**Smart Monitor: UmA FERRAMENTA para a Integração de Dados Abertos de Gestão Ambiental**

ANTÔNIO RICART JACINTO DE OLIVEIRA MEDEIROS (IFPB, Campus Cajazeiras), RUAN MIGUEL COELHO CARDOSO (IFPB, Campus Cajazeiras), FABIO GOMES DE ANDRADE (IFPB, Campus Cajazeiras)

**E-mails:** ricart.medeiros@academico.ifpb.edu.br, ruan.miguel@academico.ifpb.edu.br, fabio@ifpb.edu.br.

**Área de conhecimento:(Tabela CNPq)**: 1.03.03.03-0 Bancos de Dados.

**Palavras-Chave**: crowdsourcing; dados abertos governamentais; monitoramento ambiental; sustentabilidade

1. **Introdução**

A sustentabilidade é uma das grandes preocupações da atualidade, o que tem estimulado o desenvolvimento de vários projetos que permitam um monitoramento mais efetivo e, consequentemente, uma melhor gestão do meio ambiente. Essas ferramentas abordam diversos problemas, tais como o controle de poluição (OUARDIGHI et al., 2020; YE et al. 2020), a detecção de queimadas (GREBNEV et al., 2020; HUANG & DU, 2020), a previsão e monitoramento de enchentes (KHALAF et al., 2020; WANG et al., 2020), entre outros. Hoje, o desenvolvimento de sistemas de monitoramento ambiental é facilitado devido à grande quantidade de dados disponíveis. Um fator que contribuiu para o aumento do número de dados foi o avanço tecnológico, que hoje permite o armazenamento e o processamento de grandes quantidades de dados a um custo relativamente baixo. Além disso, nas últimas décadas, muitos governos do mundo inteiro, de diversos níveis (federal, estadual, regional e municipal), têm adotado a política de disponibilizar, de forma livre e gratuita, os conjuntos de dados que são produzidos por suas agências.

Uma importante fonte de informação disponibilizada pelos governos são as infraestruturas de dados espaciais (IDE). Elas foram desenvolvidas como uma solução para facilitar a localização e o compartilhamento dos dados espaciais produzidos por um conjunto de organizações, como forma de incentivar o reuso desses dados e evitar os custos na produção de novos conjuntos de dados. Para facilitar a localização dos dados que se encontram disponíveis, as infraestruturas de dados espaciais oferecem um serviço de catálogo. As agências produtoras de dados espaciais usam esse serviço para anunciar os seus conjuntos de dados. Por outro lado, os clientes podem usar o serviço de catálogo para localizar os dados nos quais estão interessados. Desde a sua proposição, várias IDEs já foram desenvolvidas no mundo inteiro. No Brasil, foi criada a Infraestrutura Nacional de Dados Espaciais[[1]](#footnote-1) (INDE), a partir da qual é possível obter dados espaciais de diversas organizações, como o Instituto Brasileiro do Meio Ambiente (IBAMA), a Agência Nacional de Águas (ANA), o Instituto Brasileiro de Geografia e Estatística (IBGE), entre outras.

Além do desenvolvimento das IDEs, o movimento para a divulgação de dados abertos governamentais também tem ganhado força no mundo inteiro. Esse movimento começou em 2009, com o então presidente dos Estados Unidos, Barack Obama (THE WHITE HOUSE, 2009), e defende a disponibilização dos dados produzidos por organizações públicas, permitindo que esses dados possam ser acessados por qualquer cidadão. Para facilitar a divulgação e o acesso a dados governamentais, vários governos têm criado portais de dados abertos para a disponibilização dos seus conjuntos de dados. Entretanto, diferentemente das IDEs, os dados desses portais não são essencialmente espaciais, embora alguns conjuntos de dados possuam, em seu conteúdo, referências a localizações geográficas. Desde o início do movimento dos dados abertos governamentais (OGD, do inglês *Open Government Data*), vários portais de dados abertos têm sido desenvolvidos no mundo inteiro. No Brasil, o portal de dados abertos do governo federal[[2]](#footnote-2) atualmente dá acesso a mais de dez mil conjuntos de dados sobre os mais variados domínios de aplicação.

Apesar da grande quantidade de dados publicados, as agências públicas não são as únicas fontes de dados que podem ser utilizadas para o monitoramento ambiental. Outro tipo importante de fonte de dados são as aplicações de *crowdsourcing*, nas quais os dados do sistema são gerados pela população de forma colaborativa. Um importante subtipo dessas aplicações são as aplicações de informações geográficas voluntárias, que permitem que cidadãos atuem como “sensores” e forneçam informações acerca dos locais existentes em uma determinada região, como, por exemplo, um bairro ou um município. Os dados gerados por esse mapeamento voluntário são disponibilizados de forma gratuita e podem ser livremente usados para o desenvolvimento de novas aplicações de software. Dois exemplos importantes de aplicações que fornecem dados espaciais de *crowdsourcing* são o *Wikimapia* e o *OpenStreetMap* (OSM).

