

ASSIGNMENT 1 FRONT SHEET

Qualification	BTEC Level 5 HND Diploma in Computing			
Unit number and title	Unit 13:Computing Research Project			
Submission date	3/5/2024	Date Received 1st submission	8/5/2024	
Re-submission Date		Date Received 2nd submission		
Student Name	Tran Xuan Truong	Student ID	BH00127	
Class	IT0502	Assessor name	Nguyen Thanh Trieu	

Student declaration

I certify that the assignment submission is entirely my own work and I fully understand the consequences of plagiarism. I understand that making a false declaration is a form of malpractice.

Student's signature	Truong

Grading grid

P1	P2	P3	P4	P5	M1	M2	M3	D1	D2



☐ Summative Feedback:		☐ Resubmission Feedback:			
Grade:	Assessor Signature:		Date:		
Internal Verifier's Commen	nts:				
Signature & Date:	Signature & Date:				



Content of Tables

A.	INT	RODUCTION	4
В. Н У		ODUCE A RESEARCH PROPOSAL THAT CLEARLY DEFINES A RESEARCH QUESTION OF HESIS SUPPORTED BY A LITERATURE REVIEW.(P1)	
	1.	Research Topic	5
	2.	Abstracts	5
	3.	Situation	5
	4.	Define the main aims and objectives of the project.	5
	4.	.1 Aims	5
	4.	.2 Objectives	6
	5.	Project Plan	6
	5.	.1 Work Breakdown	6
	5.	.2 Time Estimates	7
	5.	.3 Milestone identification	8
	5.	.4 Scheduling	9
		AM APPROPRIATE RESEARCH METHODS AND APPROACHES TO PRIMARY AND	
		DARY RESEARCH (P2)	
1		Research Methods	
	1.1.	•	
	1.2.	·	
	1.3.		
	1.4.	Qualitative Research	. 13
	1.5.		
	1.6.	Compare Qualitative with Qualitative	. 17
	1.7.	Scientific method	. 17
	1.8.	Research process	. 20
	1.9.	Population research	. 21
FO]	R A (CONDUCT PRIMARY AND SECONDARY RESEARCH USING APPROPRIATE METHODS COMPUTING RESEARCH PROJECT THAT CONSIDER COSTS, ACCESS AND ETHICAL	
1		econdary Research	
	1.1.	Sources:	. 23



2	. Pr	imary research	26
	2.1.	Interview	27
	2.2.	Evaluation Survey	28
E.	P4 A 29	PPLY APPROPRIATE ANALYTICAL TOOLS, ANALYSE RESEARCH FINDINGS AND D)ATA
1	. In	terView	29
2	. Int	terView Summary	37
3	. Su	ırvey	38
4	. Su	rvey Summary	41
	4.1.	Analyze the results of the primary research.	42
F.	P5 C	COMMUNICATE RESEARCH OUTCOMES IN AN APPROPRIATE MANNER FOR THE	
IN		ED AUDIENCE.	43
1	. Co	onclusion	43
2	. Re	ecommendations	44
G.	CON	ICLUSION	44
Н.	REF	ERENCE	45
Ta	ble o	f Figure	
Figu	ıre 1: \	WBS	7
_		Milestone	
Figu	ıre 3: 9	Scheduling	9
Figu	ıre 4: 9	Survey 1	39
Figu	ıre 5: S	Survey	40
Figu	ıre 6: S	Survey	40
Figu	ıre 7: S	Survey	41
Figu	ıre 8: 9	Survey	41

A. INTRODUCTION

The concept of "big data" refers to the process of collecting, storing and analyzing data on a large scale, playing an important role in many aspects of modern life. Although it brings many benefits, this process also causes unwanted impacts on the environment, especially the use of cheap resources and pollution. The article will give an overview of the challenges that data poses to the environment and explore potential



solutions, including finding and using more robust alternative materials, trying to reduce minimize these negative effects.

B. PRODUCE A RESEARCH PROPOSAL THAT CLEARLY DEFINES A RESEARCH QUESTION OR HYPOTHESIS SUPPORTED BY A LITERATURE REVIEW.(P1)

1. Research Topic

"Environmental impact and search for alternative materials in big data storage models."

2. Abstracts

The increasing volume of data generation and storage requires increased electricity consumption and additional hardware, thereby increasing the environmental impact of the digital sector. This growing concern highlights the environmental footprint of devices and services, as IT services become increasingly indispensable to people's lives. These apparent environmental impacts have prompted a study of the impact of big data models. This project aims to investigate the environmental consequences of big data models and propose alternative materials to mitigate these impacts, focusing on increasing energy and resource consumption, greenhouse gas emissions from data centers, e-waste pollution, and inefficient data management.

3. Situation

The advent of Big Data has revolutionized how organizations manage, analyze, and leverage data. However, the increased usage and storage of data have significantly impacted the environment, contributing to electronic waste and enhancing the greenhouse effect. These developments pose challenges to the natural environment, underlining the need for sustainable practices in data management and storage to mitigate these adverse effects.

4. Define the main aims and objectives of the project.

4.1 Aims

• The primary goal is to quantify and understand the environmental impact of current big data storage models. This involves determining the amount of energy consumed, greenhouse gas emissions produced, and electronic waste generated during the manufacture, operation, and disposal of data storage devices. This assessment aims to identify key issues and pinpoint areas where improvements can be made to lessen the environmental impact.



- The aim is to explore and identify potential alternative materials to minimize the environmental impact of big data storage models. This involves the research and development of new, eco-friendly materials that require fewer resources and produce less waste during the production and disposal of data storage devices. The objective is to discover alternative material solutions that can foster a more sustainable big data storage model, thereby reducing its ecological footprint.
- The ultimate goal is to foster the development and implementation of environmentally friendly big data storage solutions. By comprehensively understanding environmental impacts, researching alternative materials, and proposing solutions and policies, the aim is to encourage the adoption of sustainable big data storage practices. This endeavor strives to ensure the efficient use of natural resources and the protection of the environment for future generations, highlighting the importance of sustainability in technological advancements.

4.2 Objectives

In the realm of big data, the extensive use of data centers and cloud computing necessitates significant energy and material resources, leading to notable environmental impacts. This includes the entire lifecycle of data storage devices, from manufacturing to disposal, which contributes to energy consumption, greenhouse gas emissions, and electronic waste. The first objective involves a detailed environmental impact assessment of these storage models, aiming for a deeper understanding and identification of major concerns. The second objective focuses on discovering and developing alternative, eco-friendly materials for data storage to reduce these environmental footprints. This research and development effort focuses on materials that are less resource-intensive and generate less waste. Ultimately, the project seeks to promote sustainable big data storage solutions, encouraging practices that ensure the efficient use of natural resources and the preservation of our environmental heritage for future generations.

5. Project Plan

5.1 Work Breakdown



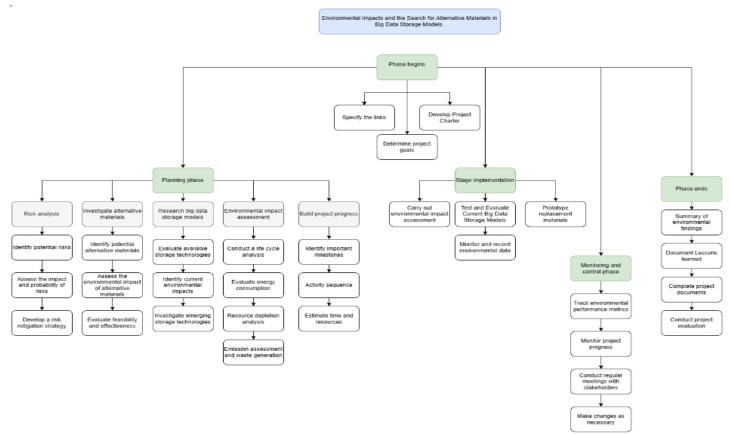


Figure 1: WBS

5.2 Time Estimates



- Planning Phase(2,5 Month)
- + Risks Analysis (2 Week)
- + Investigate alternative materials (2 Week)
- + Research big data storage models (2 Week)
- + Environmental impact assessment (2 Week)
- + Build project progress (2 Week)
- Phase begins (1 Week)
- Stage implementation (1 Week)
- Monitoring and control phase (3 Week)
- Phase ends (1 Week)
- Completion report: 5 Week
- Total Effort: 16 Week, 4,5 Month

5.3 Milestone identification



Figure 2: Milestone



5.4 Scheduling

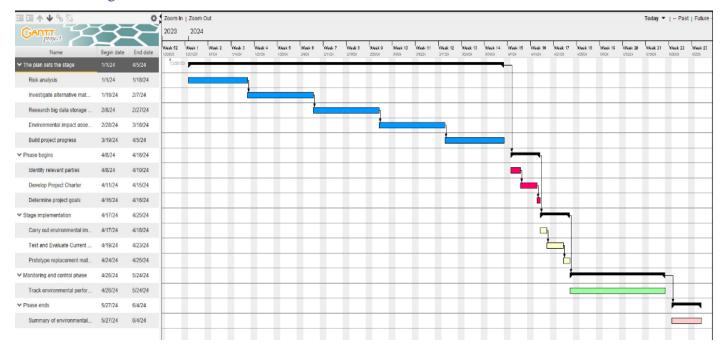


Figure 3: Scheduling

C. EXAM APPROPRIATE RESEARCH METHODS AND APPROACHES TO PRIMARY AND SECONDARY RESEARCH (P2)

1. Research Methods

"Research methods" are the strategies, techniques, and procedures that researchers use to collect, analyze, and process data to answer research questions, test instrumental hypotheses, or investigate the phenomenon. This method is very important to ensure the accuracy, reliability and generalizability of research results. Research methods may vary depending on the field of study, the nature of the research question, and available resources.

