

## **DESENVOLVIMENTO DE UM PROGRAMA JAVA PARA TRANSFORMAR DADOS BRUTOS EXPORTADOS DE BANCO DE DADOS EM HIPERTEXTOS**

**Pedro Monteiro da Silva Eleutério 1\***

<sup>1</sup> Polícia Federal – Setor Técnico-Científico da Superintendência Regional em Mato Grosso do Sul, Campo Grande, MS  
\*e-mail: [pedro.pmse@pf.gov.br](mailto:pedro.pmse@pf.gov.br)

### **RESUMO**

Este trabalho ilustra o desenvolvimento de um programa computacional em Java, realizado durante a elaboração de um Laudo Pericial em Computação Forense, para o tratamento e a organização dos dados brutos copiados durante o cumprimento de um mandado de busca e apreensão (cerca de 6,5 milhões de registros em 3,5 TB), transformando-os em mais de 300 mil hipertextos de processos judiciais com cerca de 1,5 milhões de arquivos de peças processuais linkadas, permitindo a análise, a identificação e as buscas por palavras-chaves pela equipe de investigação.

**Palavras-chave:** Computação Forense, Banco de Dados, Programação Java.

### **Introdução**

Durante uma operação de repercussão no Estado de Mato Grosso do Sul, que investigou possíveis fraudes em processos judiciais, foi realizada a cópia dos dados de todos os processos judiciais existentes em um órgão público, incluindo suas peças. Entretanto, tal ação apenas exportou as tabelas do banco de dados em grandes arquivos textuais brutos (CSVs) e copiou os arquivos referentes às peças processuais (PDFs, Imagens, DOCs, XLSs). Portanto, ao final da operação, os dados brutos foram copiados, porém não formatados, nem organizados, impossibilitando que a equipe de investigação conseguisse utilizá-los, saber quais peças pertenciam a quais processos ou realizar pesquisas por palavras-chaves.

### **Objetivos**

Este trabalho tem por objetivo mostrar como um programa Java foi desenvolvido durante a realização de um Laudo Pericial, transformando dados brutos exportados de um banco de dados em hipertextos navegáveis e pesquisáveis.

### **Métodos**

Seis passos metodologicamente bem definidos foram executados e detalhados no Laudo Pericial:

- (i) Validação dos dados: verificar a integridade dos dados copiados, com uso de funções hash;
- (ii) Tratamento dos dados: identificar os campos de cada arquivo CSV, tratando problemas para a correta importação ao programa, como a presença de caracteres inválidos, vírgulas, agrupamento de campos necessários e organização dos arquivos;
- (iii) Desenvolvimento do programa: desenvolvido em Java, com a modelagem das classes em UML, possui cerca de 1.500 linhas de código-fonte;
- (iv) Execução do programa: o programa foi executado, os dados originais foram importados e transformados em uma nova estrutura organizada de pastas, hipertextos navegáveis e suas peças;
- (v) Validação: garantir que os dados contidos nos hipertextos gerados correspondem fielmente aos dados contidos nos arquivos textuais CSV originais;
- (vi) Processamento no IPED: a nova estrutura organizada de pastas e arquivos (HTML/peças) foi processada e indexada no programa IPED.

### **Resultados e Discussão**

Como resultado, os cerca de 6,5 milhões de registros foram transformados em 337.205 hipertextos de processos, 1.561.032 arquivos de peças e 5.183 hipertextos de navegação. Todos os dados foram validados a partir de amostragens realizadas durante a fase de validação.

### **Conclusão**

A nova estrutura gerada pelo programa Java desenvolvido no escopo do Laudo Pericial possui os dados devidamente tratados e validados, permitindo o devido acesso, navegação, buscas por palavras-chaves e a correta análise pela equipe investigativa. Ademais, a metodologia de desenvolvimento e validação apresentada neste trabalho pode ser utilizada e replicada em casos semelhantes.

Realização