

# Exposer: Framework para Detecção de Anomalias em Licitações Públicas

Ednael F. V. da Silva<sup>1</sup>, Gabriel A. M. Fragoso<sup>1</sup>, Liviany R. Rodrigues<sup>1</sup>,  
Nicksson Ckayo Arrais de Freitas<sup>2</sup>, Tiago Da Silva Vinuto<sup>2</sup>

<sup>1</sup>Centro de Informática – Universidade Federal de Pernambuco (UFPE)  
Recife - PE - Brasil

<sup>2</sup>SiDi - Recife - PE - Brasil

{efvs, gamf, lrr}@cin.ufpe.br, {nicksson.a, t.vinuto}@sidi.org.br

**Abstract.** *Public spending must be constantly monitored by government agencies to minimize losses from bid rigging. In the literature, to our awareness, there are a gap on approaches or systems that are freely available to detect anomalies in bidding directing investigations to the legality of the amounts bid. Facing this scenario, we created Exposer, tool that makes it possible to detect anomalies and irregularities in bidding processes, with the perspective of use by officials and authorities and serving as a decision support system. In this sense, we conducted experiments with public data from the Transparency Portal of the Federal Government and Exposer revealed potential irregular bids that total values in the range of billions of Brazilian real.*

**Resumo.** *Os gastos públicos devem ser constantemente monitorados pelos órgãos governamentais para minimizar os prejuízos acerca dos desvios de dinheiro em licitações. Na literatura, para nossa consciência, existe uma lacuna quanto a ferramentas livres que possam detectar anomalias de legalidade nos valores de licitações públicas. Diante deste cenário, criamos o Exposer<sup>1</sup>, ferramenta que possibilita detectar anomalias e irregularidades em processos licitatórios, que tem perspectiva de uso por oficiais e autoridades e servir como um sistema de apoio à decisão. Neste sentido, conduzimos experimentos com dados públicos do Portal da Transparência do Governo Federal e o Exposer revelou indícios de licitações irregulares que somam bilhões de reais.*

## 1. Introdução

Licitação é a forma na qual a administração pública pode comprar e vender itens ou serviços [Transparencia 2021a]. No Brasil, foram contratados acima de 18 mil licitações movimentando cerca de 55,28 bilhões de reais somente no ano de 2021 [Transparencia 2021b]. Devido ao grande volume de transações financeiras, os gastos públicos precisam ser constantemente fiscalizados pelos órgãos governamentais para detectar anomalias e tomar decisões para minimizar os prejuízos acerca dos desvios de dinheiro em licitações. Uma anomalia é um evento anormal ocorrido fora dos padrões esperados [Bueno 2017, Velasco et al. 2021]. Analisar este volume de dados requer, geralmente, o uso de algoritmos e técnicas de aprendizado de máquina ou mineração de dados.

<sup>1</sup>Vídeo demonstrativo: <https://youtu.be/XDEsuSHWGso>

Nesta perspectiva, propusemos a criação do Exposer, uma ferramenta de apoio a decisão que exhibe o panorama nacional das licitações e pode ser utilizada pelos órgãos de auditoria e fiscalização para detectar possíveis anomalias em licitações, tais como: discordância com a Lei e presença de montantes exorbitantes. Esta ferramenta é baseada em algoritmos e técnicas de mineração de dados desenvolvidos pelos autores, e tem como base a Lei nº 8.666/93, responsável por estabelecer as normas gerais de licitações e contratos administrativos.

As principais contribuições deste trabalho são: (1) criação de uma nova abordagem de mineração de dados para detecção de anomalias; (2) elaboração de um software livre, que detecta irregularidades em compras públicas, chamado Exposer; (3) propomos três *Key Performance Indicators* (KPI's) por estado para monitorar as possíveis fraudes por dimensões geográficas; (4) uso do algoritmo Apriori nos itens das licitações com irregularidades, e por fim, (5) criação de uma lista dos *Top Losers*<sup>2</sup> e *Top Winners*<sup>3</sup> das empresas, assim como seu número de vitórias em processos categorizados como ilegais.

## 2. Exposer - Framework para Detecção de Anomalias em Licitações Públicas

Nesta seção, apresentamos o Exposer, cujo objetivo é facilitar o processo de descoberta de irregularidades e anomalias, desde o tratamento, pré-processamento, agrupamento e extração de informações úteis das fontes de dados, de modo a contribuir na problemática por meio de algoritmos e funcionalidades. A ferramenta terá livre acesso, e tem a perspectiva social de uso por oficiais e autoridades como um sistema de apoio à decisão.

