seus deveres. Participação social – um governo participativo promove a participação ativa da sociedade nos processos de formulação das políticas públicas e a criação de novos espaços de interlocução que favoreçam o protagonismo e o envolvimento dos cidadãos. Além disso, viabiliza que as administrações públicas se beneficiem do conhecimento, das ideias e da experiência dos cidadãos. Transparência – um governo transparente proporciona informações sobre o que está fazendo, sobre seus planos de ação, suas fontes de dados e suas atribuições perante a sociedade. Além disso, estimula a prestação de contas da administração e as ações permanentes de controle social. Tecnologia e Inovação – um governo inovador entende o papel das novas tecnologias na inovação e busca assegurar aos cidadãos a capacidade de utilizarem novas ferramentas de desenvolvimento disponíveis e o acesso a elas. (CGU, 2018, p. 6).

O tema traz uma forte relação com as compras e contratações governamentais. Nesse sentido, o Governo Federal publicou no ano de 2017 o Referencial de Governança e Gestão do Sistema de Serviços Gerais – SISG (2017). Um dos sistemas que compõem o SISG é o Comprasnet. Os indicadores divulgados na referida publicação (Quadro 4)

ilustram o volume das transações realizadas em alguns módulos do sistema, e revelam a importância de se garantir a integridade de suas informações:

Quadro 4 - Dados da utilização do SISG

Descrição	Valor
UASG (ativas)	5.033 SISG e 5.067 Não SISG;
Usuários ativos	109.361;
Fornecedores Cadastrados (ativos)	367.403;
Pregoeiros cadastrados	18.840;
Processos de Compras em 2016	105 mil (órgãos SISG);
Valores Homologados em 2016	R\$ 35 bilhões (órgãos SISG) até 02/12/2016;
Contratos Assinados em 2016	17.896 (órgãos SISG) até 02/12/2016;
Valores Contratos em 2016	R\$ 30 bilhões (órgãos SISG) até 02/12/2016.

Fonte: Brasil (2017, p. 15).

Nesse contexto, a política de dados abertos do Governo Federal pode e deve se tornar fonte de buscas e pesquisas em compras governamentais, pois é uma ferramenta de controle social que alcança volumes financeiros representativos. Na prática, esta política permite que qualquer cidadão tenha acesso aos dados governamentais, basta solicitar ao órgão responsável pela guarda dos dados.

Para que isso se torne uma realidade, faz-se necessário intensificar campanhas de conscientização da população sobre temas como dados abertos, governança e controle social. Além disso, é preciso implementar mais obrigatoriedade e rigidez na divulgação dos dados abertos por parte de diferentes instâncias do Governo Federal. Considera-se que isso proporcionará mais alinhamento com diretrizes internacionais e princípios mundiais de governança.

4 MINERAÇÃO DE DADOS GOVERNAMENTAIS ABERTOS

Ter acesso aos dados é uma etapa deste trabalho, tratar e encontrar as informações nesse contexto é a outra parte. As ferramentas para tratamento e mineração de dados podem ser úteis nestas atividades. Segundo Lin (2005), a mineração de dados é um processo que desenha padrões úteis de dados e busca expor quatro

questionamentos fundamentais: (1) quais os dados, (2) quais os padrões, (3) qual a lógica para desenhar os padrões dos dados e (4) como os padrões estão relacionados ao mundo real (sua utilidade).

Já para Hand, Mannila e Smyth (2001), Mineração de Dados pode ser entendida como a análise de conjuntos de dados (geralmente grandes volumes) observacionais para encontrar relações não previsíveis numa análise comum e resumir os dados de maneiras novas que sejam compreensíveis e úteis para o proprietário dos dados. Edelstein (1999) diz que a mineração de dados é um processo que usa uma variedade de ferramentas de análise de dados para descobrir padrões e relacionamentos em dados que podem ser usados para fazer previsões válidas.

