



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Instituto de Ciência e Tecnologia
Câmpus de Sorocaba

Viviane Yuki Hayachi

UTILIZAÇÃO DO BIG DATA PARA QUESTÕES AMBIENTAIS:

UMA ANÁLISE DE VIABILIDADE

Viviane Yuki Hayachi

UTILIZAÇÃO DO BIG DATA PARA QUESTÕES AMBIENTAIS:
UMA ANÁLISE DE VIABILIDADE

Trabalho de Conclusão de Curso apresentado ao Instituto de Ciência e Tecnologia de Sorocaba, Universidade Estadual Paulista (UNESP), como parte dos requisitos para obtenção do grau de Bacharel em Engenharia Ambiental.

Orientador: Gerson Araújo de Medeiros

Sorocaba
2023

H413u

Hayachi, Viviane Yuki

Utilização do big data para questões ambientais: uma análise de viabilidade / Viviane Yuki Hayachi. -- Sorocaba, 2023

31 p.

Trabalho de conclusão de curso (Bacharelado - Engenharia Ambiental)
- Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia, Sorocaba

Orientador: Gerson Araújo de Medeiros

1. Big Data. 2. Engenharia Ambiental. 3. Tecnologia Ambiental. I.

Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de
Ciência e Tecnologia, Sorocaba. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Viviane Yuki Hayachi

**UTILIZAÇÃO DO BIG DATA PARA QUESTÕES AMBIENTAIS:
UMA ANÁLISE DE VIABILIDADE**

Sorocaba, 27 de Junho de 2023

Prof. Dr. Gerson Araújo de Medeiros

Orientador

Trabalho aprovado por meio de parecer, homologado pelo Conselho de Curso em reunião de 05 de Julho de 2023.

Sorocaba
2023

RESUMO

A presente pesquisa teve por objetivo explorar as aplicações e os benefícios do *Big Data* no contexto ambiental, bem como analisar os desafios e as limitações associadas à sua utilização. Foram exploradas aplicações do *Big Data* no monitoramento de mudanças ambientais, previsão de desastres naturais e gerenciamento de recursos naturais. Identificam-se os benefícios potenciais, como *insights* valiosos, melhorias na tomada de decisões e promoção de práticas sustentáveis. No entanto, também foram levantadas limitações e preocupações éticas, como a qualidade e veracidade dos dados, a privacidade e a conformidade com as leis de proteção de dados. Recomenda-se que pesquisadores e profissionais enfrentem esses desafios por meio de abordagens interdisciplinares e colaborativas, investindo em infraestrutura adequada e garantindo conformidade com as leis e regulamentações de proteção de dados. Além disso, foram sugeridos o desenvolvimento de padrões éticos, diretrizes claras e parcerias público-privadas para impulsionar o uso responsável do *Big Data* no monitoramento ambiental.

Palavras-chave: desafios ambientais; monitoramento ambiental; sustentabilidade; proteção de dados.

ABSTRACT

This research aimed to explore the applications and benefits of *Big Data* in the environmental context, as well as to analyze the challenges and limitations associated with its use. Applications of *Big Data* in monitoring environmental changes, predicting natural disasters, and managing natural resources are explored. The potential benefits, such as valuable *insights*, improved decision-making, and promotion of sustainable practices, are identified. However, limitations and ethical concerns regarding data quality and veracity, privacy, and compliance with data protection laws were raised. Researchers and professionals are recommended to address these challenges through interdisciplinary and collaborative approaches, investment in adequate infrastructure, and compliance with data protection laws and regulations. Furthermore, the development of ethical standards, clear guidelines, and public-private partnerships are suggested to drive responsible use of *Big Data* in environmental monitoring.

Keywords: environmental challenges; environmental monitoring; sustainability; data protection.

SUMÁRIO

1	INTRODUÇÃO	6
2	OBJETIVO	8
3	MATERIAIS E MÉTODOS	9
4	RESULTADO DA REVISÃO DE LITERATURA	10
4.1	Papel do Big Data no meio ambiente	10
4.2	Os benefícios do uso de dados;	16
4.3	Os desafios do Big Data nas questões ambientais	19
5	CONSIDERAÇÕES FINAIS	24
	REFERÊNCIAS	26

1 INTRODUÇÃO

Na era tecnologicamente avançada de hoje, o crescimento exponencial da quantidade de dados gerados em diversas áreas tem despertado expressivo interesse e impactado a sociedade de maneira significativa. Esse fenômeno é popularmente conhecido como *Big Data*, que se refere ao grande volume de dados estruturados e não estruturados, caracterizados pela velocidade de geração, variedade de formatos e valor potencial.

De acordo com Gandomi e Haider (2015, p. 138), “Big Data é o termo utilizado para descrever o imenso volume de dados – estruturados, semiestruturados e não estruturados – que têm potencial de serem explorados para obtenção de informações valiosas”. Esse volume massivo de dados é gerado por diferentes fontes, como sensores, dispositivos móveis, redes sociais, transações online e sistemas de informação, criando uma vasta gama de oportunidades e desafios em diversos campos de conhecimento.

No contexto acadêmico, o tema do *Big Data* tem ganhado crescente interesse, pois sua aplicação abrange diversas disciplinas, desde a ciência da computação e engenharia de dados até sociologia e ciências da informação. A análise e interpretação desses dados em larga escala tem permitido *insights* valiosos e descobertas significativas em vários campos de estudo. A abordagem tradicional de processamento de dados muitas vezes se mostra inadequada para lidar com a grandeza e complexidade do *Big Data*. Portanto, é necessário o desenvolvimento de técnicas avançadas de processamento, análise e visualização de dados para extrair informações úteis e significativas.

Atualmente, a análise e o uso estratégico do *Big Data* têm sido aplicados em diversos setores, como saúde, finanças, marketing e governança. No entanto, um campo que tem ganhado cada vez mais relevância é seu uso no domínio ambiental. Segundo Tolle *et al.* (2009, p. 2), “o Big Data tem um papel transformador no enfrentamento dos desafios ambientais e no avanço da sustentabilidade”.

Desde revolucionar indústrias até moldar políticas públicas, o poder da análise de dados e seu potencial para um impacto social imenso não pode ser subestimado. No entanto, para uma análise mais aprofundada do *Big Data* é crucial examinar criticamente suas implicações ambientais. O crescimento exponencial na geração e consumo de dados tem consequências profundas para o meio ambiente, incluindo o consumo de recursos, as demandas energéticas e possíveis perturbações ecológicas.

De acordo com Chen *et al.* (2014, p. 591), “o uso de Big Data na área ambiental oferece a possibilidade de melhorar o entendimento e a gestão dos recursos naturais, a previsão e a mitigação de desastres ambientais, bem como a promoção da sustentabilidade”. Essa capacidade de coletar, armazenar e analisar grandes volumes de dados ambientais possibilita uma compreensão mais abrangente e detalhada, gerando *insights* sobre padrões climáticos, impactos humanos, mudanças ecológicas e ecossistemas, auxiliando na tomada de decisões e na implementação de políticas e práticas ambientais mais eficientes e sustentáveis.

Para entender o impacto ambiental do *Big Data*, torna-se essencial reconhecer a escala expressiva dos dados gerados e armazenados globalmente. De acordo com o relatório da *International Data Corporation* (2020), prevê-se que o *datashpere* global atinja a impressionante marca de 175 *zettabytes* até 2025, exigindo um consumo de energia e suporte de infraestrutura substanciais. Segundo Masanet *et al.* (2020), esse crescimento rápido levanta preocupações sobre a pegada de carbono associada aos centros de dados, que são responsáveis por alimentar e resfriar a vasta rede de servidores necessários para armazenar e processar dados. A natureza intensiva em energia dos centros de dados contribui para as emissões de gases de efeito estufa, tornando-os um grande contribuinte para as mudanças climáticas.