1.1.Primary Research

Primary research is a research method that relies on direct data collection, rather than relying on data that's already been collected by someone else. In other words, primary research is any type of research that you undertake yourself, firsthand, while using data that has already been collected is called secondary research.

↓ Type of primary research



Primary research can take many forms, but the most common types are:

- Surveys and questionnaires
- Observational studies
- ➤ Interviews and focus groups

Surveys and questionnaires

Surveys and questionnaires collect information about a group of people by asking them questions and analyzing the results. They are a solid choice if your research topic seeks to investigate something about the characteristics, preferences, opinions, or beliefs of a group of people.

Surveys and questionnaires can take place online, in person, or through the mail. It is best to have a combination of open-ended and closed-ended questions, and how the questions are phrased matters. Be sure to avoid leading questions, and ask any related questions in groups, starting with the most basic ones first.

Observational studies are an easy and popular way to answer a research question based purely on what you, the researcher, observes. If there are practical or ethical concerns that prevent you from conducting a traditional experiment, observational studies are often a good stopgap.

There are three types of observational studies: cross-sectional studies, cohort studies, and case-control studies. If you decide to conduct observational research, you can choose the one that's best for you. All three are quite straightforward and easy to design just beware of confounding variables and observer bias creeping into your analysis.

Likewise, to surveys and questionnaires, interviews and focus groups also rely on asking questions to collect information about a group of people. However, how this is done is slightly different. Instead of sending your questions out into the world, interviews and focus groups involve two or more people one of whom you are, the interviewer, who asks the questions.

There are 3 main types of interviews:

Structured interviews ask predetermined questions in a predetermined order.
 Unstructured interviews are more flexible and free flowing, conducted based on the interviewee's previous answers.



Semi-structured interviews fall in between, offering a mix of predetermined and random questions. While interviews are a rich source of information, they can also be challenging to do well. Be careful of interviewer bias creeping into your process. The best way to minimize this is to avoid double-barreled questions and pay close attention to your tone and phrasing when asking questions.

Additionally, a focus group is a group interview, led by a moderator. Focus groups can provide more nuanced interactions than individual interviews, but their small sample size means that external validity is low.

Advantages of primary research

- + The ability to conduct really tailored, thoroughly research, down to the "nitty-gritty" of your topic. You decide what you want to study or observe and how to go about doing that.
- + You maintain control over the quality of the data collected, and can ensure firsthand that it is objective, reliable, and valid.
- + The subsequent results are yours, for you to disseminate as you see fit. You maintain proprietary control over what you find out, allowing you to share your findings with like-minded individuals or those conducting related research that interests you for replication or discussion purposes.

♣ Disadvantage of primary research.

- + To be done well, primary research can be very expensive and time consuming. If you are constrained in terms of time or funding, it can be very difficult to conduct your own high-quality primary research.
- + Primary research is often insufficient as a standalone research method, requiring secondary research to bolster it. Primary research can be prone to various types of research bias. Bias can manifest on the part of the researcher as observer bias, Pygmalion effect, or demand characteristics. It can occur on the part of participants as a Hawthorne effect or social desirability bias.

1.2. Secondary Research

Secondary research is a research method that uses data that was collected by someone else. In other words, whenever you conduct research using data that already exists, you are conducting secondary research. On the other hand, any type of research that you undertake yourself is called primary research.

Here are the steps involved in conducting secondary research:



➤ Define Your Research Topic and Question:

A research topic is a specific part of study in a broader area of study. For instance, for your research topic, the broader research area is malaria prevention in households. A research question aims to further narrow down the scope of the study.

> Select a Dataset:

Use resources like the Society of General Internal Medicine's Online Compendium to select a dataset.

Consider choosing a dataset that hasn't been extensively utilized in your field or combining datasets for a fresh perspective. Consider the complexity, cost, and time involved in obtaining the dataset. It may be helpful to seek advice from your mentor if they have experience with a particular dataset.

✓ *Become Familiar with Your Dataset:*

✓ Gain an understanding of key aspects of the dataset:

- + What is the purpose of the dataset?
- + Who is responsible for reporting the data?
- + What motivations exist for accurate reporting?
- + Is the data subject to auditing?
- + Is there potential for linking it to other extensive datasets?
- + Thoroughly research the database and verify if your metrics have been validated against alternative sources.
- + Become acquainted with the data through analysis or by reviewing outputs if programming is handled by another party.

✓ *Structure Your Analysis and Presentation:*

- + Consider the clinical implications of your discoveries.
- + Exercise caution when interpreting statistical significance (p-values). Large sample sizes can produce statistically significant correlations that may lack clinical relevance. Seek guidance from a statistician for intricate datasets and analyses.
- + Utilize visual aids, such as figures, to effectively communicate your findings.

Advantages of secondary research

- + Secondary data is very easy to source and readily available.
- + It is also often free or accessible through your educational institution's library or network, making



it much cheaper to conduct than primary research.

- + As you are relying on research that already exists, conducting secondary research is much less time consuming than primary research. Since your timeline is so much shorter, your research can be ready to publish sooner.
- + Using data from others allows you to show reproducibility and replicability, bolstering prior research and situating your own work within your field.

♣ Disadvantages of secondary research

- + Ease of access does not signify credibility. It's important to be aware that secondary research is not always reliable and can often be out of date. It's critical to analyze any data you're thinking of using prior to getting started, using a method like the CRAAP test.
- + Secondary research often relies on primary research already conducted. If this original research is biased in any way, those research biases could creep into the secondary results.

1.3. Compare Primary Research with Secondary Research

Primary research and secondary research are two types of research methods that can be used to collect and analyze data. The main difference between them is the source of the data. Here is a brief comparison of primary research and secondary research:

Primary research is the process of gathering original data directly from the participants or phenomena of interest. For example, conducting surveys, interviews, experiments, or observations. Primary research is often done to answer specific research questions or test hypotheses. Primary research can provide recent, relevant, and customized data, but it can also be time-consuming, costly, and difficult to generalize. Secondary research is the process of using existing data that was collected by someone else for a different purpose. For example, reviewing academic journals, government reports, statistics, or online sources. Secondary research is often done to gain background knowledge, identify trends, or find gaps in the literature. Secondary research can provide comprehensive, reliable, and accessible data, but it can also be outdated, irrelevant, or biased.

Both primary research and secondary research have their advantages and disadvantages, and they can be used together or separately depending on the research objectives, resources, and limitations. Most research projects use a combination of both primary and secondary research methods to achieve the best results.

1.4.Qualitative Research

Qualitative research involves collecting and analyzing non-numerical data e.g., text, video, or audio to understand concepts, opinions, or experiences. It can be used to gather in-depth insights into a problem or



generate new ideas for research.

Qualitative research is the opposite of quantitative research, which involves collecting and analyzing numerical data for statistical analysis.

♣ Qualitative research methods

- ➤ Each of the research approaches involve using one or more data collection methods. These are some of the most common qualitative methods:
 - + **Observations:** recording what you have seen, heard, or encountered in detailed field notes.
 - + **Interviews:** personally, asking people questions in one-on-one conversations.
 - + **Focus groups:** asking questions and generating discussion among a group of people.
 - + **Surveys:** distributing questionnaires with open-ended questions.
 - + **Secondary research:** collecting existing data in the form of texts, images, audio or video recordings, etc.