Diante das definições apresentadas, podemos compreender que a mineração de dados configura-se com um processo apoiado por ferramentas tecnológicas, como softwares que possuem alta capacidade de processamento e reestruturação dos dados impossíveis de serem alcançada pelo ser humano. Tais ferramentas conseguem identificar padrões e relações não facilmente identificados de outra maneira. A área de mineração tem inúmeras possíveis aplicações, como: padrão de consumo, mercado de ações e capitais, dados espaciais, análise de riscos, segurança, catástrofes naturais, entre tantas outras.

Sendo assim, a mineração de dados apresenta-se como um meio para a investigação da problemática levantada neste estudo. Um primeiro ponto que deve ser considerado é que o Comprasnet possui uma base de dados que concentra grande parte das compras governamentais em âmbito nacional, o que gera um grande volume de dados estruturados. O segundo ponto é devido a essa base de dados se enquadrar nos princípios dos dados abertos, o que possibilita amplo acesso por qualquer cidadão interessado. Em terceiro, a identificação do uso de *bots* é um problema passível de ser detectado, através da análise dos dados, usando campos como tempos de resposta entre lances, diferença de valor entre lances ou licitações vencidas pelos mesmos fornecedores

No contexto da identificação do uso de *bots* em pregões eletrônicos, a mineração pode ser utilizada principalmente no rastreamento de padrões repetitivos, que caracterizem um comportamento não humano, através do processo sistemático chamado Descoberta de Conhecimento em Bases de Dados (do inglês *Knowledge Discovery in Databases*, ou KDD). Neste estudo, adotamos mineração de dados e KDD como sinônimos.

Han, Kamber e Pei (2012, p. 6/8, grifo dos autores, tradução nossa) descrevem o processo de descoberta de conhecimento como uma sequência interativa das seguintes etapas:

- 1) **Limpeza de dados:** para remoção de ruído e dados inconsistentes;
- 2) **Integração de dados:** onde várias fontes de dados podem ser combinadas;
- 3) **Seleção de dados:** onde os dados relevantes para a tarefa de análise são recuperados do banco de dados;
- 4) **Transformação de dados:** em que os dados são transformados e consolidados em formulários apropriados para mineração, executando operações de resumo ou agregação;
- 5) **Mineração de dados:** processo essencial em que métodos inteligentes são aplicados para extrair padrões de dados;
- 6) **Avaliação de padrões:** para identificar os padrões verdadeiramente interessantes que representam o conhecimento com base em medidas de interesse);
- 7) **Apresentação do conhecimento:** onde técnicas de visualização e representação de conhecimento são usadas para apresentar conhecimento minado aos usuários.

Já para os autores De Castro e Ferrari (2017, l. 660-673, grifos dos autores), o processo de KDD possui quatro etapas:

- 1) **Base de dados:** coleção organizada de dados, ou seja, valores quantitativos ou qualitativos referentes a um conjunto de itens, que permite uma recuperação eficiente dos dados. Conceitualmente, os dados podem ser entendidos como o nível mais básico de abstração a partir do qual a informação e, depois, os conhecimentos podem ser extraídos [...];
- 2) **Preparação ou pré-processamento de dados:** são etapas anteriores à mineração que visam preparar os dados para uma análise eficiente e eficaz. Essa etapa inclui a limpeza (remoção de ruídos e dados inconsistentes), a integração (combinação de dados obtidos a partir de múltiplas fontes), a seleção ou redução (escolha dos dados relevantes à análise) e a transformação (transformação ou consolidação dos dados em formatos apropriados para a mineração);
- 3) **Mineração de dados:** essa etapa do processo corresponde à aplicação de algoritmos capazes de extrair conhecimentos a partir dos dados pré-processados. [...] [Inclui] técnicas de análise descritiva (medidas de distribuição, tendência central e variância, e métodos de visualização), agrupamento (segmentação de bases de dados), predição (classificação e estimação), associação (determinação de atributos que concorrem) e detecção de anomalias;
- 4) **Avaliação ou validação do conhecimento:** avaliação dos resultados da mineração objetivando identificar conhecimentos verdadeiramente úteis e não triviais.