### 2.1. Principais Funcionalidades

Dentre as funcionalidades do Exposer destacam-se:

- *Identificar a legalidade das licitações com base na legislação vigente;*
- *Detectar relacionamentos relevantes entre variáveis;*
- *Apresentar KPI's Estaduais de índice de irregularidade e prejuízo monetário;*
- *Detectar os TopWinners e TopLosers;*
- *Investigar as associações entre itens das licitações identificadas como ilegais.*

### 2.2. Arquitetura e Algoritmos do Exposer

Nesta seção, apresentamos a arquitetura, as técnicas de pré-processamento e os algoritmos utilizados no Exposer, sua arquitetura pode ser visualizada na Figura 1.

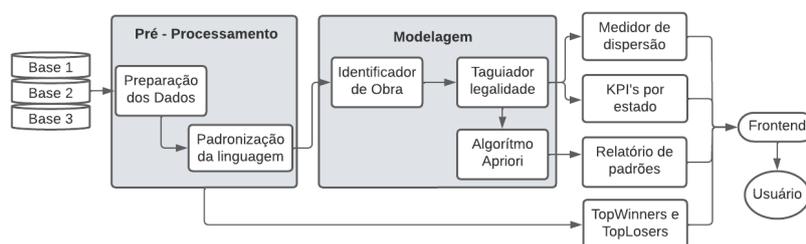
#### 2.2.1. Pré-Processamento

Nesta etapa, se fez necessário preparar, organizar e estruturar os dados, objetivando aumentar a eficiência e velocidade de obtenção dos resultados finais.

**Tratamento dos dados:** Na preparação dos dados, é realizado um ajuste nas bases, com objetivo de facilitar as etapas seguintes do trabalho, por meio da eliminação/correção de imperfeições elementares nos dados. Tratamos as seguintes imperfeições: tipos dos dados colunares e remoção das instâncias com dados faltantes e/ou dados duplicados.

<sup>2</sup>empresas que perdem todos ou quase todos os processos que participam

<sup>3</sup>empresas que ganham todos ou quase todos os processos que participam

**Figura 1. Arquitetura da abordagem**

**Padronização da linguagem:** A base de dados de itens apresentou várias instâncias da *feature* “Objeto” com escritas sem padronização, erradas e/ou com a presença de símbolos, sendo necessário duas ações: retirada de pontuações diversas dos textos e padronização de caracteres e escrita.

### 2.2.2. Modelagem

Nesta subseção, explicaremos o processo de descoberta e extração de novos recursos para os dados brutos, a partir do conhecimento do domínio, por intermédio do desenvolvimento de três algoritmos autorais.

**Identificador de obras:** Cria uma nova coluna, denominada “Tipo”, para classificar as instâncias em obras ou serviços. Utiliza como entrada a coluna “objeto” e executa uma busca por um conjunto de palavras-chave, classificando o processo de compra.

**Tagger de legalidade:** Cria uma nova coluna nos dados para classificar as instâncias em legais ou ilegais. Utiliza como entrada as colunas “Tipo” e “Total”, relacionando-as de acordo com as regras pré-estabelecidas no artigo 23 da Lei 8666/93 [Lei 2022].

**Identificador de padrões:** Utiliza o algoritmo Apriori para identificar associações entre os itens das licitações classificados como ilegais. Este foi implementado através da biblioteca “Mlxtend” e usa como entrada dados colunares “Número da Licitação” e “Descrição dos itens”.

### 2.2.3. Análises e Relatórios

Nesta subseção, detalharemos como foram realizadas as operações que resultaram nas análises apresentadas na Figura 1.

**Medidor de dispersão:** Utilizamos os valores totais de cada processo licitatório (VT) e dos limites superiores legais (LSL), previstos na Lei 8.666, para cada tipo de aquisição, o algoritmo calcula a seguinte razão:

$$\text{Medida de dispersão} = VT/LSL$$

**KPI's estaduais:** Desenvolvemos índices percentuais (KPI's), por estado brasileiro, com objetivo de condensar e simplificar a compreensão de algumas informações acerca dos prejuízos atrelados as licitações irregulares. Os quais foram:

- *Índice de irregularidade*: revela o percentual do total de licitações que está fora dos limites legais;
- *Índice de prejuízo monetário*: descreve o percentual do montante licitado que é irregular.

