



UERN – Universidade do Estado do Rio Grande do Norte
Curso de Ciências da Computação

Antoniél da Silva Alves (20250017003)
Pedro Gabriel Machado (20250038783)
Pedro Henrique Da Silva (20250046471)
Geovane Guilherme (20250038522)

Data Science e Big Data: novos paradigmas epistemológicos para a construção do conhecimento científico.

Professor: Mestre Dácio Michel Da Cruz Souza

Natal-RN

2025

Sumário

Sumário

| | |
|--|-----------|
| 1.Introdução | 3 |
| 2.Objetivos..... | 4 |
| 2.1 Objetivos Específicos | 4 |
| 3.Qual é o conceito de epistemologia? | 5 |
| 4.Como funciona a Data Science(Ciência de Dados). | 5 |
| 5.O papel do conhecimento hoje. | 6 |
| 6.Epistemologia e Ciência..... | 6 |
| 7.O que é Big Data..... | 7 |
| 8.Big Data e Epistemologia..... | 7 |
| 9.Por que Data Science precisa da Epistemologia? | 7 |
| 10.Desafios Éticos e Epistemológicos. | 8 |
| 11.Exemplo aplicado: | 8 |
| 11.1-Covid 19..... | 8 |
| 11.2 Previsão do Tráfego Urbano..... | 8 |
| 13.Considerações finais. | 10 |
| 14.Referências | 11 |

1.Introdução

Este trabalho foi elaborado com base no tema proposto pelo professor: “Big Data” e “Data Science”. No entanto, a forma como iremos abordar o assunto não será estritamente técnica, sob a perspectiva do desenvolvimento tecnológico, mas sim a partir de uma visão epistemológica.

De maneira sintética, podemos compreender essa abordagem como a busca pelo resultado de uma pesquisa, entendendo, em um sentido mais literário, o discurso, a teoria ou a tese sobre determinado conhecimento. Assim, discutiremos o estudo crítico dos principais problemas e hipóteses para a obtenção de resultados, mostrando como a *Data Science* e o *Big Data* podem se valer desses mecanismos, tanto no âmbito digital quanto em áreas diversas, como a saúde e suas aplicações práticas.

2.Objetivos

Este trabalho tem o intuito de apresentar, de maneira sintética, os conceitos de Big Data, Data Science e Epistemologia. A proposta não é abordar aspectos técnicos do desenvolvimento tecnológico, mas sim refletir sobre como esses campos se relacionam com a produção de conhecimento, a formulação de hipóteses e a busca por resultados em diferentes áreas, como o ambiente digital e setores como a saúde.

2.1 Objetivos Específicos

- Explicar de forma básica o conceito de Epistemologia.
- Apresentar o que é Data Science e sua relevância atual.
- Discutir o papel do conhecimento na sociedade contemporânea.
- Relacionar Epistemologia e Ciência, destacando sua importância no processo de validação do saber.
- Conceituar Big Data e suas principais características.
- Analisar a relação entre Big Data e Epistemologia.
- Justificar por que a Data Science necessita da Epistemologia como base crítica e metodológica.
- Refletir sobre os desafios éticos e epistemológicos relacionados ao uso de grandes volumes de dados.
- Apresentar exemplos práticos de aplicações do Big Data e da Data Science em diferentes áreas, como saúde, tecnologia e sociedade.

3.Qual é o conceito de epistemologia?

A Epistemologia pode ser entendida como o ramo da filosofia que estuda a natureza, a origem e os limites do conhecimento, podemos até falar que a base da epistemologia do seu significado seria discurso, a teoria, a tese sobre a ciência. Em que ela busca responder a questões fundamentais como: o que é conhecimento? Como o conhecimento pode ser validado? Em que condições algo pode ser considerado verdadeiro?

Segundo Frické (2015), a epistemologia é essencial para analisar como o conhecimento científico é construído e validado, especialmente diante do desafio do Big Data, em que grandes volumes de dados podem gerar informações sem necessariamente garantir entendimento ou verdade.

Para Desai et al. (2022), a epistemologia oferece uma base crítica para a Data Science, pois não basta apenas extrair padrões de dados: é necessário compreender quais métodos são confiáveis e em que medida os resultados podem ser considerados conhecimento válido.

Já Silva Júnior, Karpinski e Dutra (2020) ressaltam, a partir da epistemologia de Karl Popper, que a ciência avança pelo processo de formulação e refutação de hipóteses. Isso significa que, mesmo em um contexto de Big Data, o conhecimento não se resume a acumular dados, mas sim a submeter hipóteses a testes rigorosos para aproximar-se da verdade científica

4.Como funciona a Data Science(Ciência de Dados).