Qualitative data analysis

- Qualitative data encompasses various forms such as texts, photos, videos, and audio. For
 instance, you may be dealing with interview transcripts, survey responses, fieldnotes, or
 recordings from natural settings. Most qualitative data analysis follows five key steps:
- 1. Prepare and organize your data: This involves tasks like transcribing interviews or digitizing fieldnotes to ensure they are easily accessible and manageable.
- 2. Review and explore your data: Take time to thoroughly examine the data for emerging patterns, recurring themes, or significant ideas.
- 3. Develop a data coding system: Based on initial insights, establish a coding framework to systematically categorize and organize the data.
- 4. Assign codes to the data: Apply the coding system to the data, tagging specific segments or elements with relevant codes. This process may involve going through each piece of data and labeling it accordingly.
- 5. Identify recurring themes: Analyze coded data to identify overarching themes or patterns that emerge across different segments.

Advantages of qualitative research

Qualitative research often tries to preserve the voice and perspective of participants and can be adjusted as new research questions arise. Qualitative research is good for:

+ **Flexibility:** The data collection and analysis process can be adapted as new ideas or patterns



emerge. They are not rigidly decided beforehand.

- + **Natural settings:** Data collection occurs in real-world contexts or in naturalistic ways.
- + **Meaningful insights:** Detailed descriptions of people's experiences, feelings and perceptions can be used in designing, testing or improving systems or products.
- + **Generation of new ideas:** Open-ended responses mean that researchers can uncover novel problems or opportunities that they wouldn't have thought of otherwise.

♣ Disadvantages of qualitative research

Researchers must consider practical and theoretical limitations in analyzing and interpreting their data. Qualitative research suffers from:

- + **Unreliability:** The real-world setting often makes qualitative research unreliable because of uncontrolled factors that affect the data.
- + **Subjectivity:** Due to the researcher's primary role in analyzing and interpreting data, qualitative research cannot be replicated. The researcher decides what is important and what is irrelevant in data analysis, so interpretations of the same data can vary greatly.
- + **Limited generalizability**: small samples are often used to gather detailed data about specific contexts. Despite rigorous analysis procedures, it is difficult to draw generalizable conclusions because the data may be biased and unrepresentative of the wider population.
- + **Labor-intensive**: Although software can be used to manage and record large amounts of text, data analysis often has to be checked or performed manually.

1.5.Quantitative research.

Quantitative research is the process of collecting and analyzing numerical data. It can be used to find patterns and averages, make predictions, test causal relationships, and generalize results to wider populations.

4 Quantitative research methods

You can use quantitative research methods for descriptive, correlational or experimental research.

- + In descriptive research, you simply seek an overall summary of your study variables.
- + In correlational research, you investigate relationships between your study variables.
- + In experimental research, you systematically examine whether there is a cause-and-effect relationship between variables.
- + Correlational and experimental research can both be used to formally test hypotheses, or predictions, using statistics. The results may be generalized to broader populations based on the



sampling method used.

+ To collect quantitative data, you will often need to use operational definitions that translate abstract concepts (e.g., mood) into observable and quantifiable measures (e.g., self-ratings of feelings and energy levels).

Quantitative data analysis

Once data is collected, you may need to process it before it can be analyzed. For example, survey and test data may need to be transformed from words to numbers. Then, you can use statistical analysis to answer your research questions.

Descriptive statistics will give you a summary of your data and include measures of averages and variability. You can also use graphs, scatter plots and frequency tables to visualize your data and check for any trends or outliers.

Advantages of Quantitative research

Quantitative research is often used to standardize data collection and generalize findings. Strengths of this approach include:

- + **Replication:** Repeating the study is possible because of standardized data collection protocols and tangible definitions of abstract concepts.
- + **Direct comparisons of results:** The study can be reproduced in other cultural settings, times or with different groups of participants. Results can be compared statistically.
- + **Large samples:** Data from large samples can be processed and analyzed using reliable and consistent procedures through quantitative data analysis.
- + Hypothesis testing: Using formalized and established hypothesis testing procedures means that you have to carefully consider and report your research variables, predictions, data collection and testing methods before coming to a conclusion.

Disadvantages of Quantitative research

Despite the benefits of quantitative research, it is sometimes inadequate in explaining complex research topics. Its limitations include:

- + **Superficiality:** Using precise and restrictive operational definitions may inadequately represent complex concepts. For example, the concept of mood may be represented with just a number in quantitative research but explained with elaboration in qualitative research.
- + **Narrow focus:** Predetermined variables and measurement procedures can mean that you ignore other relevant observations.



- + **Structural bias:** Despite standardized procedures, structural biases can still affect quantitative research. Missing data, imprecise measurements or inappropriate sampling methods are biases that can lead to the wrong conclusions.
- + **Lack of context:** Quantitative research often uses unnatural settings like laboratories or fails to consider historical and cultural contexts that may affect data collection and results.

1.6. Compare Qualitative with Qualitative

Qualitative Research:

- + Focus: Qualitative research deals with words and meanings.
- + Purpose: It is used to understand concepts, thoughts, or experiences.
- + Insights: Qualitative methods enable you to gather in-depth insights on topics that are not well understood.

Methods:

- + Interviews: Conducted with open-ended questions.
- + Observations: Described in words.
- + Literature Reviews: Explore concepts and theories.
- + Biases: Qualitative research is susceptible to biases such as the Hawthorne effect, observer bias, recall bias, and social desirability bias.

Quantitative Research:

- + Focus: Quantitative research deals with numbers and statistics.
- + Purpose: It is used to test or confirm theories and assumptions.
- + Generalizability: Quantitative research establishes generalizable facts about a topic.

Methods:

- + Surveys: Closed or multiple-choice questions.
- + Experiments: Manipulating variables to establish cause-and-effect relationships.
- + Observations: Recording data as numbers.
- + Biases: Quantitative research is at risk for biases including information bias, omitted variable bias, sampling bias, or selection bias.

1.7. Scientific method

- The scientific method in technology and computers

The scientific method serves as a structured approach utilized by scientists and researchers across



various disciplines to explore natural phenomena, formulate hypotheses, conduct experiments, and analyze outcomes. In the realms of technology and computer science, this method is instrumental in the development, testing, and enhancement of theories, algorithms, and systems.

Here's a breakdown of how the scientific method is commonly applied in research within the fields of technology and computer science:

- + Observation and Inquiry: Researchers keenly observe phenomena or identify challenges within the domains of technology or computer science, often arising from practical hurdles, industry requirements, or gaps in existing knowledge.
- + Question Formulation: Drawing from observations, researchers articulate precise questions or issues suitable for research exploration.
- + Literature Review: Researchers systematically review existing knowledge domains through extensive literature exploration, encompassing theories, technologies, and methodologies pertinent to the identified challenges. This stage establishes a foundational understanding and identifies gaps in current knowledge.
- + Hypothesis Development: Based on the literature review, researchers propose hypotheses or sets of hypotheses suggesting potential solutions or explanations for the observed challenges. In technology and computer science, these hypotheses may pertain to algorithms, system architectures, or software design principles.
- + Experimental Design: Researchers meticulously design experiments, simulations, or studies aimed at scrutinizing the proposed hypotheses. This phase often involves prototyping software, implementing algorithms, or conducting simulations to accumulate data.
- + Variable Identification: Researchers discern and regulate variables that may influence experiment outcomes, ensuring controlled and systematic investigations.
- + Data Collection: Through experiments or simulations, researchers gather pertinent data germane to the research questions, which may encompass performance metrics, user feedback, or other measurable parameters.
- + Data Analysis: Researchers employ statistical methods and data analysis techniques to decipher the accumulated data, discerning patterns and trends to ascertain the validity of the hypotheses and draw substantive conclusions.
- + Interpretation and Conclusion: Researchers interpret the outcomes within the framework of the original hypotheses, evaluating whether the data corroborates or refutes the proposed explanations



or solutions.

- + Conclusion Formulation: Predicated on the analyses, researchers delineate conclusions and deliberate on the implications of their findings.
- + Peer Review and Dissemination: The research undergoes rigorous peer review, where subject matter experts scrutinize the methodology, outcomes, and conclusions. Subsequently, if accepted, the research is disseminated via scientific journals or conference proceedings, thereby enriching the collective knowledge pool.
- + Iteration and Enhancement: Researchers receive feedback from the scientific community, industry stakeholders, or end-users, which they utilize to refine hypotheses, methodologies, or technologies, thereby iterating the scientific method.
- + Application and Implementation: Successful research outcomes are frequently translated into real-world applications, often entailing the development of novel software, algorithms, or systems, thus fostering tangible advancements in technology and computer science.