Para realizar análise dos pregões eletrônicos, a primeira tarefa é realizar o cadastro solicitando acesso aos dados ao Governo, através do Sistema Eletrônico do Serviço de Informação ao Cidadão (e-SIC), mantido pelo Ministério da Transparência e Controladoria Geral da União (CGU). Outro ponto é conhecer a API (em português: Interface de Programação de Aplicativos) da plataforma Comprasnet (Disponível em:

<http://compras.dados.gov.br/docs/home.html>). Em termos simples, API pode ser considerada como a documentação de um aplicativo, plataforma ou sistema que permite o usuário buscar dados a respeito das regras de negócio e possibilidades previstas de uso do sistema. Desse modo, qualquer pessoa pode solicitar as extrações de dados do Comprasnet. É importante ressaltar que ainda há um espectro limitado das informações disponibilizadas sobre os eventos de pregões eletrônicos no âmbito do Governo, principalmente referentes aos metadados, variedades de atributos e dicionário de dados. Ou seja, uma possível ampliação dos dados disponíveis para autoextração também poderia contribuir com a melhoria da transparência dos dados das compras governamentais federais.

A mineração de dados se fundamenta em dados guardados em Banco de Dados (Database ou simplesmente DB). Para fins práticos deste estudo, DBs podem ser entendidos como repositórios onde todas as informações sobre um processo, inclusive metadados, podem ser gravadas e mantidas para consulta posterior. Geralmente os dados são armazenados em um DB do tipo relacional e estão estruturados em tabelas. Por sua vez, as tabelas possuem linhas e colunas sobre um determinado objeto (entidade) analisado. As colunas representam as características deste objeto e as linhas representam o conjunto dessas características para uma determinada ocorrência do objeto. As tabelas podem ser cruzadas entre si, revelando novos fatos e padrões entre os dados analisados.

Para exemplificar, no caso de dados de um pregão eletrônico, podemos imaginar a existência de uma tabela com fornecedor que possui N atributos (CNPJ, endereço, telefone, data de cadastro, etc.), em que cada linha representa um conjunto de informações sobre determinado fornecedor cadastrado.

No processo de mineração de dados, o termo atributos representa as palavras que aparecem numa base de dados coletada. Nesse sentido, De Castro e Ferrari (2017) consideram que ferramentas simples, como histogramas ou gráficos, que possibilitem a visualização da distribuição de valores de um atributo, ou que permitam a comparação entre vários atributos, são bastante úteis para um entendimento inicial da base de dados. Para o pré-processamento, recomenda-se que, na importação dos dados, todas as colunas sejam tratadas como nominais, facilitando a visualização das informações relacionadas à distribuição, quantidade e frequência de valores, assim como a incidência de valores idênticos, entre outros, pois em colunas do tipo numéricas apresentam

métricas como média aritmética, mínimo, máximo, média, entre outros. A análise inicial de distribuição dos valores pode facilitar uma melhor compreensão do que representa cada atributo e sua utilidade para uma pesquisa a ser realizada, assim como contribui com a eliminação de duplicidades nos atributos.

Atualmente, há uma série de ferramentas que realizam mineração de dados, muitas delas voltadas para fins específicos ou focadas em etapas específicas da análise de dados. Algumas outras são aplicadas mais amplamente. São exemplos dessas ferramentas: WEKA, Lucene, RapidMiner, DataMelt, OpenRefine, Tableau, Orange, Knime Analytics Platform, entre tantas outras. Há soluções tanto pagas como gratuitas, e há ferramentas que podem exigir um conhecimento mais avançado dos usuários, como domínio de linguagem de programação específica, como Java, Python, R, etc. Já outras exigem apenas conhecimentos básicos em análise estatística de dados e manuseio do computador, sem necessidade de conhecimento em programação.

Dessa forma, uma ferramenta bastante conhecida e com licença educacional gratuita de um ano é a RapidMiner (Disponível em: <http://www.belge.com.br/rapidminer.php>), a qual pode ser uma opção para quem deseja analisar os dados governamentais abertos, tais como dados dos pregões. Além disso, não é necessário saber linguagem de programação nem conhecimento específico para operar na ferramenta e alcançar resultados satisfatórios.