**Padrões em itens de compra:** Investiga a associação entre os itens das licitações com indícios de irregularidades, utilizando como base o algoritmo Apriori. Este, apresenta como saída os itens relacionados e seus respectivos valores de suporte <sup>4</sup> e confiança <sup>5</sup>.

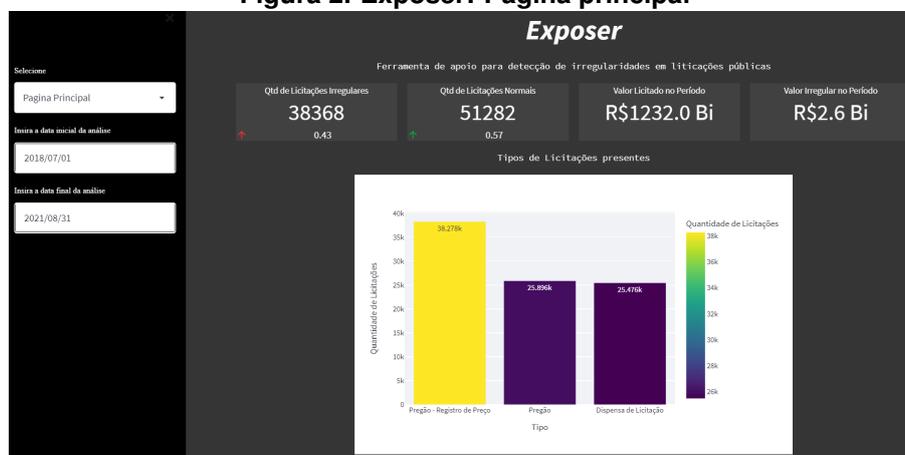
**Deteção dos *topWinners* e *topLosers*:** A fim de identificar as empresas que mais participam de processos de licitações fraudulentos, foram listados os *TopWinners* e *TopLosers*. Estes indicadores sugerem um alto risco de manipulação no processo e dar uma aparência de competitividade, fator assegurado por lei nessas modalidades.

### 2.3. Interface com o Usuário

Nesta seção, exibiremos os principais recursos presentes no Exposer, como também, a interface utilizada para apresentar as análises.

A interface gráfica é desenvolvida para ser prática, intuitiva e de fácil utilização. Como ilustrado na Figura 2, a ferramenta é composta por cinco páginas, são elas: Página principal, Lista dos TOPS, Itens, Relação dos itens e Licitações por estado.

Figura 2. Exposer: Página principal



**Página principal:** É a página de apresentação do Exposer, que mostra informações gerais acerca das licitações para o período selecionado.

**Lista dos tops:** Contém informações quanto as empresas que participam dos processos licitatórios (*TopWinners*, *TopLosers*) e a lista das TOP 20 licitações que mais ultrapassam o teto estabelecido por lei.

**Itens:** Contém informações dos itens mais licitados por estado.

**Relação entre os itens:** Contém informações sobre a associação entre os itens licitados nas licitações identificadas com indícios de irregularidades.

<sup>4</sup>Indica a porcentagem de frequência daquele produto

<sup>5</sup>Cálculo do número de registro de itens A e itens B, dividido pelo número de registros de A

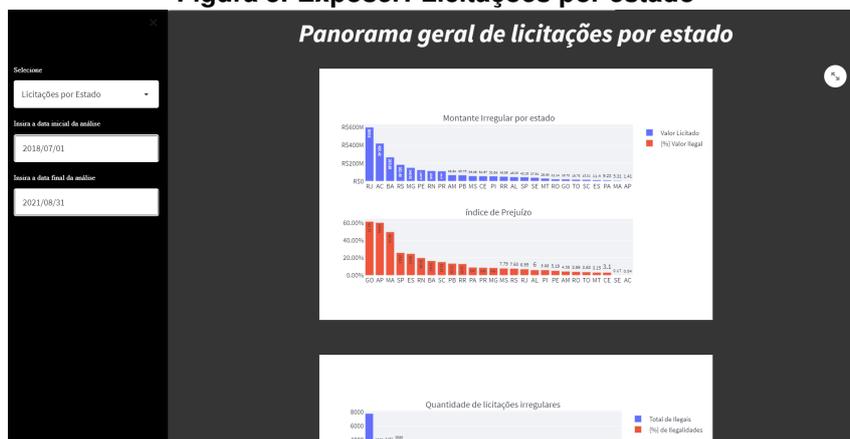
**Licitações por estado:** Contém os KPI's estaduais (índice de irregularidade e prejuízo monetário).