4 Steps of the scientific method

The scientific method represents a systematic framework employed by scientists and researchers to scrutinize natural phenomena, formulate hypotheses, conduct experiments, and scrutinize outcomes. While minor variations may exist, the ensuing steps encapsulate the fundamental stages of the scientific method:

- + Observation: Researchers meticulously observe the natural world or a specific phenomenon, discerning intriguing patterns, occurrences, or quandaries that pique scientific curiosity.
- + Question Formulation: Grounded in observations, researchers craft specific, focused questions poised to elucidate or scrutinize phenomena under investigation, ensuring testability and clarity.
- + Hypothesis Formulation: Researchers postulate tentative explanations or answers to the research questions, known as hypotheses, which are rooted in extant knowledge and amenable to experimental verification.
- + Prediction Generation: Derived from hypotheses, researchers formulate explicit predictions or expectations that are quantifiable and facilitative of hypothesis validation.
- + Experimentation Design: Researchers meticulously devise experiments delineating variables, controls, and experimental protocols to orchestrate systematic and equitable tests of hypotheses.
- + Data Collection: Researchers meticulously execute experiments, methodically amassing data, and recording observations and measurements.



- + Data Analysis: Employing statistical methodologies or other pertinent techniques, researchers scrutinize amassed data, discerning patterns, trends, and relationships to ascertain the support or refutation of hypotheses.
- + Conclusion and Interpretation: Researchers interpret analytical outcomes vis-à-vis the original hypotheses, deliberating whether the data buttresses the hypotheses or warrants modifications.
- + Conclusion Drawing: Predicated on analyses, researchers draw conclusions pertaining to the research questions, cognizant of limitations and uncertainties inherent in the findings.
- + Communication: Researchers disseminate results via scientific papers, presentations, or other mediums, elucidating methodologies, data, and conclusions to the scientific community for scrutiny and validation.
- + Peer Review: The research is subjected to meticulous peer review, where experts in the field evaluate its validity, methodology, and conclusions.
- + Iteration and Revision: Cognizant of feedback from the scientific community, replication endeavors, or subsequent observations, researchers evaluate whether modifications to hypotheses, experimental designs, or interpretations are warranted.
- + Iteration: If necessitated, researchers embark on iterative cycles, refining research designs or delving into emergent queries stemming from findings.
- + Application: Validated findings find application in real-world scenarios, catalyzing further research and technological advancements.

These sequential steps underscore the cyclical and iterative nature of the scientific method, as researchers continually refine hypotheses, conduct additional experiments, and build upon existing knowledge. This systematic approach underpins the development of scientific theories and fosters the cumulative advancement of scientific understanding.

1.8. Research process

The research process often begins with a very broad idea for a topic you'd like to know more about. You do some preliminary research to identify a problem. After refining your research questions, you can layout the foundations of your research design, leading to a proposal that outlines your ideas and plans.





1. Choose Your Topic:

- + Begin by brainstorming ideas related to your area of interest. Your topic can start out broad.
- + Read widely to build background knowledge and narrow down your ideas.
- + Conduct an initial literature review to identify gaps, contradictions, and relevant sources.
- + Discuss your topic with your thesis supervisor to refine it further.

2. Identify a Problem:

- + Define a specific research problem within your chosen topic.
- + The problem could be practical (e.g., a process not working well) or theoretical (e.g., an underexplored phenomenon).
- + Write a problem statement to contextualize the issue, explain its significance, and outline how your research will contribute to solving it.

3. Formulate Research Questions:

- + Based on the problem, create clear and focused research questions.
- + These questions guide your investigation and help you gather relevant data.

4. Create a Research Design:

- + Methodology: Will you use qualitative, quantitative, or mixed methods?
- + Data Collection: How will you collect data (surveys, experiments, observations)?
- + Sampling: Define your sample size and selection criteria.
- + Data Analysis: Specify how you'll analyze the collected data.

5. Write a Research Proposal:

- + Summarize your research plan in a proposal.
- + Include the problem statement, research questions, methodology, and expected outcomes.
- + Seek feedback from peers or advisors to refine your proposal.

1.9.Population research

Population research is a scientific inquiry aimed at understanding the dynamics of a population. It encompasses various aspects related to population size, structure, growth, distribution, and dispersal.

Collecting data from a population

Collecting data from a population involves gathering information from every member of a specific group:

1. Population:



- + A population refers to the entire group you want to study or draw conclusions about.
- + It can include people, objects, events, organizations, species, or any other elements relevant to your research.
- + Collecting data from the entire population is straightforward when it is small, accessible, and cooperative.
- + Example: Analyzing the final exam scores of all graduating seniors in a high school to identify trends.

2. Sample:

- + A sample is a specific subset of the population from which you collect data.
- + The size of the sample is always smaller than the total population size.
- + Researchers use sampling when collecting data from larger or dispersed populations.
- + Example: Surveying 300 undergraduate students from three Dutch universities to explore psychology research questions.

3. Sampling Methods:

- + Random Sampling: Selecting a sample randomly to ensure representativeness.
- + Stratified Sampling: Dividing the population into subgroups (strata) and sampling from each subgroup.
- + Cluster Sampling: Randomly selecting clusters (e.g., schools, neighborhoods) and sampling within those clusters.
- + Convenience Sampling: Choosing participants based on convenience (e.g., surveying people nearby).
- + Purposive Sampling: Selecting participants intentionally based on specific criteria.

D. P3 CONDUCT PRIMARY AND SECONDARY RESEARCH USING APPROPRIATE METHODS FOR A COMPUTING RESEARCH PROJECT THAT CONSIDER COSTS, ACCESS AND ETHICAL ISSUES

1. Secondary Research

To gain a deeper understanding of the environmental impacts of Big Data and search for alternative materials and sustainable solutions for the future, I undertook a comprehensive secondary research effort. This study was undertaken with the aim of collecting and synthesizing existing information from a variety of authoritative and relevant academic sources, including journals, industry reports, official government statistics and reputable articles.



Date of publication: Prioritize recent sources as they provide insights into the most up-to-date developments and trends in Big Data, environments and alternative materials. However, historical perspectives are also considered if they provide valuable contextual information.

Reputable Sources: Sources from reputable publications, organizations, academic institutions are preferred to ensure that the information collected comes from trustworthy and reliable sources.

Author's professional qualifications: The author's professional qualifications and qualifications are evaluated. Information provided by experts in the field is considered more trustworthy.

Relevance to the research topic: Each source is evaluated for its direct relevance to the research topic of Big Data's impact on the environment and alternative materials. Irrelevant or tangential sources were excluded. By adhering to these strict selection criteria, I have aimed to compile a comprehensive and reliable body of information that will serve as a solid foundation for research on this important topic. Their diverse range of sources and collective expertise provided a comprehensive view of the topic, allowing for a thorough analysis of potential impacts and solutions related to Big Data, the environment and alternative materials.

1.1.Sources:

What is big data?

Link: Big data defined: Examples and benefits | google cloud (no date) Google.

Available at: https://aws.amazon.com/what-is/big-data/ (Accessed: 07 March 2023).

+ **Data collected:** big data signifies the vast, intricate collections of data that expand at an exponential rate, encompassing structured, semi-structured, and unstructured types. These datasets are characterized by their enormous size, rapid growth, and the diversity in the data forms they include, making them beyond the handling capacity of conventional data processing systems. The proliferation of digital advancements, such as enhanced connectivity, mobile technology, the Internet of Things (IoT), and artificial intelligence (AI), has been a significant driver in the surge of data generation and availability.

What impact does big data have on the environment?

Link: Lue, C. and Kimmy, K. (2022) Environmental sustainability in the age of big data: Opportunities and challenges for business and industry - environmental science and pollution research, SpringerLink.

Available at: Big Data Can Lead to Big Environmental Impacts (yale.edu) (Accessed: 02 March 2024).



- + **Data collected:** big data's role in environmental impact is multifaceted, offering both challenges and solutions. Here's a deeper look:
 - 1. Consumption of Resources:
 - Data Centers' Demand: Hosting and computing big data necessitates extensive energy and water, significantly impacting resources.
 - Operational Emissions and Usage: The operation of cooling systems, servers, and other infrastructural elements elevates carbon emissions and water consumption.
 - Designing Eco-friendly Data Centers: Prioritizing eco-conscious design in data centers is vital for reducing environmental footprints.