Iniciativas de mineração de dados aplicadas em processos de pregões eletrônicos podem também ser de muita valia para funcionários/servidores que trabalham em organizações Estatais vinculadas à Administração Pública Federal, na identificação de padrões de anormalidades em compras públicas, evitando assim beneficiamento de fornecedores fraudulentos, em prejuízo ao erário e toda sociedade. Vale ressaltar que, embora o foco deste trabalho tenha sido a base de dados do Governo Federal, o uso de análise de dados abertos a favor da construção de uma sociedade mais justa e consciente dos seus deveres pode ocorrer em qualquer esfera do poder executivo (estadual, municipal, federal), assim como nos demais poderes (judiciário e legislativo), além de organizações paraestatais e do terceiro setor.

5 CONSIDERAÇÕES FINAIS

Ressalvadas as hipóteses legais, através das licitações são selecionados os fornecedores para as contratações na esfera pública, seja para aquisição de bens, fornecimento de serviços ou realização de obras. Os recursos que custeiam essas despesas são públicos e executados através de agentes de governo, por isso é necessário que sejam estabelecidos princípios e regras que busquem garantir a integridade e finalidade desses processos.

Há princípios constitucionais que podem ser evocados para garantir mais controle sob os processos licitatórios, assim como leis específicas que versam sobre o tema Licitação e contratos públicos (Lei nº 8.666) e a Lei do Pregão Eletrônico (Lei nº 10.520), entre outras. Mesmo assim, constatamos muitos aspectos ainda tratados sem a clareza necessária para o discernimento sobre questões polêmicas, como o exemplo abordado neste trabalho que apura a utilização *bots* de em pregões eletrônicos.

Com isso, percebe-se que a mineração de dados pode ajudar na solução de alguns problemas referentes à análise de grandes volumes de dados, identificando conhecimentos úteis e não triviais, que seriam dificilmente observados por uma análise puramente humana. Assim, pode exercer papel fundamental na identificação de anomalias e indícios do uso de *bots* em pregões eletrônicos.

Apesar da complexidade e quantidade de massa de dados com que as ferramentas de mineração de dados conseguem trabalhar, o uso e a compreensão dessas ferramentas podem estar ao alcance de qualquer cidadão, não necessitando ser um pesquisador nem um especialista em análise de dados. Em muitos casos, nem mesmo é necessário ter domínio de alguma linguagem de programação, o que ajuda bastante na difusão dos benefícios a serem obtidos com análise desse tipo de dados.

A ausência do fornecimento de um dicionário de dados pode gerar como resultado uma extração de dados de baixa qualidade, o que evidencia a característica de um serviço não orientado ao usuário. Dessa forma, há a necessidade de um comprometimento maior por parte dos órgãos governamentais na disponibilização de dados e metadados decorrentes dos processos de pregões eletrônicos.

Por fim, verifica-se que o objetivo deste estudo de identificar o uso de *bots* foi alcançado, tendo em vista que ficou evidenciado o potencial da análise de dados extraídos de pregões. Estes podem ser submetidos ao tratamento e análise por

ferramentas de mineração de dados, ajudando a identificar ações fraudulentas, ou na obtenção de vantagens indevidas, como no uso dos *bots*.

Destaca-se a função de controle social apoiada por uma política estruturada de dados abertos, estabelecendo-se uma via de mão-dupla: à medida que entendermos a importância dos dados abertos e os reflexos em suas vidas, enquanto cidadãos, poderemos melhor cobrar e apontar irregularidades na gestão dos recursos públicos. Por outro lado, à medida que os Entes Públicos disponibilizarem dados com mais qualidade e precisão, ajudarão a motivar que novos estudos e padrões sejam identificados em várias vertentes do conhecimento, por um número cada vez maior de pessoas, desde pesquisadores e acadêmicos, até cidadãos comuns.

REFERÊNCIAS

AMORIM, V. A. J. **Licitações e contratos administrativos**: teoria e jurisprudência. Brasília: Senado Federal, Coordenação de Edições Técnicas, 2017.

BRASIL. Lei nº 8.666, de 21 de junho de 1993. **Lei de Licitações**. Brasília. Disponível em: http://www.planalto.gov.br/ccivil_03/Leis/L8666compilado.htm. Acesso em: 14 jan. 2018.