### 3. Resultados e Validação da Ferramenta

Nesta seção, realizamos um estudo de caso com intuito de validar as contribuições pertinentes ao Exposer. Utilizamos três bases de dados de licitações públicas do Brasil, ocorridas no período de julho de 2018 à agosto de 2021<sup>6</sup>. Estas foram baixadas no formato CSV, e posteriormente, carregadas na ferramenta.

Para validar a ferramenta, mostraremos dois resultados obtidos pelo Exposer. As Figuras 3 e 4 ilustram as telas do Exposer nas quais foram retiradas, respectivamente, as informações referentes as análises do índice de prejuízo e do índice de irregularidade. As demais análises apresentadas pelo Exposer podem ser visualizadas no link do GitHub: <https://github.com/ednaelxd/residencia2.git>

**Figura 3. Exposer: Licitações por estado**



**Análise do índice de prejuízo:** Através da tela apresentada na Figura 3 notamos que o estado do Acre, por exemplo, possui um índice de prejuízo monetário de 0,04%, o Exposer ainda revela que, apesar da baixa porcentagem, isto representa um montante expressivo de 419 milhões de reais, perdendo apenas para o Rio de Janeiro. Nota-se também que em Goiás mesmo com elevados valores percentuais (60,9%) possui baixo montante atrelado (19,7 milhões de reais).

**Análise do Índice de irregularidade:** A tela representada na Figura 4, por sua vez, revela que Goiás e Mato Grosso tem taxas de irregularidades licitatórias acima dos 69%. Enquanto isso, os estados de São Paulo e Rio de Janeiro registram 27% e 33% de índice de irregularidade, respectivamente. Estes últimos somam, em gastos anômalos, 173 e 654, milhões de reais.

Observa-se também que novamente o Acre aparece como destaque, quanto ao valor do índice e a quantidade de licitações. No estado, o órgão superior que mais licita é o Ministério da Defesa, seguido pelo Ministério da Educação e da Economia. Mais de 1300 compras públicas do Ministério da Defesa no estado, foram classificadas pelo Exposer como indícios de irregularidades, somando R\$ 129 milhões.

<sup>6</sup>Disponíveis em: <https://www.portaldatransparencia.gov.br/download-de-dados/licitacoes>

**Figura 4. Exposer: Licitações por estado**

#### 4. Considerações Finais

A identificação de licitações com práticas potencialmente fraudulentas é um problema na sociedade, neste sentido, o objetivo geral deste trabalho foi atacar o problema com a criação do Exposer. Por meio do Exposer, conseguimos apresentar um sistema capaz de detectar possíveis fraudes em processos licitatórios. A extensão desse estudo pode ser realizada por meio de atualizações da ferramenta conforme mudanças na legislação, visto que a lei vigente é atualizada periodicamente pelas autoridades. Somado a isso, pretende-se adicionar recursos de processamento de linguagem natural nos processos de identificação de obras e uniformização da escrita dos itens licitados.

#### Agradecimentos

Os resultados apresentados neste artigo foram desenvolvidos como parte de um projeto em parceria com o SiDi e financiado pela Samsung Eletrônica da Amazônia Ltda., sob os auspícios da Lei Federal de Informática nº. 8248/91.

#### Referências

- Bueno, W. F. (2017). *Analysis of Brazilian deputies expenses claims from 2013 to 2016*. PhD thesis, Dublin, National College of Ireland.
- Lei (2022). Lei 8.666/93 de licitação. [http://www.planalto.gov.br/ccivil\\_03/leis/18666cons.htm](http://www.planalto.gov.br/ccivil_03/leis/18666cons.htm). Acesso em 01-janeiro-2022.
- Transparencia, P. (2021a). Definição de licitação. <https://www.portaltransparencia.gov.br/entenda-a-gestao-publica/licitacoes-e-contratacoes>. Acesso em 26-novembro-2021.
- Transparencia, P. (2021b). Licitações contratadas. <https://www.portaltransparencia.gov.br/licitacoes?ano=2021>. Acesso em 30-novembro-2021.
- Velasco, R. B., Carpanese, I., Interian, R., Paulo Neto, O. C., and Ribeiro, C. C. (2021). A decision support system for fraud detection in public procurement. *International Transactions in Operational Research*, 28(1):27–47.