

PEP Scan - Ferramenta de Auxílio à Identificação de Possíveis Recebimentos Indevidos de Benefícios por PEPs no Brasil

Michel De Oliveira Guijarro¹, Cláudio Fabiano Motta Toledo²
ICMC-USP

1 Introdução

A identificação de pessoas politicamente expostas (PEPs) e seu possível envolvimento em esquemas de fraude e lavagem de dinheiro é crucial para mitigar os riscos associados a essas atividades ilícitas. O envolvimento de PEPs em crimes financeiros representa um desafio significativo para governos e instituições financeiras em todo o mundo [1]. De acordo com a Resolução nº 29, de 07/12/2017 [2], PEPs são indivíduos que ocupam posições públicas proeminentes ou possuem influência significativa sobre ações governamentais, ou transações comerciais devido a suas conexões, ou recursos financeiros. Portanto, o objetivo do presente trabalho em desenvolvimento é a criação de uma aplicação web, utilizando técnicas de mineração de dados e teoria dos grafos para se analisar múltiplas fontes de dados publicamente disponíveis sobre PEPs no Brasil e identificar possíveis recebimentos indevidos de benefícios, esquemas de fraude e lavagem de dinheiro.

2 Materiais e Métodos

2.1 Bases de Dados

Para a construção da prova de conceito desta aplicação web, estão sendo consideradas as seguintes bases de dados disponíveis publicamente pelo governo brasileiro:

- **PEP** [março/2023] [3]
- **Auxílio Emergencial** [abril/2022 - fevereiro/2023] [4]
- **Bolsa Família** [janeiro/2013 - novembro/2021] [5]
- **CNPJ** [março/2023] [6]

¹michelguijarro@usp.br

²claudio@icmc.usp.br

2.2 Mineração de Dados e Análise de Grafos

Para extrair, padronizar, integrar e analisar os dados, utilizou-se uma metodologia robusta que combinou técnicas de mineração de dados, transformação, automação e web scraping. Segundo Fayyad et al. (1998) [7], a mineração de dados (DM) é o núcleo do processo de Extração de Conhecimento em Bases de Dados (KDD), que visa identificar padrões e modelos válidos, novos e interpretáveis, guiando a descoberta de conhecimento nos dados.

Para aprofundar a análise e identificar conexões entre as PEPs que receberam auxílios governamentais e empresas, foi empregada a análise de grafos, destacando o algoritmo de busca em largura (Breadth-First Search - BFS). Esse método permitiu a construção de redes de relacionamentos entre empresas, PEPs e benefícios governamentais, facilitando a identificação de possíveis conexões suspeitas.

Além disso, a integração de técnicas de visualização ao processo de mineração, conhecida como Mineração Visual de Dados (VDM), reforça a análise ao permitir a exploração gráfica dos dados e a geração de hipóteses visuais que identificam padrões desconhecidos. A abordagem visual amplia o entendimento de dados heterogêneos, possibilitando insights e apoio à tomada de decisão na análise de redes complexas de relacionamentos, segundo Keim e Kriegel (1996)[8].

2.3 Arquitetura

A arquitetura do PEP Scan foi projetada para incluir etapas de extração, seleção, transformação e agregação de dados, visando uma análise focada. Além dessas etapas, foi aplicada a metodologia de Visualização Subsequente (SV), que permite aprimorar a interpretação dos padrões extraídos automaticamente por algoritmos de Data Mining, através de novo *input* do usuário.

De acordo com Keim et al. (2002) [9], a SV envolve o uso de um algoritmo de DM que identifica padrões no conjunto de dados e, em seguida, apresenta esses padrões de forma visual. Esse tipo de visualização permite ao analista revisar e refinar os parâmetros de entrada no algoritmo, orientando-o a ajustes iterativos que resultem em uma interpretação mais rica e significativa dos dados.

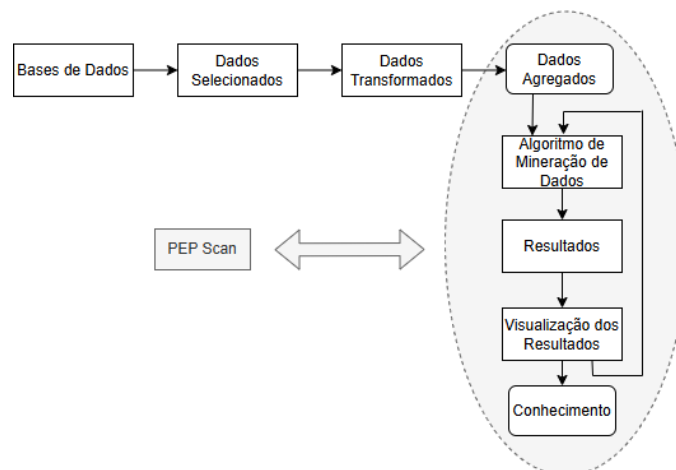


Figura 1: Arquitetura da aplicação em desenvolvimento.

3 Resultados Parciais

Na página "PEPs no Brasil" do PEP Scan, são apresentadas análises visuais da distribuição de PEPs políticos e não políticos por cargo e localização, facilitando a compreensão da concentração regional e funcional desses indivíduos. Além disso, a página oferece a possibilidade de consultar e baixar a lista completa de PEPs que receberam auxílios governamentais, ampliando o acesso aos dados e viabilizando investigações mais detalhadas.