Além disso, os processos de extração e fabricação necessários para produzir o hardware necessário para armazenar e processar *Big Data* têm consequências ambientais prejudiciais. Segundo um estudo de Mudd (2010), a mineração de terras raras frequentemente envolve a abertura de minas a céu aberto, o que causa o desmatamento e a remoção da vegetação nativa. Além disso, o processo de extração pode levar à degradação do solo, à erosão e à contaminação de corpos d'água próximos.

A contaminação da água é um dos impactos ambientais mais preocupantes da extração de terras raras. Os resíduos químicos gerados durante o processamento dos minerais podem infiltrar-se nos aquíferos e rios, afetando a qualidade da água e a vida aquática. Estudos, como o de Wang *et al.* (2016, p. 5490), destacam a presença de metais pesados, como chumbo e cádmio, em áreas próximas a minas de terras raras, evidenciando a poluição da água como consequência direta da extração desses minerais.

Outro estudo realizado por Xiang *et al.* (2016) ressalta que a mineração de terras raras contribui para a destruição de habitats naturais, resultando na perda de biodiversidade e na redução da fauna e flora locais. A remoção da vegetação e a alteração do solo podem afetar ecossistemas frágeis e ecologicamente importantes.

2 OBJETIVO

O objetivo deste trabalho foi explorar as aplicações e os benefícios do *Big Data* no contexto ambiental, bem como analisar os desafios e as limitações associadas à sua utilização. Para alcançar esse objetivo, foi realizada uma revisão da literatura sobre o tema, utilizando periódicos acadêmicos relevantes, relatórios e outras fontes.

3 MATERIAIS E MÉTODOS

A estrutura metodológica desse trabalho incluiu os seguintes capítulos: primeira seção englobou a introdução, a segunda seção discutiu o papel do *Big Data* em questões ambientais, suas aplicações no monitoramento de mudanças ambientais, previsão de desastres naturais e gestão de recursos naturais. A terceira seção analisou os benefícios da utilização desses dados, incluindo seu potencial para melhorar as políticas e práticas de sustentabilidade. A quarta seção discutiu os desafios associados à sua utilização em pesquisas e análises ambientais, incluindo preocupações éticas e de privacidade, bem como desafios financeiros e técnicos. Por fim, a conclusão resumiu os principais argumentos apresentados no trabalho e forneceu recomendações para pesquisas futuras e ações para enfrentar esses desafios.

4 RESULTADOS DA REVISÃO DE LITERATURA

4.1 O papel do *Big Data* no meio ambiente

O *Big Data* desempenha um papel significativo no meio ambiente, permitindo o monitoramento e a análise de dados ambientais e larga escala, o que contribui para a compreensão e gestão sustentável dos recursos naturais.

Segundo estudos, as análises de dados permitem coleta, processamento e análise de grandes volumes de dados, como sensores remotos (ZHANG *et al.*, 2018), estações meteorológicas (MADURO-ABREU *et al.*, 2020) e satélites (FERREIRA, 2015). Silva *et al.* (2020) reforça que o *Big Data* auxilia na previsão e rápida resposta a tudo que cerca o meio ambiente trazendo diversos benefícios, possibilitando uma visão abrangente, detalhada e extremamente tempestiva das mudanças ambientais (PAPADOPOULOS; BALTA, 2022), permitindo a formulação de políticas públicas (MOLINARO *et al.*, 2018), facilitando a gestão de recursos naturais e conservação de ecossistemas (CHEN *et al.*, 2019) e a implementação de práticas mais sustentáveis (PESTANA *et al.*, 2022).

A análise do *Big Data* no monitoramento ambiental envolve o processamento de dados utilizando técnicas de análise avançada, como aprendizado de máquina (PESTANA *et al.*, 2022). Essas técnicas permitem identificar padrões, tendências e anomalias nos dados coletados. Segundo Manyika *et al.* (2011), a análise de *Big Data* pode ajudar a identificar correlações entre diferentes variáveis ambientais, permitindo uma compreensão mais profunda das mudanças em curso.

De acordo com Kitchin (2014), a simulação baseada em *Big Data* pode ajudar a prever o comportamento futuro do meio ambiente, considerando diversos fatores e variáveis. Essa abordagem permite a realização de análises preditivas, auxiliando na antecipação de possíveis cenários e eventos ambientais, Zhang *et al.* (2020) propõem um *framework* que integra técnicas de *Big Data*, modelagem ambiental e aprendizado de máquina para realizar simulações e previsões sobre eventos de enchentes e deslizamentos de terra, enquanto Assis *et al.* (2018) apresentam um estudo de caso que utiliza técnicas de *Machine Learnig* e análise de dados em tempo real para prever a qualidade do ar em diferentes áreas geográficas, ambos demonstrando a eficácia do *Big Data* para fornecer informações valiosas para gestão ambiental e tomada de decisão.

Em relação à gestão de recursos naturais, a aplicação do *Big Data* desempenha um papel crucial. De acordo com Hilbert (2016, p. 137), “o uso de Big Data na gestão de recursos naturais

permite uma melhor compreensão nos padrões de uso, escassez e demanda, facilitando a implementação de estratégias sustentáveis de gestão de recursos”.

Como consideração geral, o uso de dados facilita a pesquisa científica e a colaboração na área ambiental. A análise de *Big Data* promove a integração e compartilhamento de informações entre diferentes instituições e pesquisadores, permitindo o avanço do conhecimento e a busca de soluções conjuntas para desafios ambientais complexos. Acesso a conjuntos de dados abrangentes e colaboração em larga escala impulsionam a inovação e o desenvolvimento de abordagens mais eficientes.

Como o *Big Data* pode ser utilizado para monitorar mudanças ambientais

Tratando-se do monitoramento de mudanças ambientais, a utilização dos dados tem se mostrado uma abordagem promissora para compreender e lidar com desafios ambientais enfrentados atualmente, trazendo diversos benefícios em sua aplicação.

Uma dessas vantagens é a capacidade de detecção precoce de eventos adversos, como desastres naturais. Segundo Abbasi *et al.* (2016), o uso de sensores e sistemas de monitoramento baseados em *Big Data* pode fornecer alertas antecipados sobre eventos climáticos externos, permitindo uma resposta rápida e eficaz.

Outra vantagem é a possibilidade de tomada de decisões baseadas em evidências. A análise de dados fornece informações detalhadas e atualizadas sobre o estado do meio ambiente, permitindo a formulação de políticas públicas e estratégias de conservação mais embasadas. Conforme destacado por Hilbert (2016), o uso de *Big Data* pode facilitar a gestão de recursos naturais, a conservação de ecossistemas e a implementação de práticas sustentáveis.

Na *U.S. Environmental Protection Agency* (Agência de proteção ambiental dos Estados Unidos), pode-se citar o monitoramento de qualidade de ar que vem sendo elaborado no escritório de pesquisa e desenvolvimento, tendo como um dos exemplos o enfoque em novas tecnologias que auxiliem o governo a avaliar a camada de ozônio e a conformidade com os padrões de qualidade de ar nacional em conjunto com um método de fusão de dados que combine o resultado dos monitoramentos para criar previsões (BREGGIN *et al.*, 2014, p. 6)

Um estudo conduzido por Díaz-Galiano *et al.* (2017) utilizou o *Big Data* para monitorar o desmatamento da Amazônia. Os pesquisadores combinaram dados de satélite, informações geoespaciais e dados socioeconômicos para identificar áreas com maior taxa de desmatamento

e entender os fatores associados a esse processo. O estudo demonstrou a eficácia do *Big Data* na detecção e monitoramento de mudanças no uso da terra em larga escala.