BRASIL. Lei nº 10.520, de 17 de julho de 2002. **Lei do Pregão**. Brasília. Disponível em: http://www.planalto.gov.br/CCivil_03/leis/2002/L10520.htm. Acesso em: 14 jan. 2018.

BRASIL. Decreto nº 5.450, de 31 de maio de 2005. **Regulamento do Pregão Eletrônico**. Brasília. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/decreto/d5450.htm. Acesso em: 15 jan. 2018.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Logística e Tecnologia da Informação. **Estratégia de Governança Digital da Administração Pública Federal 2016-19**. Brasília: MP, 2016.

BRASIL. Ministério do Planejamento, Desenvolvimento e Gestão. **Referencial de Governança e Gestão do Sistema de Serviços Gerais – SISG** / Secretaria de Gestão. Brasília, 2017a.

CGU. Ministério da Transparência e Controladoria-Geral da União. **4º Plano de Ação Nacional em Governo Aberto**. Brasília, 2018. Disponível em: http://governoaberto.cgu.gov.br/no-brasil/planos-de-acao-1/copy_of_3o-plano-de-acao-brasileiro/4o-plano-de-acao-nacional_portugues.pdf. Acesso em: 3 dez. 2018.

DE CASTRO, L.N.; FERRARI, D.G. **Introdução a Mineração de Dados**. São Paulo: Saraiva. Edição do Kindle, 2017.

DONG, F.; SHATZ, S. M.; XU, H. Combating Online In-Auction Fraud: Clues, Techniques and Challenges. **Computer Science Review**, v. 3, n. 4, p. 245-258. University of Illinois at Chicago, Chicago, U.S. National Science Foundation. 2009.

EDELSTEIN, Herbert A. **Introduction to Data Mining and Knowledge Discovery**. 3. ed. Potomac: Two Crows Corporation, 1999.

HAN, Jiawei; KAMBER, Micheline; PEI, Jian. **Data Mining: Concepts and Techniques**. 3. ed. Waltham: Elsevier, 2012.

HAND, David; MANNILA, Heikki; SMYTH, Padhraic. **Principles of Data Mining**. Cambridge, The MIT Press, 2001.

JUSTEN FILHO, M. **Comentários à lei de licitações e contratos administrativos /** Marçal Justen Filho. 15. ed. São Paulo: Dialética, 2012.

LIN, Tsau Young. Mathematical Foundation of Association Rules – Mining Associations by Solving Integral Linear Inequalities. *In*: LIN, Tsau Young *et al.* (ed.). **Foundations of Data Mining and Knowledge Discovery**. Berlim: Springer, 2005. (Studies in Computational Intelligence, v. 6).

MELLO, C. A. B. **Curso de direito administrativo**. 32. ed. São Paulo: Malheiros, 2015.

MONAHAN, D. Bot Defense: Insights Into Basic And Advanced Techniques For <https://painelcompras.planejamento.gov.br> Acesso em: 14 jan. 2018.

PAINEL DE COMPRAS. Ministério do Planejamento, Orçamento e Gestão. **Painel**

Thwarting Automated Threats. **Enterprise Management Associates**. 2016. Disponível em: <https://www.enterprisemanagement.com/research/asset.php/3317/Bot-Defense:-Insights-Into-Basic-and-Advanced-Techniques-for-Thwarting-Automated--Threats>. Acesso em: 4 nov. 2018.

OCDE. **OECD Digital Government Toolkit: 12 principles**. 2018. Disponível em: <http://www.oecd.org/governance/digital-government/toolkit/12principles/>. Acesso em: 25 nov. 2018.

OCDE. **The E-government Imperative**. Paris: OCDE, 2003.

OPEN KNOWLEDGE INTERNACIONAL. **What is Open?**, Disponível em: <https://okfn.org/about/>. Acesso em: 29 mar. 2019.

PAINEL DE COMPRAS. Ministério do Planejamento, Orçamento e Gestão. **Painel de Compras do Governo Federal**. Disponível em: **de Compras do Governo Federal**. Versão 1.4. 2018. Disponível em: <http://painelcompras.planejamento.gov.br>. Acesso em: 26 nov. 2018.