2. Generation of Electronic Waste:

- Obsolescence and Waste: The swift advancement in big data technology renders hardware outdated rapidly, escalating electronic waste.
- Recycling and Disposal Measures: Implementing effective recycling and disposal strategies for outdated electronics is critical for mitigating environmental damage.

3. Energy Utilization:

- Computing Demands: The intensive computing required for big data analytics amplifies energy consumption.
- Innovations in Efficiency: Adoption of energy-saving algorithms and hardware is key to diminishing energy use.

4. Privacy and Data Security:

- Safeguarding Data: Ensuring the security of big data is essential, requiring stringent security protocols.
- Risks of Data Breaches: Security breaches and unauthorized data access pose risks, potentially leading to environmental repercussions.

5. Monitoring the Environment:

- Enhanced Environmental Surveillance: Big data facilitates improved tracking of environmental parameters, such as air pollution levels and deforestation rates.
- Support for Environmental Sustainability: Utilizing big data for early threat detection and promoting eco-friendly practices is beneficial for environmental conservation.
- Alternative materials for big data to protect the environment.



Link: Abrantes, D. (2020) Environmental sustainability in the age of big data: Opportunities and challenges for business and industry - environmental science and pollution research, SpringerLink. Available at: Opportunities for big data in conservation and sustainability | Nature Communications (Accessed: 12 March 2023).

- + Data collected: To mitigate the environmental impact of big data, we can explore alternative materials and practices. Here are some considerations:
 - 1. Energy-Efficient Data Centers:
 - Instead of traditional data centers, consider using renewable energy sources (such as wind, hydro, or solar) to power data centers.
 - Optimize or upgrade technology to improve efficiency and operating temperature.
 - 2. Tackling Plastic Waste:
 - Beyond Big Data's Direct Impact: While indirectly related, combating plastic waste is vital
 for overall environmental health.
 - 3. Alternatives to Reduce Plastic Use:
 - Paper Solutions: An age-old, recyclable choice.
 - Polylactic Acid (PLA) Options: Made from renewable resources; requires industrial composting for degradation.
 - Plant-based Compostable Plastics: Offers a compostable solution, yet some types may release greenhouse gases during decomposition.
 - 4. Investigating Eco-friendly Materials:
 - Pioneering Research on Non-plastic Substitutes: Efforts are underway to find sustainable alternatives that can minimize the ecological impacts associated with big data infrastructure.

Interpretation and implications of the findings.

The findings from the exploration of big data's environmental impact reveal significant considerations across several domains:

- 1. Resource Utilization:
 - Data Center Demands: The operation of data centers for big data storage and processing is
 resource-intensive, significantly consuming energy and water. The emissions and resource
 use from cooling systems, servers, and other infrastructure underscore the necessity for ecofriendly data center designs to lessen environmental repercussions.
- 2. Generation of Electronic Waste:



 Technology Turnover: The swift evolution in big data technology precipitates a cycle of hardware obsolescence, escalating the generation of electronic waste. Effective management, including the recycling and responsible disposal of outdated equipment, is crucial to curtailing environmental degradation.

3. Consumption of Energy:

• Computational Intensity: The computation-heavy nature of big data analytics amplifies energy use. Implementing energy-efficient computing practices, such as optimizing algorithms and investing in low-power hardware, is pivotal in mitigating energy demands.

4. Energy Expenditure in Data Transmission:

 Network Load: The act of transmitting voluminous data sets across networks incurs significant energy consumption. Strategies like improving data compression and favoring localized processing can substantially reduce energy usage.

5. Data Privacy and Security:

 Cybersecurity Concerns: The imperative to safeguard big data from breaches is paramount, not only for data integrity but also to prevent environmental damage arising from data leaks or compromised systems.

6. Enhancements in Environmental Monitoring:

Ecosystem Surveillance: Leveraging big data enhances the monitoring of critical
environmental indicators such as air quality and deforestation rates. This capability is
instrumental in detecting ecological threats early and promoting practices that are aligned
with sustainability goals.

2. Primary research

The study framework, focusing on the environmental repercussions of big data and the exploration of eco-friendly alternatives, is carefully crafted for a thorough and effective data gathering and analysis process.

Interviews with Experts: Structured dialogues were held with professionals in environmental science and big data sectors. These discussions aimed to capture valuable perspectives, authoritative advice, and firsthand experiences concerning big data's environmental impacts and the exploration of sustainable alternatives. The selection of interviewees was based on their comprehensive expertise and professional background, ensuring a rich diversity of views and profound insights.



Expert Surveys: A specialized questionnaire was developed and disseminated among a carefully chosen group of specialists in environmental and big data disciplines. This survey sought to gather empirical data regarding the awareness of big data's environmental footprint, the efficacy of existing measures for environmental preservation, and the most urgent issues within the big data ecosystem. The formulation of the survey items was strategically done to draw out detailed information and feedback pertinent to the study's goals.

Ethical Guidelines: Adhering to ethical guidelines is paramount in this research to maintain the integrity of the research process and safeguard the rights and privacy of the participants. Key ethical principles

- Informed Consent: Before commencing data collection, detailed information about the study's aims, participant involvement, and data usage was provided to all participants. Consent was secured from participants ahead of their contribution.
- Anonymity and Confidentiality: To foster open and frank communication, the identities and inputs of participants are anonymized and treated with the utmost confidentiality. All collected data is securely stored and exclusively utilized for the purposes of this research.
- Voluntary Participation: Engagement in the study through interviews and surveys was entirely voluntary, with participants having the freedom to withdraw at any point without facing any repercussions.

2.1.Interview

The interview process for this study was designed to gather expert insights and practical experiences related to Big Data. Structured interviews will be conducted with experts in the field. The questions are designed to encourage detailed responses and facilitate a comprehensive understanding of the subject.

- Questions:

- 1. What is your perspective on the dynamic environments of big data processing and storage jobs?
- 2. How do you rate the detection of fractional environments from discontinuous large data expansion?
- 3. What solutions can help data centers operate more sustainably, reducing pressure on natural resources?



- 4. Do you know of any cases where data centers have been successful in adopting renewable energy?
- 5. In the growing landscape of big data, how can we solve the e-waste problem?
- 6. What role do energy-saving algorithms play in minimizing energy consumption for big data analysis?
- 7. What are the effects of larger data transfers over the network on energy consumption and what can we do to optimize this process?
- 8. What does data privacy and security in big data management mean for environmental protection?
- 9. How can big data be used to improve environmental monitoring and promptly detect threats to ecosystems?
- 10. What materials can be used to replace plastic components in data centers and electronic equipment, helping to reduce environmental impact?

2.2.Evaluation Survey

The goal of this survey is to gain a deeper understanding of Big Data experts' perspectives and knowledge. It is configured to collect computational and quantitative information, conduct corrections on implementation practices, facing formulas and emerging trends.

1. Survey Questions:

- Position and Experience

- What role in Big Data would you definitely get?
- How many years of experience do you have working with Big Data?

- Field and Organization

- To what extent does your field of work relate to Big Data?
- Has your organization established any policies or goals to minimize the impact of Big Data on the environment?

- Impact awareness

- Are you aware of the negative effects of Big Data on the environment? What tools do you know?
- Do you think these influences will impact your work and life? If so, what is the impact?

- Main Concern



- What issue are you most concerned about regarding the impact of Big Data on the environment?
- What solutions do you think are needed to solve this problem?

- Challenges in Finding Alternative Materials

- Have you ever looked for or used any alternative solutions to Big Data? If so, what material is it and how do you use it?
- Do you have any recipes when looking for or applying these alternatives?

- Effectiveness of alternative materials in environmental sanitation

- Do you believe that using alternative materials for Big Data can be effective in protecting the environment? If so, can you share any examples or evidence?
- What conditions do you think are needed to use alternative materials for Big Data to become more popular and widely applied?

E. P4 APPLY APPROPRIATE ANALYTICAL TOOLS, ANALYSE RESEARCH FINDINGS AND DATA

1. InterView

Moreover, the digital nature of the survey aligns with the contemporary digital landscape, enabling a seamless and efficient data collection process. Participants are empowered to respond at their own pace and convenience, removing any potential barriers associated with traditional paper based surveys. This ease of use not only enhances the likelihood of a higher response rate but also encourages participants to share their valuable experiences and insights openly.

The online survey platform also provides functionalities that aid in structuring and designing the survey in a manner that elicits meaningful and relevant responses. The incorporation of multiple question types, such as yes/no, scaled, and multiple-choice questions, allowed me to create a comprehensive set of inquiries. These questions were thoughtfully designed to explore various dimensions of my research topic, ensuring a well-rounded understanding of the subject matter. Additionally, the use of Google Forms facilitated real-time data collection, enabling me to monitor and analyze the responses as they were being submitted. This real-time monitoring proved immensely beneficial, allowing for prompt adjustments in the survey design if needed or identifying patterns and trends in the responses.