Além das fontes de dados supracitadas, existem outras fontes importantes que oferecem dados úteis para o monitoramento ambiental. Por exemplo, algumas organizações disponibilizam livremente os dados obtidos por meio de suas estações de sensoriamento remoto. Alguns exemplos de informações coletadas por meio dessa tecnologia são os dados sobre condições meteorológicas, incidência de raios, incidência de queimadas, entre outras. No Brasil, por exemplo, esses dados podem ser obtidos por meio do sítio de instituições especializadas como o Instituto Nacional de Meteorologia (INMET) ou o Instituto Nacional de Pesquisas Espaciais (INPE).

Apesar da enorme quantidade de dados disponíveis que podem ser livremente utilizados para o desenvolvimento de ferramentas de monitoramento e gestão ambiental, ainda existem barreiras importantes que limitam o potencial de utilização desses dados. Uma dessas barreiras está relacionada ao processo de localização dos dados. Hoje, as informações disponíveis para essa tarefa estão espalhadas em diversas fontes de dados. Por exemplo, dados espaciais referentes a temas como hidrografia, relevo, uso do solo, tipo de solo e vegetação podem ser encontrados a partir de infraestruturas de dados espaciais. Por outro lado, dados abertos governamentais referentes a hospitais, localização de empresas e o número de pessoas que exploram um determinado recurso natural podem ser localizados a partir de portais de dados abertos governamentais. Já os dados fornecidos pela população, como a identificação das construções e locais que se encontram próximos aos recursos gerenciados podem ser encontrados a partir de sistemas de *crowdsourcing*. Finalmente, dados de sensoriamento remoto podem ser obtidos a partir de sítios de organizações especializadas. Dessa forma, para encontrar os dados do seu interesse, atualmente um gestor ambiental precisa consultar várias fontes de dados e integrá-los por conta própria. Para resolver essa limitação, esse trabalho propõe uma ferramenta chadama *Smart Monitor*, que permite que todos esses dados sejam centralizados, integrados e visualizados em um único ambiente.

1. **Materiais e Métodos**

Na primeira fase do desenvolvimento da pesquisa foi criado o ambiente de integração. Nessa etapa, o objetivo era ter um ambiente inicial no qual fosse possível adicionar, nas etapas posteriroes, informações de diferentes fontes de dados. Esse ambiente foi implementado utilizando-se a ferramenta *OpenLayers*, que facilita a exibição e o gerenciamento de mapas dinâmicos em páginas web. Uma vez criado o ambiente de integração, foram desenvolvidos os módulos para a inclusão dos diferentes tipos de dados usados para o monitoramento ambiental.

O primeiro módulo implementado permite a inclusão de dados de *crowdsourcing*. Para isso, foram incluídos no ambiente de integração os dados do projeto *OpenStreetMap*. Essa fonte de dados foi escolhida porque oferece gratuitamente uma grande quantidade de dados relevantes para o monitoramento ambiental e permite que os seus dados sejam facilmente acessados por meio de uma API baseada no protocolo *REST.* Esse acesso dinâmico aos dados é importante porque permite que o sistema de monitoramento possa ter sempre as informações mais recentes a respeito da região que está sendo monitorada. O segundo módulo desenvolvido permite a adição de dados espaciais. A solução desenvolvida pemite a visualização de dados espaciais vetoriais oferecidos por diversas agências. Essas informações são acessadas por meio de serviços web ofertados nos formatos *Web Map Service* (WMS) e *Web Feature Service* (WFS).

Depois da inclusão dos dados espaciais, foi implementado um módulo que permite a adição de dados abertos governamentais. Esse módulo permite que o usuário adicione dados produzidos por organizações públicas de nível municipal, estadual ou federal. Os dados governamentais são incluídos no formato CSV, no qual os dados são organizados e descritos de forma tabular. Assim, para que a inclusão desses dados seja possível, o módulo desenvolvido realiza a geocodificação do arquivo adicionado pelo usuário. Durante esse processo, cada linha de dados contida no arquivo é associada a uma localização no globo terrestre. Para realizar essa tarefa, o módulo analisa primeiro se o arquivo contém informações geográficas explícitas como, por exemplo, coordenadas de latitute e longitude. Caso esse tipo de informação não seja encontrado, o processo de geocodificação é feito com base em referências a nomes de lugares encontradas dentro do arquivo, tais como endereços, nomes de cidades, nomes de estados, entre outros. Nesse caso, o módulo utiliza um banco de dados local, formado a partir dos dados do *OpenStreetMap*, para idenficar a localização de cada linha de dados contida no arquivo. Ao fim do processo de geocodificação, o módulo gera um arquivo no formato *GEOJSON*, que pode ser visualizado no ambiente de integração.