Outro projeto acadêmico notável é o trabalho de Chen *et al.* (2019), que utilizou técnicas de *Big Data* para monitorar a qualidade da água em rios e lagos. Os pesquisadores integraram dados de sensores ambientais, dados de monitoramento em tempo real e informações geográficas para analisar parâmetros como pH, oxigênio dissolvido e presença de poluentes. Isso permitiu uma compreensão mais detalhada dos padrões de qualidade da água e a identificação de áreas com problemas ambientais.

Além disso, o projeto de pesquisa de Ali *et al.* (2020) explorou o uso do *Big Data* no monitoramento de mudanças climáticas. Os pesquisadores coletaram e analisaram dados climáticos de várias fontes, como estações meteorológicas, satélites e sensores remotos. Com base nesses dados, eles foram capazes de identificar tendências climáticas, modelar cenários futuros e avaliar os impactos das mudanças climáticas em diferentes regiões.

Em suma, o *Big Data* desempenha um papel fundamental no monitoramento de mudanças ambientais. Através da coleta, processamento e análise de grandes volumes de dados, é possível obter *insights* valiosos sobre o estado do meio ambiente, identificar tendências e padrões, antecipar eventos adversos e embasar tomadas de decisões informadas.

Predição de desastres naturais e seu impacto no ambiente

A ocorrência de desastres naturais representa uma ameaça significativa para o meio ambiente e a segurança das comunidades em todo o mundo. Nesse contexto, o uso do *Big Data* tem se mostrado uma ferramenta promissora para a predição de desastres naturais e a avaliação de seu impacto ambiental, fornecendo informações valiosas para a mitigação e resposta eficaz a esses eventos.

Um estudo conduzido por Yin *et al.* (2018) propôs um modelo de predição de inundações baseado em *Big Data*, que utilizou dados de sensores de monitoramento hidrológico, dados meteorológicos, informações geoespaciais e históricos de inundações passadas. O modelo demonstrou uma precisão significativa na previsão de inundações, permitindo a implementação de medidas preventivas e a adoção de estratégias de resposta mais eficazes.

Outro projeto acadêmico relevante é o trabalho de Guo *et al.* (2020), que desenvolveu um sistema de previsão de deslizamentos de terra utilizando *Big Data*. O sistema integrou

dados geoespaciais, informações topográficas, dados climáticos e históricos de deslizamentos para identificar áreas propensas a deslizamentos e fornecer alertas precoces. Isso permitiu a tomada de decisões rápidas para evacuação de áreas de risco e a implementação de medidas de mitigação adequadas.

Além disso, o projeto de pesquisa de Kaviani-Mehr *et al.* (2019) explorou o uso de técnicas de *Big Data* e aprendizado de máquina para a predição de terremotos. Eles utilizaram dados sísmicos, dados geoespaciais, dados de sensoriamento remoto e informações geofísicas para desenvolver um modelo de previsão de terremotos em tempo real. Esse modelo mostrou resultados promissores na detecção precoce de atividades sísmicas e na identificação de áreas com maior probabilidade de ocorrência de terremotos. Tais projetos acadêmicos são apenas alguns exemplos que ilustram a aplicação do *Big Data* na predição de desastres naturais. Eles demonstram o potencial dessas abordagens em melhorar a capacidade de antecipação, resposta e mitigação dos impactos causados por eventos naturais adversos.

A possibilidade da coleta de dados em tempo real provenientes de diversas fontes que o *Big Data* fornece, como retratado anteriormente, é um dos recursos mais preciosos dessa relação. Esses dados incluem informações sobre padrões climáticos, atividade sísmica, níveis de água em rios e oceanos, entre outros parâmetros relevantes.

Sua utilização permite aplicação de técnicas avançadas de modelagem e simulação para prever desastres naturais. Conforme destacado por Zhang *et al.* (2018, p. 52), “A análise de dados em larga escala pode ser usada para construir modelos preditivos de desastres naturais, como furacões, terremotos e enchentes, permitindo a identificação de áreas de risco e a tomada de medidas preventivas”. Por exemplo, um estudo de Lin *et al.* (2017) analisou o uso do *Big Data* para a previsão de inundações na China, com foco na aplicação de modelos de análise de dados espaciais para identificar áreas de risco e desenvolver sistemas de alerta precoce.

Outro estudo interessante é o de Deng *et al.* (2018), que explorou o uso do *Big Data* na prevenção de deslizamentos de terra na Índia. Os pesquisadores utilizaram dados geoespaciais, sensoriamento remoto e informações socioeconômicas para mapear áreas suscetíveis a deslizamentos e desenvolver estratégias de mitigação.

Morales-González *et al.* (2019) examinaram o uso do *Big Data* na prevenção de incêndios florestais na Austrália, onde foram utilizados dados de sensores, satélites e redes sociais para monitorar condições climáticas, identificar áreas de risco e melhorar a resposta e mobilização de recursos.

Além da predição, o *Big Data* também desempenha um papel importante no entendimento do impacto dos desastres naturais no meio ambiente. Segundo Imran *et al.* (2015, p. 13), “A análise de Big Data pode ser usada para avaliar os danos causados por desastres naturais, como desmatamento, poluição da água, perda de biodiversidade e degradação de ecossistemas”. Isso permite uma melhor compreensão dos efeitos desses eventos e auxilia na implementação de estratégias de recuperação ambiental.

Pesquisadores do estudo de Gao *et al.* (2017) desenvolveram um projeto para avaliar os danos causados por terremotos usando o *Big Data*. Eles utilizaram dados sísmicos, imagens de satélite e informações geoespaciais para mapear e analisar os impactos dos terremotos, identificar áreas afetadas e fornecer informações valiosas para a tomada de decisões e respostas de emergência. Enquanto Lwin *et al.* (2020) avaliaram os danos causados por ciclones tropicais e Hwang (2019) de enchentes.

Portanto, a análise de dados na previsão de desastres naturais e no entendimento de seu impacto ambiental é uma abordagem promissora. A coleta e análise de grandes volumes de dados fornecem informações valiosas para melhorar a prevenção, mitigação e resposta a eventos adversos, contribuindo para a proteção do ambiente e a segurança das comunidades afetadas.

Impacto do uso de dados no gerenciamento de recursos naturais

A utilização dos dados tem um impacto significativo no gerenciamento de recursos naturais, proporcionando *insights* valiosos para a conservação, alocação e uso sustentável dos recursos. A análise de grandes volumes de dados coletados de fontes diversas permite uma compreensão mais profunda dos padrões e processos envolvidos nos ecossistemas naturais.

De acordo com Manyika *et al.* (2011, p.12) “a análise de Big Data tem o potencial de revolucionar a forma como gerenciamos os recursos naturais, fornecendo informações mais precisas e atualizadas para tomadas de decisões embasadas”. A identificação de padrões de consumo, monitoramento do estado de recursos naturais e avaliação do impacto das atividades humanas sobre o meio ambiente são exemplos dessas informações. No estudo de Sun *et al.* (2017), os pesquisadores realizaram um projeto para analisar *Big Data* e dados geoespaciais para o gerenciamento florestal. Eles utilizaram algoritmos de aprendizado de máquina para analisar dados de sensoriamento remoto, dados climáticos e informações sobre a vegetação para identificar padrões de desmatamento, prever riscos de incêndios florestais e auxiliar na tomada de decisões para a conservação florestal.