About the interview participants:



- Mr. Nguyen Hoang Quan | Director of BKAV Cyber Security Technology Joint Stock Company
- Mr. Le Van Lam | Leading expert in network security and big data, Director of Network and Information Security Company Limited (VNISA).
- Mr. Nguyen Van Luc | Director of Boba Cyber Company Limited
- Mr. Nguyen Manh Ha | Leading expert in big data and artificial intelligence (AI). He is Director of Artificial Intelligence Technology Application and Development Joint Stock Company (VICORP)
- Mr. Nguyen Ha Minh | Security expert, Viettel Cyber Security.

1.1.InterView 1

Name: Nguyen Hoang Quan

Age: 52

Occupation: Director at BKAV Cyber Security Technology Corporation

Company: BKAV Cyber Security Technology Corporation

1. The Environmental Impacts of Big Data and the Search for Sustainable Alternatives:

- A. Can you discuss the environmental effects of big data within its ecosystem?
- B. Big data's environmental footprint is noticeable through increased energy demands, heightened greenhouse gas emissions, the accumulation of electronic waste, and the diminution of biodiversity, all of which compromise ecosystem balance and sustainability.
- C. What obstacles do you encounter in reducing big data's environmental impact?
- D. The primary obstacles involve sourcing renewable energy, constructing green data centers, managing e-waste recycling and reuse, and ensuring biodiversity's preservation and rehabilitation, all while upholding big data's efficiency and integrity.

2. Sustainable Materials and Environmental Safeguards:

- A. How do you perceive the contribution of sustainable materials to diminishing big data's ecological impact?
- B. Sustainable materials are crucial in lessening big data's ecological footprint by lowering the consumption of virgin resources, minimizing energy use and waste, and enhancing material recovery and reuse, thereby aiding in environmental preservation and improvement.
- C. Could you illustrate the application of sustainable materials in big data settings for environmental enhancement?
- D. In big data realms, sustainable materials are innovatively utilized through the adoption of biodegradable, recycled, organic, and nanomaterials as substitutes for conventional substances like



metals, plastics, and silicon, offering energy-saving, emission-reducing, and sustainability-boosting solutions.

3. Innovative Technologies and Approaches for Sustainable Alternatives:

- A. What innovative technologies facilitate the exploration and adoption of sustainable alternatives in big data?
- B. Technologies like artificial intelligence (AI), machine learning (ML), the Internet of Things (IoT), cloud, and edge computing bolster the exploration and utilization of sustainable alternatives by enabling advanced data analysis, predictive modeling, process enhancement, and the fostering of collaborative ecosystems.
- C. How do these technologies enhance the effectiveness and overcome the challenges of sustainable materials in big data contexts?
- D. Leveraging big data, these technologies assist in identifying and implementing suitable sustainable materials, thereby reducing expenditure, time, and risk, while elevating the quality, dependability, and flexibility of sustainable alternatives in big data environments.

4. Developing an Enhanced System for Sustainable Materials:

- A. From your perspective, what strategies should organizations employ to cultivate an advanced system for sustainable materials in big data environments?
- B. Organizations should embrace a forward-thinking stance on sustainable materials by engaging in continuous research, assessment, and enhancement of these materials, adhering to environmental and safety standards, encouraging stakeholder participation, and nurturing a culture of creativity and innovation to improve the use and sustainability of sustainable materials in big data.
- C. Can you share any successful practices or methodologies for developing systems that effectively employ sustainable materials in big data scenarios?
- D. Embracing user-centric design principles, leveraging interdisciplinary approaches, forming strategic partnerships, and investing in necessary resources and infrastructure stand out as successful strategies for developing systems that better utilize sustainable materials in big data contexts.

1.2.InterView 2

- Name: Mr. Nguyen Van Luc

- Age: 43

- Occupation: Director at Boba Cyber Company Limited

- Company: Boba Cyber Company Limited



1. Understanding Big Data's Environmental Impact and Exploring Sustainable Alternatives

- A. Can you elaborate on your comprehension of big data and its environmental implications?
- B. My understanding is that big data encompasses vast and complex datasets which lead to significant environmental challenges, including high energy use, increased greenhouse gas emissions, and the proliferation of electronic waste.
- C. How committed are you to identifying and adopting sustainable alternatives to mitigate big data's environmental impact?
- D. My commitment to seeking and utilizing sustainable alternatives for big data is robust, as I view this as a crucial and urgent task for environmental preservation.

2. Sustainable Alternatives and Environmental Safeguards

- A. Are you aware of any sustainable alternatives to traditional big data practices, and could you mention a few?
- B. I am acquainted with various sustainable materials as alternatives in the big data sphere, such as using DNA, proteins, or other organic compounds for data storage and processing.

3. Innovative Technologies and Approaches for Sustainable Alternatives

- A. What innovative technologies and methods are you familiar with that facilitate the discovery and application of sustainable materials in big data?
- B. Several cutting-edge technologies and approaches are on my radar for supporting sustainable alternatives in big data, including quantum computing, artificial intelligence, and blockchain technology.
- C. What objectives do you have for employing these innovative technologies and methods in the context of sustainable materials for big data, and can you provide examples?
- D. My aim in leveraging these innovative solutions is to enhance big data's quality, security, and efficiency. Examples include employing advanced encryption for data security, sophisticated analytics for data interpretation, and efficient storage solutions to minimize environmental impact.

4. Developing an Enhanced System for Sustainable Materials in Big Data

A. Why do you believe it is essential and viable to establish a system for sustainable materials tailored to big data, and could you illustrate with examples?



B. Establishing a system for sustainable materials within the big data ecosystem is both necessary and achievable, addressing environmental, security, and efficiency concerns. For instance, adopting energy-efficient storage solutions, enhancing cybersecurity measures to safeguard data, and utilizing faster data processing technologies can all contribute to a more sustainable and efficient big data infrastructure.

1.3.InterView 3

Name: Mr. Nguyen Ha Minh

Age: 30

Occupation: Security Expert at Viettel Cyber Security

Company: Viettel

1. Exploring the Environmental Impact of Big Data and Sustainable Alternatives

A. Can you explain your understanding of big data and its environmental impacts?

- B. My grasp of big data involves extensive and intricate datasets that lead to environmental challenges like heightened energy usage, substantial greenhouse gas emissions, and significant generation of electronic waste.
- C. How important is the exploration and utilization of sustainable alternatives for big data in environmental conservation to you?
- D. My commitment to exploring and implementing sustainable alternatives for big data in environmental conservation is profound, recognizing it as a critical and immediate concern.

2. Sustainable Alternatives and Measures for Environmental Protection

- A. Could you describe your familiarity with the attributes of big data and some examples?
- B. I'm familiar with big data's attributes, such as its vast volume, diverse variety, and rapid velocity. Examples include the multifaceted sources of data, the different forms it can assume, and the swift generation and processing of data.
- C. What measures for environmental protection do you employ in your big data usage, and could you offer some instances?
- D. In my usage of big data, I adopt environmental protection measures like minimizing data traffic, enhancing computing efficiency, and advocating for the recycling of electronic devices.

3. Innovative Technologies and Approaches for Sustainable Alternatives

A. Are you aware of the characteristics of big data and can you list some?



- B. Understanding big data's nature involves recognizing its large volume, diversity, and quick processing. Instances include the variety of data sources, the array of data formats, and the speed at which data is produced and analyzed.
- D. How do you intend to use or are using emerging technologies and methods to find and apply sustainable alternatives for big data, and can you provide examples?
- C. My use or intended use of emerging technologies and methods aims at addressing big data challenges, refining solutions, and leveraging big data opportunities, with examples such as enhancing data compression, visualizing data more effectively, and facilitating data sharing.

4. Establishing a Superior System for Sustainable Materials in Big Data

- A. Why do you believe it is crucial and practical to create a system of sustainable materials for big data, and can you provide examples?
- B. Creating a system of sustainable materials for big data is deemed crucial and practical due to its potential to address environmental, security, and efficiency issues associated with big data. Examples include initiatives aimed at reducing the carbon footprint, bolstering data security, and optimizing data storage solutions.