A Data Science ou no português brasileiro, Ciência de Dados, é uma área que combina estatística, matemática, ciência da computação e conhecimentos de domínio específico para extrair informações e gerar conhecimento a partir de dados. Seu objetivo principal é transformar grandes volumes de dados, em algum úteis para a tomada de decisão, como por exemplo pegar diversos dados para uma implementação de uma vacina em um determinado local, estudo para criação de grandes mercado, adicionar empresas, é muito amplo a quantidade de possibilidade possível a partir desse grande números de dados, não é atoa que é bem comum vermos que hoje em dia um do nossos principais bens são

as nossas informações que infelizmente no meio digital com as redes sociais se tornam bem expostas e de certa forma vulneráveis e balizadas.

Agora com base em algumas das nossas pesquisas, segundo Desai et al. (2022), a Data Science não se limita a técnicas de processamento ou análise, mas envolve também uma reflexão crítica sobre a confiabilidade e a validade dos métodos empregados. Já para Frické (2015), ela se relaciona diretamente com a epistemologia, pois levanta questões sobre quando um conjunto de dados realmente pode ser considerado conhecimento.

Na prática, a Data Science é aplicada em diversas áreas, como saúde, economia, indústria e redes sociais, oferecendo suporte para prever comportamentos, identificar padrões e resolver problemas complexos, ou pelo menos tentar solucionar principalmente que a nossa área se baseia muito em tentar solucionar um problema existente, como a maior parte da área da ciência da computação é justamente solucionar um problema existente ou tentar prever um problema que possa ocorrer e já solucionar ele, como o exemplo em foi dado em aula sobre a Margaret Hamilton em que ela previu e já solucionou um problema em uma viagem à lua.

5.O papel do conhecimento hoje.

Atualmente, o conhecimento exerce um papel central na sociedade, funcionando como recurso estratégico para o desenvolvimento científico, tecnológico e social. Em um mundo marcado pela abundância de informações e pela rapidez das transformações digitais, o conhecimento se torna essencial para interpretar dados, orientar decisões e promover inovações. Mais do que acumular informações, é necessário saber filtrar, analisar criticamente e aplicar esse conhecimento em diferentes contextos, garantindo avanços de forma responsável e ética.x

6.Epistemologia e Ciência.

Dentro da epistemologia, a reflexão sobre o conhecimento se dá através de várias vertentes, como o empirismo, o racionalismo e o construtivismo, cada uma oferecendo uma visão diferente sobre o processo de aquisição de conhecimento.

A **ciência**, por sua vez, pode ser vista como uma das formas mais estruturadas e sistemáticas de busca por conhecimento.

7.O que é Big Data.

Big Data refere-se a conjuntos de dados massivos, complexos e de rápida evolução, que são difíceis de gerir e analisar com ferramentas tradicionais de software. É caracterizado pelos "Vs" como Volume, Velocidade e Variedade de dados, provenientes de diversas fontes como redes sociais, sensores e transações

8.Big Data e Epistemologia.

Big Data refere-se ao processamento de enormes volumes de dados, em alta velocidade e variedade, usados para gerar previsões e insights. Do ponto de vista epistemológico, isso levanta a questão: ter muitos dados significa ter conhecimento? A reflexão mostra que dados só se tornam conhecimento quando interpretados criticamente e inseridos em teorias, caso contrário permanecem apenas como informação bruta.

9.Por que Data Science precisa da Epistemologia?

A Data Science lida com grandes quantidades de dados e utiliza métodos estatísticos, computacionais e matemáticos para gerar informações relevantes. No entanto, transformar dados em conhecimento confiável exige mais do que técnica: é necessário refletir sobre como os resultados são produzidos, validados e interpretados. É justamente nesse ponto que a epistemologia se torna essencial.

A epistemologia fornece critérios para diferenciar o que é apenas informação do que pode ser considerado conhecimento. Ela ajuda a questionar se os métodos utilizados são adequados, se os resultados obtidos são válidos e se as conclusões realmente representam a realidade ou apenas padrões superficiais.

Assim, sem o suporte epistemológico, a Data Science corre o risco de reduzir-se a simples processamento de dados, sem garantir a solidez do conhecimento gerado. Com a epistemologia, por outro lado, é possível dar

fundamentação crítica e científica às descobertas, assegurando maior confiabilidade e relevância às suas aplicações em áreas como saúde, economia, política e tecnologia.

10.Desafios Éticos e Epistemológicos.