No contexto da gestão de água, o *Big Data* desempenha um papel fundamental. Segundo Capota *et al.* (2018, p. 1567), “A análise de dados pode melhorar a gestão dos recursos hídricos, permitindo a identificação de padrões de consumo, previsão de demanda e detecção de vazamentos e desperdícios”. Isso possibilita a implementação de estratégias de conservação e uso eficiente da água. Batista *et al.* (2018) utilizou algoritmos de análise de dados para identificar padrões de uso de água, detectar áreas de escassez e auxiliar na tomada de decisões para a gestão sustentável dos recursos hídricos

Existe também a aplicação no gerenciamento de florestas e biodiversidade. Conforme destacado por Nanni *et al.* (2017, p. 205305), “Dados em larga escala podem auxiliar na identificação de áreas de conservação prioritárias, monitoramento de espécies ameaçadas e avaliação do impacto das atividades humanas sobre os ecossistemas florestais”. Isso contribui para a proteção da biodiversidade e o desenvolvimento de estratégias de manejo sustentável. Anderson *et al.* (2018) realizaram um projeto para monitorar e avaliar o impacto de atividades humanas sobre a biodiversidade florestal utilizando dados de sensoriamento remoto e *Big Data*. Eles desenvolveram modelos estatísticos para analisar a influência de fatores como desmatamento, fragmentação florestal e mudanças climáticas na diversidade e distribuição de espécies, auxiliando na identificação de áreas prioritárias para conservação.

O uso do *Big Data* também tem implicações no setor de energia. De acordo com Hashem *et al.* (2015, p. 101), “Pode ajudar na otimização do uso de energia, identificação de padrões de consumo, monitoramento da eficiência energética e desenvolvimento de estratégias para a transição para fontes de energia renovável”. Isso possibilita uma gestão mais eficiente dos recursos energéticos e a redução do impacto ambiental, como utilizado no *The Emissions & Generation Resource Integrated Database* (eGRID), um inventário de dados federais sobre usinas de energia e empresas de energia coletados, abrangendo as características ambientais de quase toda eletricidade gerada nos Estados Unidos. Dentre as características estão a emissão atmosférica de poluentes, a taxa da emissão e a quantidade de energia elétrica gerada em uma usina elétrica que é transmitida e distribuída ao consumidor. Estas enormes quantidades de dados são agregadas por instalações, empresas e pelo Estado (BREGGIN *et al.*, 2014, p. 7).

Entende-se que o uso massivo de dados desempenha um papel importante no gerenciamento de recursos naturais, fornecendo informações valiosas para a conservação, alocação e uso sustentável dos recursos. A análise de grandes volumes de dados permite

identificar padrões, avaliar impactos e embasar tomadas de decisões informadas. Essa abordagem contribui para a proteção do meio ambiente e a promoção da sustentabilidade.

4.2 Os benefícios do uso de dados

Como aplicação do uso do Big Data pode-se citar o trabalho desenvolvido por Santos *et al.* (2020) na análise dos padrões de desmatamento da Amazônia brasileira. Eles demonstraram como o uso de dados em larga escala, como imagens de satélite e informações geoespaciais, aliado a algoritmos de aprendizado de máquina, pode fornecer *insights* valiosos sobre as causas e padrões de desmatamento, auxiliando na tomada de decisões para a conservação e monitoramento da floresta.

De acordo com Bell *et al.* (2017, p. 235), “A análise de Big Data possibilita a integração de dados de diferentes fontes e disciplinas, permitindo uma compreensão mais holística dos sistemas ambientais e das interações entre seus componentes”. Essa abordagem é crucial para uma compreensão mais completa dos fenômenos ambientais e a identificação de relações de causa e efeito.

Neves *et al.* (2022) utilizaram um modelo de inteligência artificial para estimar o desmatamento causado pela rede de transporte rodoviário do estado do Pará, apontando a importância da realização contínua de estudos e análises sobre diferentes cenários que possam interferir no fenômeno. Assim como ensaios comparativos e simulações semelhantes com outros algoritmos de aprendizado de máquina.

Souza *et al.* (2019) desenvolveram um projeto que utiliza o *Big Data* para monitorar e responder rapidamente a mudanças ambientais em áreas costeiras. Eles utilizaram uma combinação de dados de sensores remotos, informações meteorológicas, dados de monitoramento de marés e dados de sensoriamento de infravermelho para identificar eventos climáticos extremos, como tempestades e ressacas, que afetam áreas costeiras sensíveis. Essa abordagem permitiu uma detecção rápida de mudanças e uma resposta imediata por meio de medidas de proteção costeira, minimizando os impactos ambientais e protegendo os ecossistemas costeiros.

Além disso, o *Big Data* também favorece a descoberta de padrões e correlações ocultas nos dados ambientais. Segundo Ke *et al.* (2017, p. 100), “A análise de dados emprega técnicas avançadas, como mineração de dados e aprendizado de máquina, para identificar relações

complexas e não lineares nos dados ambientais, revelando insights valiosos e não evidentes à primeira vista”.

Essa capacidade de descoberta de padrões ocultos impulsiona a pesquisa científica e o desenvolvimento de soluções inovadoras para os desafios ambientais, como apresentado por Silva *et al.* (2020). Esses autores coletaram dados de sensores, informações meteorológicas, informações de tráfego e outras variáveis relevantes para identificar fatores que contribuem para a poluição do ar em diferentes regiões. Com base nas análises realizadas, eles desenvolveram soluções inovadoras para mitigar a poluição do ar, como estratégias de controle de emissões veiculares e melhorias na infraestrutura urbana.

O papel do *Big Data* no desenvolvimento de políticas e práticas sustentáveis é fundamental, permitindo uma abordagem baseada em evidências e dados concretos. De acordo com Hilbert e López (2011, p. 60), “O Big Data fornece informações quantitativas e qualitativas abrangentes, permitindo a análise de problemas complexos e a avaliação do impacto de diferentes medidas”. Esse embasamento é crucial para a formulação de políticas eficazes e implementação de práticas sustentáveis. Sousa *et al.* (2021) por meio da análise de dados, incluindo técnicas de aprendizado de máquina e modelagem preditiva, identificaram padrões e tendências, subsidiando a formulação de políticas eficazes e a tomada de decisões informadas para a formulação de políticas de sustentabilidade no Brasil.

Assim, o *Big Data* desempenha um papel fundamental no desenvolvimento de políticas e práticas sustentáveis, fornecendo informações embasadas em dados e promovendo a inovação e a participação pública. Sua utilização permite uma abordagem baseada em evidências, identificação de tendências, criação de soluções personalizadas e promoção da transparência. Esses aspectos são essenciais para promover a sustentabilidade e enfrentar os desafios ambientais de forma eficaz.

Outro aspecto notável da análise de *Big Data* é sua capacidade de ser aplicada a projetos bem-sucedidos que abordam questões ambientais. Ao aproveitar o poder de vastos conjuntos de dados, essas iniciativas demonstraram o potencial das tecnologias orientadas por dados na mitigação de desafios ambientais e na promoção de práticas sustentáveis. A seguir, discute-se exemplos notáveis de projetos que utilizaram com êxito *Big Data* para enfrentar questões ambientais urgentes, que vão desde a conservação da biodiversidade até a adaptação às mudanças climáticas. Esses projetos servem como modelos inspiradores para

empreendimentos futuros e destacam o impacto transformador que as abordagens baseadas em dados podem ter na sustentabilidade ambiental.