1.4.InterView 4

Name: Le Van Lam

Age: 33

Occupation: Director at Network and Information Security Company Limited (VNISA)

Company: Network and Information Security Company Limited (VNISA)

- 1. Exploring Big Data's Environmental Impact and Sustainable Alternatives
- A. Can you explain your understanding of big data and its environmental effects?
- B. My grasp of big data encompasses the extensive and intricate datasets that lead to environmental issues by necessitating significant energy, generating substantial greenhouse gases, and contributing to electronic waste accumulation.
- C. How keen are you on exploring and implementing sustainable alternatives to alleviate big data's environmental impact?
- D. My interest in exploring and implementing sustainable alternatives for big data is profound, given the critical and immediate need for environmental preservation.



- 2. Sustainable Alternatives and Measures for Environmental Protection
- A. What insights do you have on big data's pros and cons, and could you list some?
- B. Recognizing big data's capabilities, I see its potential to deliver valuable insights, support diverse sectors, and generate economic benefits. However, it poses challenges including environmental strain, security vulnerabilities, and data integrity issues. For instance, while big data can enhance trend analysis, behavior prediction, and process efficiency, it also leads to high energy consumption, greenhouse gas emissions, and susceptibility to cyber threats.
- C. What specific measures for environmental conservation do you adopt in your big data utilization, and can you provide examples?
- D. In my big data practices, I prioritize environmental conservation through actions like reducing data transmission, enhancing system efficiency, and promoting the recycling of electronic components.
- 3. Leveraging Emerging Technologies for Sustainable Materials
- A. Can you discuss the benefits and drawbacks of big data, including examples of both?
- B. My understanding acknowledges big data's dual nature, offering significant informational, cross-sectoral support, and economic growth opportunities, while also presenting environmental, security, and quality dilemmas. For instance, big data facilitates advanced analytics and optimizations but raises concerns over energy use, greenhouse emissions, and cybersecurity risks.
- C. How do you intend to use or currently use emerging technologies to find and apply sustainable materials in big data, and can you share examples?
- D. I aim to or currently utilize emerging technologies to address big data's drawbacks, capitalize on its benefits, and innovate new applications. Examples include employing data encryption for security, machine learning for deeper insights, and developing platforms for data exchange.
- 4. Establishing a System for Sustainable Materials in Big Data
- A. Why do you believe it's both necessary and possible to create a system for sustainable materials in big data, and can you offer examples?
- B. I believe in the necessity and feasibility of a sustainable materials system for big data because it addresses critical issues like environmental degradation, security threats, and operational efficiency. Examples include energy-saving initiatives, enhanced cybersecurity measures, and improved data processing techniques, all contributing to a more sustainable, secure, and efficient big data ecosystem.

1.5.InterView

Name: Nguyen Manh Ha



Age: 41

Occupation: Director at VICORP, specializing in Artificial Intelligence Technology Application and Development

Company: VICORP - A leader in Artificial Intelligence Technology Application and Development

1. Understanding Big Data and Its Environmental Footprint

- A. Could you explain your comprehension of big data and its environmental repercussions?
- B. My understanding encompasses big data as extensive and intricate datasets that contribute to heightened energy consumption, increased greenhouse gas emissions, and a surge in electronic waste.
- C. How significant do you find the exploration and application of sustainable alternatives for big data in environmental preservation?
- D. My commitment to identifying and employing sustainable alternatives for big data is substantial, driven by the critical and immediate necessity to safeguard the environment.

2. Sustainable Alternatives and Conservation Efforts

- A. Could you detail your grasp of big data and its characteristics, perhaps mentioning a few?
- B. Recognizing big data as voluminous and multifaceted, its notable characteristics include its vast volume, diverse variety, and rapid velocity, such as its enormous size, the variety of data types it encompasses, and its swift generation and processing speeds.
- C. What specific environmental conservation practices do you adopt in your big data operations, and could you provide examples?
- D. In managing big data, I implement several conservation measures like reducing data transfer volumes, enhancing computing efficiency, and advocating for the recycling of electronic components.

3. Utilizing Emerging Technologies for Sustainable Solutions

- A. Could you reaffirm your understanding of big data and its key characteristics, possibly elaborating on a few?
- B. Big data, in my view, is characterized by its sheer volume, diverse content, and the speed at which it is produced and processed, indicating its expansive capacity, varied forms, and rapid generation and analysis.



- C. For what purposes do you engage or plan to engage with emerging technologies and solutions to find and use sustainable alternatives for big data, and can you offer examples?
- D. My aim in leveraging emerging technologies and solutions is to address big data's challenges, fulfill its demands, and drive innovation, with practical applications like enhancing data encryption, utilizing machine learning for data analysis, and promoting efficient data sharing.

4. Developing Systems for Sustainable Materials in Big Data

- A. Why do you believe in the necessity and feasibility of creating systems for sustainable materials tailored to big data, and could you cite some reasons?
- B. I advocate for the development of sustainable material systems for big data because it addresses critical issues of environmental impact, cybersecurity, and operational efficiency. Examples include endeavors to lessen energy reliance, enhance data protection against cyber threats, and elevate data processing quality, illustrating the vital role of sustainable systems in the modern data landscape.

2. InterView Summary

Nguyen Hoang Quan:

The objective was to gather insights on the interviewee's understanding of big data, alternative materials, and innovative technologies. The interviewee is well-versed in big data's environmental impact and highly values the exploration and utilization of alternative materials for environmental protection. They are familiar with various alternative materials like DNA, proteins, or organic compounds, alongside emerging technologies such as quantum computing, AI, and blockchain. They advocate for the use of these technologies to enhance big data's quality, safety, and efficiency, emphasizing the necessity and viability of developing alternative material systems to address big data's environmental, security, and efficiency challenges.

Le Van Lam:

Participants demonstrate a deep understanding of big data and its environmental consequences, showing a strong interest in seeking and applying alternative materials for environmental safeguarding. They have knowledge of numerous alternative materials and the supportive role of emerging technologies like quantum computing, AI, and blockchain. The focus is on using these advancements to boost big data's effectiveness and security, underscoring the importance of creating systems for alternative materials to resolve big data's pivotal issues.

Nguyen Van Luc:



The survey disclosed the respondents' comprehension of big data, its adverse environmental impacts, and their enthusiasm for discovering and applying alternative materials for environmental protection. They are informed about various alternative materials and the supportive emerging technologies, aiming to use these innovations to tackle big data challenges, enhance solutions, and seize opportunities. The consensus is that establishing systems for alternative materials is essential and practical to overcome big data's environmental, security, and efficiency dilemmas.

Nguyen Ha Minh:

Respondents are aware of big data's implications on the environment and show significant interest in alternative materials for ecological conservation. They possess knowledge of several alternative materials and the supporting role of new technologies and solutions. Their aim is to leverage these technologies to navigate big data's challenges, enhance solutions, and explore new possibilities, highlighting the critical need for alternative material systems to address the environmental, security, and efficiency issues associated with big data.

Nguyen Manh Ha:

Survey participants understand big data's impact on the environment and express a high level of interest in identifying and utilizing alternative materials for environmental protection. They are acquainted with various alternative materials and the assistance of emerging technologies. Their intent is to employ these technologies to fulfill requirements, address needs, and innovate within the big data landscape, affirming the necessity and feasibility of developing alternative material systems to rectify the environmental, security, and efficiency problems inherent in big data.

3. Survey

In my research, I prioritized obtaining direct insights and data through primary research methods, with a particular focus on deploying an online survey created via Google Forms. This method was chosen for its precision in collecting accurate and reliable data directly from respondents. Utilizing an online survey facilitated by Google Forms offered several advantages for data gathering, primarily due to its accessibility and user-friendliness, allowing respondents from various locations and backgrounds to participate effortlessly.

The ability for participants to access the survey from any internet-connected location significantly broadened the potential respondent base, ensuring a diverse and inclusive sample. This diversity is crucial for acquiring a comprehensive understanding of the subject matter from multiple perspectives.



The digital format of the survey aligns with the current digital era, making the data collection process smooth and efficient. It afforded participants the flexibility to respond at their convenience, potentially increasing the response rate and encouraging more open and thorough sharing of insights.

Google Forms supported the creation of a well-structured survey by offering various question types, including yes/no, scaled responses, and multiple-choice options. This flexibility was instrumental in formulating questions that delve into different aspects of the research topic, aiming for a thorough exploration of the subject. The platform's capability for real-time data collection was particularly advantageous, as it allowed for immediate analysis and, if necessary, swift modifications to the survey based on emerging trends or feedback.

The survey was conducted over a period, starting from February 20, 2023, and concluding on March 08, 2023. This timeframe was strategically chosen to allow ample opportunity for participants to engage with the survey, ensuring a robust and informative data set for analysis.