Figura 2: Distribuição de PEPs políticos por cargo e geografia

Na página Análise de Grafos, a aplicação disponibiliza ferramentas para identificar conexões entre indivíduos e empresas, incluindo o caminho mais curto entre entidades, com uso de métodos como a busca em largura (BFS). Também é oferecida um recurso interativo de exploração de conexões, permitindo ao usuário visualizar vínculos diretos e indiretos de um nó inicial, o que facilita a identificação de possíveis relações suspeitas entre PEPs, empresas e beneficiários de auxílios governamentais, sendo um suporte visual para futuras investigações.

Figura 4: Conexão entre nós

A seguir você pode identificar tais PEPs:

Nome	CPF	Sigla da Função	UF	DT do Início do Exercício	DT do Fim do Exercício	Qtd de Meses Recebidos	Primeiro Mês Recebido	Último Mês Recebido
ADALBERTO BARBOSA DE OLIVEIRA	***996125***	VEREAD	BA	01/01/2021	31/12/2024	1	01/2022	01/2022
ADALTO LOPES DA SILVA	***127782***	VEREAD	AM	01/01/2021	31/12/2024	1	01/2022	01/2022
ADILSON ALVES DE ATAÍDE	***884778***	VEREAD	ES	01/01/2017	31/12/2020	6	07/2020	12/2020
ADILSON FERNANDES CARLEME	***185858***	VEREAD	SP	01/01/2017	31/12/2020	5	06/2020	12/2020
ADILSON FERRAZ BUZARRA	***235344***	VEREAD	FB	01/01/2021	31/12/2024	1	07/2022	07/2022
ADILSON ROBERTO MESSIAS	***311888***	VEREAD	SP	01/01/2017	31/12/2020	6	06/2020	12/2020
ADRIANO GRACIANO	***965028***	VEREAD	SP	01/01/2021	31/12/2024	1	01/2022	01/2022
AFERIANO VIGNER WISCONCELLOS DA VARA	***915109***	VEREAD	RS	01/01/2017	31/12/2020	1	11/2020	11/2020
AGUNALDO MARTINS SAMPAIO DE LANES	***848196***	VEREAD	MG	01/01/2021	31/12/2024	1	01/2022	01/2022
ALAN RICARDO BATISTA	***621409***	VEREAD	SC	01/01/2021	31/12/2024	1	01/2022	01/2022

Você pode baixar a lista completa de PEPs que receberam Auxílio Emergencial em formato CSV [Download CSV](#)

Figura 3: Lista de PEPs Políticos que receberam Auxílio Emergencial

Figura 5: Caminho mais curto entre nós

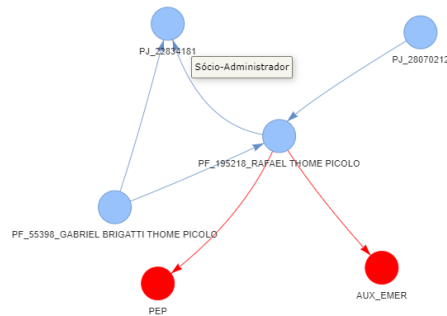


Figura 6: Caminho no grafo a partir do nó de entrada

4 Conclusões Parciais

Os resultados preliminares do PEP Scan já revelam casos de políticos em exercício que receberam auxílios governamentais, sugerindo potenciais irregularidades. Os recursos de análise de conexões mostram vínculos suspeitos entre PEPs e beneficiários, enquanto a exploração de grafos facilita a identificação visual dessas relações, direcionando futuras investigações. As próximas etapas incluem a disponibilização do recurso mostrado na Figura 6, integração de novos dados, e uso do API de compras governamentais, assim como tornar o PEP Scan acessível publicamente, ampliando seu impacto.

Referências

- [1] Guidance, FATF. Politically Exposed Persons (Recommendations 12 and 22). Documents-Financial Action Task Force (FATF)(fatf-gafi.org).(27.05.2022).
- [2] <https://www.gov.br/coaf/pt-br/aceso-a-informacao/Institucional/a-atividade-de-supervisao/regulacao/supervisao/normas-1/resolucao-no-29-de-7-de-dezembro-de-2017>
- [3] <https://portaldatransparencia.gov.br/download-de-dados/pep>
- [4] <https://dados.gov.br/dados/conjuntos-dados/auxilio-emergencial-ae>
- [5] <https://dados.gov.br/dados/conjuntos-dados/bolsa-familia>
- [6] <https://dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj>
- [7] Fayyad, U.; Platetsky-Shapiro, G.; Smyth, P. - Knowledge Discovery and Data Mining: Towards a Unifying Framework. In: Knowledge Discovery and Data Mining. 1996. p. 82-88.
- [8] D.A. Keim and H.-P. Kriegel, "Visualization Techniques for Mining Large Databases: A Comparison," IEEE Trans. Knowledge and Data Eng., vol. 8, no. 6, p. 923-936, Dec. 1996.
- [9] KEIM, Daniel A.; WARD, Matthew O. Visual data mining techniques. 2002.