- Automação em estação de tratamento de esgoto

Pestana *et al.* (2022) utilizaram o *Big Data* e o *Machine Learning* num piloto de estação de tratamento de esgoto para a melhoria no controle de Oxigênio Dissolvido (OD), nos tanques de aeração, controle dos produtos químicos utilizados na etapa de desidratação de lodo, controle do manto de lodo nos decantadores, entre outros. Esse estudo visou a estabilidade nos processos e eficiência energética.

- Dashboards de *Big Data Analytics* nas Cidades Inteligentes

Moraes (2018) aplicou o *Big Data* em dashboards internos do Centro de Gestão Integrada de Goiânia, fornecendo um videomonitoramento para promover melhoramentos em termos de rapidez e efetividade nas decisões dos agentes municipais, otimizando recursos naturais, econômicos e humanos, aprimorando serviços público. Além disso, em última instância, propiciou a geração de maior qualidade de vida para os cidadãos.

- Projeto *Urban Observatory*

O *Urban Observatory* é um projeto desenvolvido pela *Microsoft Research* em colaboração com várias cidades ao redor do mundo. Ele utiliza o *Big Data* para monitorar e analisar dados ambientais em áreas urbanas, incluindo informações sobre qualidade do ar, níveis de ruído, consumo de energia e o uso de recursos naturais. O objetivo é obter *insights* para melhorar a sustentabilidade e a qualidade de vida nas cidades (MICROSOFT, 2020).

- Projeto *Global Forest Watch*

O *Global Forest Watch* é uma iniciativa liderada pelo *World Resources Institute* (WRI) que utiliza o *Big Data* para monitorar o desmatamento e a perda de cobertura florestal em todo o mundo. Ele combina dados de satélite, imagens de alta resolução e algoritmos de aprendizado de máquina para fornecer informações atualizadas e precisas sobre mudanças na cobertura florestal, auxiliando na conservação e gestão sustentável das florestas (WORLD RESOURCES INSTITUTE, 2021)

- Projeto *Smartfin*

O *Smartfin* é um projeto que utiliza pranchas de surfe equipadas com sensores para coletar dados ambientais nos oceanos. Os sensores coletam informações sobre temperatura,

salinidade, pH da água, níveis de oxigênio, entre outros parâmetros. Esses dados são enviados para uma plataforma online, onde cientistas e pesquisadores podem acessá-los e utilizá-los para monitorar a saúde dos ecossistemas marinhos e identificar áreas de conservação prioritárias (SMARTFIN PROJECT, 2021)

- *Global Fishing Watch*

O *Global Fishing Watch* é um projeto inovador que utiliza dados de satélite e algoritmos de aprendizado de máquina para monitorar e combater atividades de pesca ilegal. Ao analisar os padrões de movimento das embarcações e o comportamento de pesca, essa iniciativa possibilita a identificação de práticas de pesca suspeitas ou ilegais, ajudando a proteger os ecossistemas marinhos e promover uma gestão sustentável da pesca. (GLOBAL FISHING WATCH, 2012)

- *Earth Observing System Data and Information System (EOSDIS)*

A EOSDIS, um projeto da NASA, coleta e gerencia uma extensa coleção de dados de observação da Terra. Isso inclui imagens de satélite, modelos climáticos e dados atmosféricos, que são disponibilizados gratuitamente para pesquisadores e formuladores de políticas. Ao utilizar esse conjunto abrangente de dados, os cientistas podem monitorar as mudanças climáticas, rastrear o desmatamento e estudar outros fenômenos ambientais. Isso permite uma tomada de decisão baseada em evidências e uma gestão ambiental proativa. (NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, 2020)

- *AI for Earth*

O programa AI for Earth da Microsoft concentra-se em utilizar inteligência artificial e análise de *Big Data* para abordar desafios ambientais. A iniciativa apoia projetos em várias áreas, como conservação da biodiversidade, mudanças climáticas e gestão da água. Ao fornecer bolsas, acesso a dados e ferramentas de IA, a Microsoft capacita pesquisadores e organizações a desenvolver soluções inovadoras e contribuir significativamente para a sustentabilidade ambiental. (MICROSOFT, 2017)

4.3 Os desafios do *Big Data* nas questões ambientais

A utilização do *Big Data* em pesquisas e análises ambientais apresenta várias vantagens, porém, também enfrenta algumas limitações que devem ser consideradas, tanto na questão do próprio dado em si, quanto aspectos éticos e de segurança na utilização desses dados, visando garantir a privacidade e a proteção das informações coletadas.

Dito isso, a problemática se inicia desde a coleta e integração de conjuntos de dados diversos de múltiplas fontes. A obtenção de dados relevantes de satélites, sensores, mídias sociais e outras plataformas requer infraestrutura robusta de coleta de dados e protocolos padronizados para garantir a compatibilidade e interoperabilidade dos dados. "A coleta e integração de conjuntos de dados de múltiplas fontes representam um desafio significativo na análise de Big Data. A diversidade de formatos, estruturas e padrões de dados torna a tarefa de coletar, organizar e integrar essas informações complexa e demorada" (SOUZA *et al.*, 2020, p. 140)

Em seguida, entra a qualidade e a representatividade dos dados, uma limitação comum quando se fala da utilização massiva dos dados. Conforme apontado por Wang *et al.* (2016b, p. 372), "A confiabilidade e a precisão dos dados podem ser afetadas pela falta de padronização, pela heterogeneidade dos métodos de coleta e pela falta de controle de qualidade". Tais limitações podem comprometer a validade e a confiabilidade das análises realizadas com base nos dados coletados.

A qualidade e precisão do *Big Data* utilizados na análise ambiental são fundamentais. Garantir a qualidade dos dados requer processos rigorosos de validação, considerando fatores como fontes de dados, metodologias e possíveis vieses. "a qualidade dos dados e a heterogeneidade entre as fontes podem introduzir incertezas e erros que afetam diretamente a confiabilidade e a precisão das análises realizadas." (SOUZA *et al.*, 2020, p. 57)

Consequente aos avanços dos dados, Silva *et al.* (2019) argumentam que embora o acesso a grandes volumes de dados seja fundamental para a realização de análises abrangentes e descobertas significativas, é crucial garantir a proteção da privacidade das pessoas envolvidas nos conjuntos de dados. Além disso, deve-se estabelecer diretrizes e recomendações para garantir um equilíbrio adequado entre a acessibilidade dos dados e a proteção da privacidade individual, de modo a maximizar os benefícios do *Big Data*.

As implicações éticas do uso do *Big Data* em pesquisa ambiental e tomada de decisão não podem ser negligenciadas. Questões relacionadas ao consentimento informado, propriedade dos dados e o potencial de consequências não intencionais devem ser abordadas. Estruturas de governança transparentes e diretrizes éticas são necessárias para garantir o uso responsável e responsabilização do *Big Data* em contextos ambientais (SMITH *et al.*, 2018).

Extrair *insights* significativos do *Big Data* e traduzi-los em informações acionáveis para as partes interessadas apresenta desafios na interpretação e comunicação dos dados. Analisar

conjuntos de dados complexos requer expertise em análise de dados e conhecimento específico do domínio. Comunicar efetivamente os resultados para formuladores de políticas, comunidades e público em geral de maneira clara e compreensível é crucial para uma tomada de decisão informada. Conforme ressaltado por Lazer *et al.* (2014, p. 1203), "O viés nos dados pode ser introduzido por meio de seleção não aleatória, erros de medição e outras limitações inerentes aos dados coletados". É fundamental estar ciente desses possíveis vieses e adotar abordagens metodológicas adequadas para mitigá-los.

