



A UNIVERSIDADE DE BRASÍLIA

APRESENTA UMA INOVAÇÃO

SISTEMA DE COLETA E MINERAÇÃO DE TEXTOS DO DIÁRIO OFICIAL DA UNIÃO: UMA FERRAMENTA PARA A GESTÃO PÚBLICA

BR 51 2018 001352 0

Grupo: Gestão e gerenciamento (recursos humanos)

Subgrupo: Tecnologia do trabalho

MINERAÇÃO DE DADOS NO CENÁRIO ATUAL

A mineração de dados é uma técnica que consiste em extrair, analisar e transformar grandes volumes de dados em informações úteis para a tomada de decisões. Essa técnica pode ser aplicada em diversos domínios, incluindo o serviço público, onde pode auxiliar em práticas burocráticas que envolvem a gestão de recursos, a prestação de serviços, o controle de qualidade, a fiscalização, a transparência e a participação cidadã. No entanto, a mineração de dados no serviço público também apresenta diversos desafios, tais como: a integração e a padronização de dados provenientes de diferentes fontes e sistemas; a garantia da privacidade e da segurança dos dados pessoais dos cidadãos; a escolha e a aplicação de métodos e ferramentas adequados para cada problema e contexto; a interpretação e a comunicação dos resultados obtidos; e a avaliação do impacto e do retorno social dos projetos de mineração de dados. Esses desafios exigem uma abordagem multidisciplinar e colaborativa, que envolva profissionais de diferentes áreas do conhecimento e setores da sociedade.

Sob essa perspectiva, pesquisadores da Universidade de Brasília (UnB), desenvolveram um sistema que facilita a coleta e a mineração de textos do Diário Oficial da União (D.O.U.), uma ferramenta que publica os atos oficiais do governo brasileiro. O sistema, intitulado “Extrator de Dados”, utiliza expressões regulares para extrair e identificar os dados dos atos normativos, como leis, decretos, portarias e despachos, e preenche automaticamente os metadados específicos para cada tipo de documento. Além disso, o sistema se integra com o software DSpace, que permite agrupar, organizar e classificar as informações extraídas, facilitando a pesquisa e o acesso aos dados. O sistema é focado nas seções 1 e 2 do D.O.U., que tratam respectivamente de atos de interesse geral



A UNIVERSIDADE DE BRASÍLIA APRESENTA UMA INOVAÇÃO

e de interesse dos servidores da Administração Pública Federal. O usuário pode enviar os dados extraídos diretamente ao DSpace, com a opção de editar antes do envio, ou mudar alguns dos parâmetros de extração e classificação, para adaptar o sistema à realidade dos documentos.

VANTAGENS

- **Acessibilidade:** Facilita o acesso aos atos oficiais do governo brasileiro através da coleta e análise automatizada de documentos.
- **Precisão:** Utiliza expressões regulares para extrair e identificar dados dos documentos normativos, como leis, decretos, portarias e despachos;
- **Organização:** Integra-se com o software DSpace, garantindo a agregação, organização e classificação das informações extraídas;
- **Controle:** Oferece a opção de editar os dados extraídos antes de enviá-los ao Dspace;
- **Flexibilidade:** Permite a configuração de parâmetros de extração e classificação, para adaptar o sistema à realidade dos documentos.

Agenda 2030 da ONU:



Gostou dessa tecnologia?

Para obter mais informações entre em contato com a Agência de Comercialização de Tecnologias (ACT) da Coordenação de Inovação e Transferência de Tecnologia (CITT), por meio dos contatos a seguir:

E-mail: act@cdt.unb.br

Telefone: (61) 3107-4116