Por fim, o ultimo módulo desenvolvido permite a inserção de dados de sensoriamento remoto. A implementação desse módulo permite ao usuário a visualização de dados como temperatura, umidade do ar, direação do vento, entre outros. Os dados de sensoriamento são acessados por meio de requisições *REST* e são recuperados no formato *JSON*. Assim, o módulo desenvolvido realiza a geocodificação do arquivo obtido a partir das coordenadas de latitude e longitude de cada estação de monitoramento descrita no documento obtido como resposta da requisição. Ao fim desse processo, também é gerado um arquivo no formato *GEOJSON*, que é encaminhado para a visualização no ambiente de integração.

1. **Resultados e Discussão**

A Figura 1 mostra a aparência atual da interface principal ferramenta desenvolvida, com destaque para o estado da Paraíba. Nela, a camada base, que mostra os contornos do estado, suas cidades, estradas e outras informações foram obtidas por meio do *OpenStreetMap*. Na figura, também é possível perceber que o estado é subdivido em várias regiões. Essas regiões representam os dados das suas bacias hidrográficas, que foram obtidos a partir de dados espaciais disponibilizados pela AESA. Finalmente, os pontos do mapa representam as estações de monitoramento do Instituto Nacional de Meteorologia (INMET) e descrevem os dados de sensoriamento remoto. Ademais, o usuário pode, a partir do menu exibido no lado esquerdo da figura, escolher quais camadas devem ser apresentadas. Para cada camada, ele também pode clicar no mapa para visualizar informações mais detalhadas de qualquer objeto que está sendo apresentado.



Figura 1: Interface principal da ferramenta

1. **Considerações Finais**

Os resultados obtidos com o desenvolvimento da pesquisa são muito animadores, uma vez que foi possível integrar, em um único ambiente de monitoramento, dados ambientais hetoerogêneos oferecidos por diferentes fontes de informação. Entretanto, ainda existem muitas questões desafiadoras que precisam ser resolvidas. Uma questão importante consiste na inclusão de dados de mídias sociais. Outra questão consiste em realizar uma integração mais profunda entre os conjuntos de dados disponiveis. Isso acontece porque o trabalho atual permite apenas a integração visual desses dados. Entretanto, para que seja possível o desenvolvimento de ferramentas mais sofisticadas, que permitam, por exemplo, a análise automatizada dos dados envolvidos, a geração de modelos preditivos e a simulação de eventos, é necessário que os dados sejam integrados em um nível mais baixo. Uma possível solução para essa questão consiste na utilzação de padrões como a linguagem RDF e o conceito de *Linked Open Data* para descrever as conexões entre os dados.

**Agradecimentos**

Os autores gostariam de agradecer ao IFPB pelo auxílio financeiro recebido para o desenvolvimento da pesquisa.

**Referências**

GREBNEV, Y. et al. The practice of connectionist model for predicting forest fires in the Arctic zones of the Krasnoyarsk Territory. Int. J. Syst. Assur. Eng. Manag. v. 11, n. 1s, p. 1–9, 2020.

HUANG, X.; DU, L. Fire Detection and Recognition Optimization Based on Virtual Reality Video Image. IEEE Access v. 8, p. 77951–77961, 2020.

KHALAF, M. et al. IoT-Enabled Flood Severity Prediction via Ensemble Machine Learning Models. IEEE Access v. 8, p. 70375-70386, 2020.

OUARDIGHI, F. et al. Transboundary pollution control and environmental absorption efficiency management. Annals OR v. 287, n. 2, p. 653–68, 2020.

THE WHITE HOUSE. Memorandum on Transparency and Open Government. 2009. Disponível em: <http://www.presidency.ucsb.edu/ws/index.php?pid=85677>.

YE, F. et al. An environmental pollution management method based on extended belief rule base and data envelopment analysis under interval uncertainty. Comput. Ind. Eng. v. 144, p. 106454 , 2020.

WANG, R. et al. Tracking Flooding Phase Transitions and Establishing a Passive Hotline With AI-Enabled Social Media Data. IEEE Access v. 8, p. 103395–103404 , 2020.

1. https://inde.gov.br/ [↑](#footnote-ref-1)
2. http://dados.gov.br [↑](#footnote-ref-2)