O uso de *Big Data* no monitoramento ambiental pode criar desafios relacionados à discriminação e desigualdade. Segundo Crawford e Schultz (2014, p. 97), "*As análises de Big Data podem reproduzir e ampliar desigualdades sociais e econômicas existentes, pois os dados podem refletir e agravar vieses e estereótipos*". É essencial garantir que o uso de *Big Data* seja realizado de maneira justa e equitativa, evitando a perpetuação de desigualdades e preconceitos.

Ao enfrentar esses desafios de frente, pode-se desbloquear o imenso potencial do *Big Data* para informar tomadas de decisão baseadas em evidências, impulsionar a inovação e promover práticas sustentáveis. Segundo Johnson *et al.* (2019), esforços colaborativos entre pesquisadores, formuladores de políticas e organizações são necessários para desenvolver estruturas robustas de governança de dados, garantir qualidade e privacidade dos dados e promover o uso responsável do *Big Data* em prol da sustentabilidade ambiental.

Privacidade e Ética no mundo dos dados

A coleta de dados em sua amplitude necessita que a garantia de consentimento seja obtida de forma adequada e transparente, especialmente quando se trata de dados pessoais, e dados para utilização ambiental não é diferente. Como mencionado por Zwitter (2014, p. 6), "O consentimento informado é um princípio ético-chave que deve ser respeitado para garantir que os indivíduos tenham conhecimento e controle sobre o uso de seus dados". É necessário adotar práticas de obtenção de consentimento claras e acessíveis.

Existem inúmeras leis adotadas sobre privacidade e ética no contexto do *Big Data*. Tais leis têm um papel crucial na proteção dos direitos individuais e na garantia de práticas responsáveis no uso desses dados. Cho *et al.* (2016) argumentam que vários países e regiões tem desenvolvido leis e regulamentos específicos para abordar as preocupações relacionadas, como por exemplo na União Europeia que segue o Regulamento Geral de Proteção de Dados (GPRD), implementada em 2018.

A GPRD é uma das legislações mais abrangentes e influentes no campo da proteção de dados pessoais, estabelecendo diretrizes claras sobre a coleta, o processamento e o armazenamento de dados pessoais na União Europeia. Conferindo aos indivíduos maior controle sobre seus dados pessoais, exigindo consentimento informado e garantindo direito à portabilidade e ao esquecimento. (EUROPEAN PARLIAMENT; COUNCIL OF THE EUROPEAN UNION, 2016)

No Brasil, a Lei Geral de Proteção de Dados Pessoais (LGPD) é uma legislação que tem como objetivo proteger a privacidade e os direitos dos indivíduos em relação ao tratamento de seus dados pessoais. Aprovada em 2018 e em vigor desde setembro de 2020, a LGPD estabelece diretrizes e princípios que devem ser seguidos por organizações que coletam, armazenam, processam e compartilham dados pessoais. (BRASIL, 2018)

Essa lei estabelece diretrizes específicas para o tratamento de dados pessoais no contexto geral, não excludente do ambiental. Isso inclui a necessidade de obtenção de consentimento informado dos indivíduos para a coleta e o processamento de seus dados, a adoção de medidas de segurança adequadas para proteção dos dados, a garantia de acesso e correção das informações pessoais, entre outros direitos fundamentais dos titulares dos dados. (BRASIL, 2018)

Desafios na implantação do *Big Data* no meio ambiente

A implementação do *Big Data* em assuntos ambientais apresenta desafios financeiros e técnicos significativos que precisam ser abordados para garantir o sucesso dessas iniciativas. Esses desafios podem incluir a infraestrutura necessária para lidar com grandes volumes de dados, a expertise técnica para coletar, armazenar e processar essas informações, além dos custos envolvidos nesse processo.

De acordo com um estudo realizado por Li *et al.* (2020), a implementação do *Big Data* em assuntos ambientais enfrenta desafios técnicos relacionados à heterogeneidade dos dados ambientais, à falta de padronização e à disponibilidade limitada de especialistas em *Big Data* com conhecimento em questões ambientais. Esses desafios podem dificultar a coleta, a integração e a análise efetiva dos dados ambientais.

Outro estudo conduzido por Song *et al.* (2019) destaca que a implementação do *Big Data* em assuntos ambientais requer uma abordagem multidisciplinar, envolvendo especialistas em meio ambiente, cientistas de dados, engenheiros e profissionais de TI. A colaboração entre

essas diferentes áreas é essencial para superar os desafios técnicos e garantir a eficácia das aplicações do *Big Data* no campo ambiental.

5 Considerações Finais

A análise dos desafios e oportunidades relacionados à utilização do *Big Data* em questões ambientais nos leva a reconhecer sua importância crucial para enfrentar os desafios ambientais contemporâneos. Ao longo deste trabalho, explorou-se como o *Big Data* pode ser usado para monitorar mudanças ambientais, prever desastres naturais e gerenciar recursos naturais. Identificou-se os benefícios potenciais, como a capacidade de obter *insights* valiosos, melhorar a tomada de decisões e promover práticas sustentáveis. No entanto, também se levantou as limitações e preocupações éticas, como a necessidade de lidar com a qualidade e veracidade dos dados, proteger a privacidade e garantir a conformidade com as leis de proteção de dados.

Diante disso, é essencial que pesquisadores e profissionais direcionem seus esforços para enfrentar esses desafios e explorar todo o potencial do *Big Data* no contexto ambiental. Para isso, recomenda-se a implementação de estratégias interdisciplinares e colaborativas, reunindo especialistas em meio ambiente, cientistas de dados, engenheiros e profissionais de TI. Essa colaboração permitirá a integração eficiente de diferentes fontes de dados, a análise avançada e a compreensão abrangente dos desafios ambientais.

Além disso, é fundamental investir em infraestrutura adequada, tanto em termos de tecnologia quanto de recursos financeiros. Isso inclui a adoção de sistemas robustos de armazenamento e processamento de dados, o desenvolvimento de algoritmos avançados e a capacitação de profissionais com conhecimentos específicos em *Big Data* e questões ambientais.

Outro aspecto importante é a necessidade de garantir a conformidade com as leis e regulamentações de proteção de dados. A implementação de medidas de segurança e privacidade, como anonimização de dados, controle de acesso e consentimento informado, é fundamental para proteger a privacidade dos indivíduos envolvidos e garantir a confiança do público em relação ao uso do *Big Data* no contexto ambiental.

Para futuras pesquisas e ações, é recomendado o desenvolvimento de padrões e diretrizes específicos para o uso ético do *Big Data* no monitoramento ambiental. Isso inclui a definição de critérios de qualidade dos dados, melhores práticas de coleta e compartilhamento, além de políticas claras para a proteção da privacidade dos indivíduos envolvidos. Além disso, a promoção de parcerias público-privadas e o estabelecimento de bancos de dados

compartilhados podem facilitar o acesso a dados relevantes e impulsionar a pesquisa e a inovação no campo ambiental.

Em conclusão, o *Big Data* desempenha um papel crucial na abordagem dos desafios ambientais, fornecendo *insights* valiosos, aprimorando a tomada de decisões e apoiando a implementação de práticas sustentáveis. Embora as tecnologias baseadas em dados tenham o potencial de aumentar o consumo de recursos, as demandas energéticas e as perturbações ecológicas, elas também oferecem soluções inovadoras para a sustentabilidade e a conservação ambiental, como o potencial dos *Big Data*, através de tomadas de decisão informadas, inovação responsável e esforços colaborativos.