Na era do Big Data, surgem dilemas que vão além da técnica. Do lado ético, aparecem questões de privacidade, uso responsável dos dados e impactos sociais das decisões algorítmicas. Do lado epistemológico, o desafio é transformar grandes volumes de informação em conhecimento confiável, evitando vieses, interpretações superficiais e a ilusão de que “mais dados” significam automaticamente “mais verdade”.

11.Exemplo aplicado:

Também é possível mostrar de forma breve como a epistemologia pode se articular com as tecnologias de dados. A partir dessa relação, pode-se pensar em aplicações práticas em diferentes áreas, onde o uso crítico do conhecimento, aliado às ferramentas da Data Science e do Big Data, permite não apenas organizar e interpretar informações, mas também gerar resultados mais consistentes e com maior valor científico e social.

11.1-Covid 19

Também é possível mostrar de forma breve como a epistemologia pode se articular com as tecnologias de dados. A partir dessa relação, pode-se pensar em aplicações práticas em diferentes áreas, onde o uso crítico do conhecimento, aliado às ferramentas da Data Science e do Big Data, permite não apenas organizar e interpretar informações, mas também gerar resultados mais consistentes e com maior valor científico e social.

11.2 Previsão do Tráfego Urbano

A previsão do tráfego urbano é um exemplo de aplicação de Data Science e Big Data, onde grandes volumes de dados sobre veículos, sensores de ruas e transporte público são analisados para antecipar congestionamentos e otimizar o fluxo de trânsito. Quando combinada com uma abordagem epistemológica, essa análise não se limita a números e padrões: permite interpretar criticamente

os dados, identificar limitações dos modelos e tomar decisões mais confiáveis e fundamentadas para melhorar a mobilidade urbana.

13.Considerações finais.

A epistemologia mostra-se fundamental para a ciência de dados e o Big Data, pois nos lembra que o conhecimento não se resume à coleta de informações ou a crenças aceitas como verdade. Embora muitas vezes nossas ideias ou hipóteses possam parecer corretas, questões sociais, científicas e éticas exigem comprovação e análise rigorosa. O estudo científico, mesmo sujeito a erros, continua sendo o meio mais confiável e estruturado para transformar dados e observações em conhecimento sólido.

Além disso, a epistemologia fornece critérios para avaliar a validade dos métodos empregados na Data Science e no Big Data, orientando a interpretação dos resultados e garantindo que as conclusões sejam fundamentadas. Ela permite analisar criticamente os processos de geração de conhecimento, relacionando hipóteses, testes e dados de forma consciente, evitando que decisões se baseiem apenas em padrões ou correlações superficiais.

Portanto a combinação entre epistemologia e ciência de dados não apenas orienta como interpretamos os dados, mas também reforça a necessidade de aplicar métodos científicos bem estabelecidos. Isso possibilita resultados mais confiáveis, decisões mais precisas e aplicações mais éticas, seja no contexto digital ou em setores diversos, como saúde, transporte e gestão social. Em suma, a epistemologia fortalece a prática da Data Science, promovendo uma compreensão crítica e responsável do conhecimento gerado a partir de grandes volumes de dados.

Referências

DESAI, Jules; WATSON, David; WANG, Vincent; TADDEO, Mariarosaria; FLORIDI, Luciano. *The epistemological foundations of data science: a critical review*. **Synthese**, v. 200, n. 6, p. 1-27, 2022. DOI: 10.1007/s11229-022-03933-2. Disponível em: <https://link.springer.com/article/10.1007/s11229-022-03933-2>.

FRICKÉ, Martin. *Big Data and Its Epistemology*. **Journal of the Association for Information Science and Technology**, v. 66, n. 4, p. 651-661, 2015. DOI: 10.1002/asi.23212. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.23212>.

SILVA JÚNIOR, Eugênio Monteiro da; KARPINSKI, Cezar; DUTRA, Moisés Lima. *Scientific knowledge in big data context: reflections from Popper's epistemology*. **Brazilian Journal of Information Science: Research Trends**, v. 14, n. 4, 2020. DOI: 10.36311/1940-1640.2020.v14n4.10936. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/10936>.

CHAHUARA QUISPE, José Carlos. *Big data, epistemological problems and epistemic contextualism*. **Escritura y Pensamiento**, v. 21, n. 45, p. 33-53, 2022. DOI: 10.15381/escrypensam.v21i45.23806. Disponível em: <https://revistasinvestigacion.unmsm.edu.pe/index.php/letras/article/view/23806>.

HOFFMANN, Daniel Sander. *Big Data and Deep Learning Models*. **Principia: an international journal of epistemology**, v. 26, n. 3, p. 597-614, 2022. DOI: 10.5007/1808-1711.2022.e84419. Disponível em: <https://periodicos.ufsc.br/index.php/principia/article/view/84419>.