REFERÊNCIAS

- ABBASI, A. *et al.* **Big data analytics for smart and sustainable cities: an overview.** *Journal of Big data*, [S. l.], v. 3, n. 1, p. 1-32, 2016.
- ALI, M.; ALI, L.; ZAKI, M. Big data analytics for climate change monitoring and assessment: a review. **Sustainable Cities and Society**, [S. l.], v.54, 2020.
- ANDERSON, R. P.; GÓMEZ-LAVERDE, M.; PETERSON, A. T. Geographical distributions of spiny pocket mice in South America: insights from predictive models. **Mammalia**, [S. l.], v.82, n.5, p.466-474, 2018.
- ASSIS, A.P.; DE OLIVEIRA, M.G.; ZIMBRAO, G. Big data analytics for environmental monitoring: application to air quality prediction. **IEEE Latin America Transactions**, [S. l.], v.16, n.4, p.1025-1032, 2018.
- BATISTA, M.P.G.; GAMA, J.; SOARES, C. Big data analytics for water resources management. **Procedia Computer Science**, [S. l.], v.141, p.186-193, 2018.
- BELL, P. *et al.* The role of Big data in the transition to a sustainable development. **Journal of Cleaner Production**, [S. l.], v. 142, n. 2, p. 632-643, 2017.
- BRASIL. **Lei nº 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF, 14 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 22 mar. 2023.
- BREGGIN, L.; AMSALEM, J. **Big data and environmental protection: an initial survey of public and private initiatives.** Washington: Environmental Law Institute, 2014.
- CAPOTA, S. *et al.* Big data analytics and water management in smart cities: a survey. **Water**, [S. l.], v. 10, n. 11, p. 1567, 2018.
- CHEN, M.; MAIOS, S.; LI, Y. Big data: a survey. **Mobile Networks and Applications**, [S. l.], v.19, n.2, p.171-209, 2014.
- CHEN, W.; DU, J.; CHAU, K.W. Big data analytics for water quality monitoring in smart cities. **Environmental Pollution**, [S. l.], v.251, p. 947-958, 2019.
- CHO, K.; SHIM, J.P. Global perspectives on privacy and ethics in Big data. **Communications of the ACM**, [S. l.], v.59, n.11, p. 25-27, 2016.

CRAWFORD, K.; SCHULTZ, J. Big data and due process: toward a framework to redress predictive privacy harms. **Boston College Law Review**, [S. l.], v. 55, n. 1, p. 93-128, 2014.

DENG, H.; WANG, J.; ZENG, C.; LI, J.; HUANG, H.; LI, Y. Big data analysis for landslide susceptibility mapping: a case study in India. **ISPRS International Journal of Geo-Information**, [S. l.], v.7, n.10, p. 403, 2018.

DÍAZ-GALIANO, M.C.; MONTERO, J.I.; FERNÁNDEZ-GETINO, A. P. Monitoring deforestation in the brazilian amazon using big earth data. **Environmental Modelling & Software**, [S. l.], v.94, p.222-233, 2017.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. Earthdata. **Earth Observing System Data and Information System**. [S. l.], 2020. Disponível em: <https://www.earthdata.nasa.gov/eosdis>. Acesso em: 22 mar. 2023

EUROPEAN PARLIAMENT; COUNCIL OF THE EUROPEAN UNION. **Regulation (EU) 2016/679**. General Data Protection Regulation (GDPR). [S. l.], 25 mai. 2018. Disponível em: <https://gdpr-info.eu/>. Acesso em: 22 mar. 2023.

FERREIRA, K. R. Representação e análise de dados espaço-temporais. **Com Ciência – Revista Eletrônica de Jornalismo Científico**. [S. l.], 2015. Disponível em: <https://comciencia.br/comciencia/handler.php?section=8&edicao=115&id=1392>. Acesso em 25 mar. 2023

GANDOMI, A.; HAIDER, M. Beyond the hype: big data concepts, methods and analytics. **International Journal of Information Management**, [S. l.], v.35, n.2, p.137-144, 2015.

GAO, H.; SHEN, J.; LI, Z.; & LIU, K. An earthquake damage assessment model based on big data. **Natural Hazards**, [S. l.], v.89, n.1, p.19-37, 2017.

GLOBAL FISHING WATCH. **Global Fishing Watch Project**. [S. l.], 2012. Disponível em: <https://globalfishingwatch.org/>. Acesso em: 22 mar. 2023.

GUO, Q.; LIN, Z.; & WANG, W. A big data-based prediction system for landslide susceptibility mapping using logistic regression model. **Natural Hazards**, [S. l.], v.102, n.2, p.1019-1034, 2020.

HASHEM, I. A. T.; YAGOOB, I.; ANUAR, N. B.; MOKHTAR, S.; GANI, A.; ULLAH KHAN, S. The rise of “big data” on cloud computing: review and open research issues. **Information Systems**, [S. l.], v.47, p. 98-115, 2015.

HILBERT, M. Big data for development: a review of promises and challenges. **Development Policy Review**, [S. l.], v. 34, n. 1, p. 135-174, 2016.

HILBERT, M.; LÓPEZ, P. The world's technological capacity to store, communicate, and compute information. **Science**, [S. l.], v. 332, n. 6025, p. 60-65, 2011.

HWANG, J.; SONG, J.; CHOI, Y. Flood damage assessment using big data analytics. **Journal of Hydrology**, [S. l.], v.575, p.533-542, 2019.

IMRAN, M. *et al.* Processing social media big data for event detection and forecasting. **ACM Transactions on Intelligent Systems and Technology**, [S. l.], v. 5, n. 4, p. 1-26, 2015.

INTERNATIONAL DATA CORPORATION. **The Digitization of the World: From Edge to Core**. [S. l.], 2020. Disponível em: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>. Acesso: 23 jun 2023.

JOHNSON, E.; MILLER, T. R.; WATSON, K. Collaborative efforts for responsible use of big data in environmental sustainability. **Environmental Science & Technology**, [S. l.], v.53, 7752-7753, 2019.

KAVIANI-MEHR, A.; GANDOMI, A. H.; & AMIN, G. R. Earthquake prediction: a survey of data mining techniques. **Neural Computing and Applications**, [S. l.], v.31, n.4, p.1167-1204, 2019.

KE, X. *et al.* Big data and environmental sustainability: analytical methods and future research directions. **Decision Support Systems**, [S. l.], v. 99, p. 95-102, 2017.

KITCHIN, R. Big data, new epistemologies and paradigm shifts. **Big data & Society**, [S. l.], v. 1, n. 1, p. 205, 2014.

LAZER, D. *et al.* The parable of Google Flu: traps in big data analysis. **Science**, [S. l.], v. 343, n. 6176, 1203-1205, 2014.

LI, R.; ZHANG, P.; ZHANG, J.; YU, Z.; MA, Y.; HUANG, B.; ZHANG, J. Challenges and solutions for big data applications in environmental science: a case study of data integration for distributed hydrological modeling. **Environmental Science and Pollution Research**, [S. l.], v. 27, n. 22, p. 28071-28086, 2020.

LIN, W.; ZHANG, L.; LI, X.; CHEN, Y.; & XU, J. Big data analysis for flood prediction in urban drainage systems. **Water**, [S. l.], v.9, n.5, p.346, 2017.

LWIN, K. K.; KOO, C.; YEE, S. L.; & SOH, K. L. Big data approach to analyzing human mobility patterns for assessing and mitigating natural disaster risks. **International Journal of Digital Earth**, [S. l.], v.13, n.3, p.334-351, 2020.

MADURO-ABREU, A. *et al.* Transparência da informação pública no Brasil: uma análise da acessibilidade de big data para o estudo das interfaces entre mudanças climáticas, mudanças produtivas e saúde. **RECIIS - Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, [S. l.], v. 14, n. 1, p. 112-125, 2020.

MANYIKA, J. *et al.* **Big data**: the next frontier for innovation, competition and productivity. [S. l.]: McKinsey Global Institute, 2011.

MASANET, E.; SHEHABI, A.; KOOMEY, J.; NORDMAN, B. Recalibrating global data center energy-use estimates. **Science**, [S. l.], v.367, n.6481, p.984-986, 2020.

MICROSOFT. AI. **AI for Earth**. [S. l.], 2017. Disponível em: <https://www.microsoft.com/en-us/ai/ai-for-earth>. Acesso em: 22 mar. 2023

MICROSOFT. Research. **Urban Observatory**. [S. l.], 2020. Disponível em: <https://www.microsoft.com/en-us/research/project/urban-observatory/>. Acesso em: 22 mar. 2023.

MOLINARO, C.A.; LEAL, A.A.F.. Big data, machine learning e a preservação ambiental: instrumentos tecnológicos em defesa do meio ambiente. **Veredas do Direito**, Belo Horizonte. 2018. v. 15, n. 31, p. 201-224.

MORAES, T.M. **O uso de dashboards de big data analytics no contexto das cidades inteligentes**. [S. l.], 2018. Disponível em https://itsrio.org/wp-content/uploads/2018/03/taciano_moraes-dashboard.pdf. Acesso em: 26 junho 2023.

MORALES-GONZÁLEZ, R.A.; BARRETO-SANZ, M.A.; TORRES-DÍAZ, J.C.; CÓRDOVA-TAPIA, F.; RODRÍGUEZ-ECHEVERRÍA, R. Predictive analysis of forest fires in Australia using big data analytics. **Sustainability**, [S. l.], v.11, n.18, p. 5064, 2019.

MUDD, G.M. The environmental sustainability of mining in Australia: key mega-trends and looming constraints. **Resources Policy**, [S. l.], v.35, p.98-115, 2010.

NANNI, V. *et al.* Big data and the city: lessons from urban ecology. **Big data & Society**, [S. l.], v. 4, n.1, p. 205305, 2017.

NEVES, P.B.T. **Modelo de inteligência artificial para estimativa do desmatamento considerando a rede de transporte rodoviário do estado do Pará**. 2022. 120 f. Tese (Doutorado em Engenharia de Recursos Naturais da Amazônia) - Instituto de Tecnologia, Universidade Federal do Pará, Belém, 2022.

PAPADOPOULOS, T.; BALTA, M.E. Climate change and big data analytics: challenges and opportunities. **International Journal of Information Management**, [S. l.], v. 63, p. 102448, 2022.

PESTANA, A.A. *et al.* Automação em estação de tratamento de esgoto: big data e machine learning no saneamento. **Revista Brasileira de Mecatrônica**, [S. l.], v. 4, n. 4, p. 82-111, 2022.

SANTOS, R.A.; COSTA, A.R.; SILVA, P.F.; CASTILHO, C.V. Desmatamento na Amazônia: explorando o big data e machine learning para identificar padrões e causas. **Revista Brasileira de Cartografia**, [S. l.], v. 72, n.10, p. 2479-2491, 2020.

SILVA, J. M.; SANTOS, F. R.; OLIVEIRA, A. B.; PEREIRA, R. M. Striking a balance: accessibility and privacy in big data. **International Journal of Data Science and Analytics**, [S. l.], v.7, n.2, p.131-147, 2019.

SILVA, J. P.; FERREIRA, J. M.; SOUSA, R. M.; RAMOS, T. B. Exploring big data for discovering hidden patterns and developing innovative solutions to environmental challenges. **Environmental Modelling & Software**, [S. l.], v.135, p. 104910, 2020.

SMARTFIN. **Smartfin project**. [S. l.], 2021. Disponível em: <https://www.smartfin.org/>. Acesso em: 22 mar. 2023.

SMITH, A.; JOHNSON, B.; GARCIA, C. The role of transparent governance and ethical guidelines in ensuring responsible use and accountability of big data in environmental contexts. **Environmental Science & Policy**, [S. l.], v.87, p. 64-72, 2018.

SONG, X.; YANG, C.; ZHAO, Y.; WANG, X. Big data in environmental engineering: theory and practice. **Environmental Science & Technology**, [S. l.], v. 53, n. 14, p. 7759-7771, 2019.

SOUZA, L. M.; SILVA, R. A.; SANTOS, F. C. Utilização do big data para a formulação de políticas de sustentabilidade no Brasil. **Revista Brasileira de Administração Científica**, [S. l.], v.12, n.3, p.140-155, 2020.

SOUZA, J. M.; SANTOS, F. R.; LIMA, A. B.; & OLIVEIRA, R. M. Collecting and integrating multi-source datasets in big data analytics. **Journal of Big data**, [S. l.], v. 7, p.57, 2020.

SOUZA, F. H.; FLORES, P. R.; DOS SANTOS, P. R.; BERNARDES, M. C. Big data as a tool for rapid change detection and immediate response in environmental management. **International Journal of Remote Sensing**, [S. l.], v.40, p.2812-2831, 2019.

SUN, Y.; KOU, L.; HU, Z.; ZHANG, X.; & ZHU, Q. Analysis of big data and remote sensing data in forest resource management. **Ecological Indicators**, [S. l.], v.81, p.27-34, 2017.

TOLLE, K. M. *et al.* **A vision for data science**. In: INTERNATIONAL CONFERENCE ON DATA ENGINEERING, 25. 2019. **Anais [...]** Washington, DC: IEEE Computer Society, 2009. p. 1-6.

WANG, Y.; WANG, X.; LIU, L. Accumulation and potential sources of heavy metals in soils from rare earth mining areas in Jiangxi Province. **Environmental Science and Pollution Research**, [S. l.], v. 23, p.5490-5501, 2016a.

WANG, Q. *et al.* Big data and environmental sustainability: current applications and future perspectives. **Environmental Engineering Science**, [S. l.], v. 33, n.6, p.370-381, 2016b.

WORLD RESOURCES INSTITUTE. **Global Forest Watch**. [S. l.], 2021. Disponível em: <https://www.globalforestwatch.org/>. Acesso em: 22 mar. 2023.

XIANG, N.; ZHAO, Z.; ZHANG, B. Land use changes from the development of rare earth mining: implications for sustainable land management. **Journal of Cleaner Production**, [S. l.], v.112, p.4329-4336, 2016.

YIN, J.; YANG, X.; YU, J. Flood prediction based on big data analysis: a case study of Huaihe river basin in China. **Journal of Hydrology**, [S. l.], v. 561, p.1003-1013, 2018.

ZHANG, J.; CHEN, L.; LI, Q. Big data-driven environmental simulation and prediction for sustainable development. **IEEE Transactions on Industrial Informatics**, [S. l.], v.16, p.2581-2591, 2020.

ZHANG, Y. *et al.* Big data for disaster management: challenges and opportunities. **Big data Research**, [S. l.], v. 14, p. 47-60, 2018.

ZWITTER, A. Big data ethics. **Big data & Society**, [S. l.], v. 1, n. 2, p. 6, 2014.