Survey Duration:

Start Date: July 12, 2022.

End Date: October 10, 2022.

How do you rate your level of understanding about big data?

Сору

11 responses

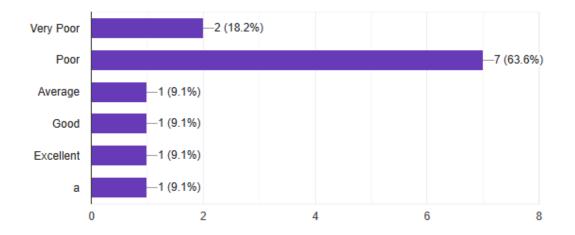


Figure 4: Survey 1

It's Surprising That There Are a Large Number of Users Who Don't Understand Big Data



Do you think big data has an impact on the environment?"

11 responses

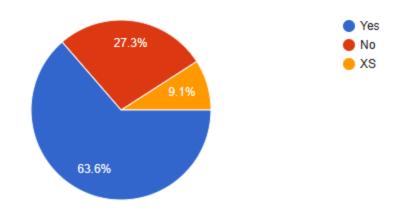


Figure 5: Survey

Most participants think big data has an impact on the environment.

Do your relatives know much about big data?

10 responses

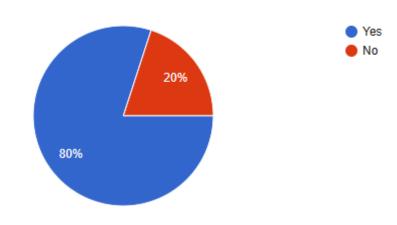


Figure 6: Survey

There are many relatives of survey participants who have knowledge of big data up to 80%



Do you think it is easy to find substitutes for big data?

10 responses

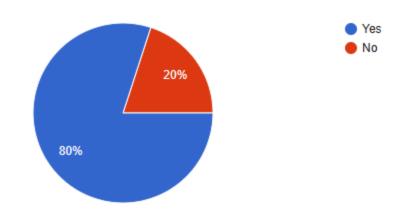


Figure 7: Survey

There are a majority of users who think that it is easy to find substitutes for big data.

There are 5 levels, to what extent do you think big data affects the environment?

10 responses

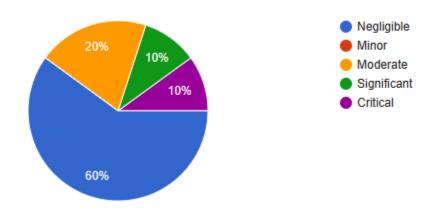


Figure 8: Survey

There are a majority of users who think that big data affects the environment at a negligible level.

4. Survey Summary

- Based on the survey charts you provided, here's a summary:

Formula about big data (Big Data): Personal Architecture: 30% of survey participants rated their knowledge of big data as average to excellent, with 63.6% saying they have intermediate knowledge of the field.

Family knowledge: 80% of participants said their relatives do not know much about big data.



- Operations of big data to the environmental field:

Perception of impact: 90% of participants believe data has a major impact on the environment. Of these, 60% perceive this impact as moderate, 20% perceive the impact as significant, and 10% perceive the impact as small or insignificant.

Assessment of impact level: Ability to analyze the level of activity of larger data on the environment, reflecting differences in individual perspectives.

- Using and replacing big data:

Usage in life and work: A large number of participants are aware that their work uses big data regularly, even though it is not allowed.

- Survey tools in charts.

Ability to find alternatives: 80% of participants said that it is not easy to find alternatives to big data, which shows the dependence on big data in current society grand.

- Development of big data in the future:

Outlook: There is a certain club about developing the potential of big data in the future, but there are no tools that can survey this problem.

- In summary, although there is some awareness about the impact of big data on the environment and difficulty in finding these types of alternatives, general knowledge about big data and its popularity is lacking. In the family, the family is still low. This may indicate that more education and information is needed so that people can better understand big data and how it affects their daily lives and the environment around them.

4.1. Analyze the results of the primary research.

Delving into initial research on the environmental ramifications of Big Data and the pursuit of sustainable alternatives, we undertake a synthesis of collected data, integrating both qualitative discourse and quantitative analysis to establish a comprehensive picture of the current landscape, discernible trends, and actionable strategies.

1. Overview of Methods:

- + Qualitative Dialogues: Engagements with informed individuals through qualitative discourse provide nuanced perspectives on the subject. Conversations with professionals or stakeholders entrenched in Big Data offer subjective insights into the nuanced challenges and opportunities within the field.
- + Quantitative Analysis: The execution of a broad-based quantitative analysis assists in identifying prevailing trends and measuring the environmental footprint of Big Data, enabling a data-driven approach to understanding its scope and effects.

2. Principal Observations:

+ Energy Utilization: The operation of Big Data necessitates a significant electrical input, predominantly for the upkeep of data repositories, which includes perpetual cooling, power management, and equipment maintenance, thereby intensifying carbon footprints and the strain on energy provisions.



- + Resource Management: The infrastructural expansion needed for Big Data escalates the consumption of varied resources like metals, plastics, and water, which in turn amplifies waste generation and resource depletion.
- + Data Volume Escalation: The exponential growth in data generation demands increased collection mechanisms, leading to the propagation of sensors and web-connected apparatuses, with their production and distribution contributing to emissions and potential wastage.
- + Confidentiality and Safeguarding: Handling Big Data, often laden with personal details, necessitates stringent security protocols, which can lead to higher energy and resource expenditure due to encryption and surveillance systems.

3. Suggested Resolutions:

- + Computational and Infrastructural Innovations: Exploration into novel substances and designs for the development of Big Data's computational framework is recommended, such as the utilization of materials that enhance energy efficiency or diminish the reliance on rare metals, alongside the search for eco-friendly alternatives for core components.
- + Data Conservation Advances: Investigations into pioneering data conservation substances are underway, aiming to identify materials with superior storage capabilities, heightened energy efficiency, and minimized ecological footprints, with DNA-based storage systems being a notable area of interest.
- + Data Handling Enhancements: Strategies are being developed for more energy-conscious data management, such as state-of-the-art data compression techniques and the adoption of distributed and cloud computing to optimize resource allocation and curtail energy demands.
- + Renewable Energy Integration: Big Data systems' energy demands are increasingly being met through renewable sources, incorporating solar, wind, hydro, or biomass energies to power operations.
- + Operational Optimization: Continuous search for methodologies that enhance the efficiency of Big Data operations, utilizing advanced algorithms and analytics to streamline resource consumption and decrease energy requirements.

F. P5 COMMUNICATE RESEARCH OUTCOMES IN AN APPROPRIATE MANNER FOR THE INTENDED AUDIENCE.

1. Conclusion

In the digital age, the prominence of big data is undeniable, yet its management and storage exert considerable environmental pressures, including significant energy use, greenhouse gas emissions, and the generation of electronic waste. To mitigate these adverse effects, it is crucial to enhance data center efficiency, incorporate renewable energy sources, and explore sustainable alternatives for electronic components. Implementing these strategies not only aids in safeguarding the environment but also paves the way for novel business and technological advancements for both companies and individuals.



2. Recommendations

The term big data encompasses the burgeoning volume of information generated, stored, and processed by electronic devices, offering numerous societal advantages but also presenting environmental challenges and the need for sustainable material alternatives. Notable environmental concerns include:

- Elevated Energy Demand: The operation and cooling of extensive data centers consume vast amounts of electricity, contributing to greenhouse gas emissions and climate change.
- Electronic Waste Accumulation: Frequent updates in digital technology lead to rapid obsolescence of devices, resulting in significant electronic waste that harbors hazardous substances and presents recycling challenges.
- Scarcity of Materials: The manufacture of electronic devices relies on scarce resources like gold, silver, platinum, cobalt, lithium, and other rare metals, which face depletion risks over time.
- ♣ Addressing the environmental footprint of big data and the quest for sustainable materials necessitates innovative solutions, such as:
 - Enhancing Energy Efficiency: Data centers should adopt energy-saving technologies and shift towards renewable energy sources like solar, wind, and hydro power.
 - Promoting Recycling and Reusability: Encouraging the recycling and reusability of digital devices can help diminish e-waste and conserve materials. Initiatives like equipment return and refurbishment programs are beneficial.
 - Advancing Material Research: There is a critical need for research into new materials that match or surpass the performance of current materials but are more abundant and environmentally friendly. Approaches could include nanotechnology, biotechnology, or synthetic alternatives.
 - Advancing Material Research: There is a critical need for research into new materials that match or surpass the performance of current materials but are more abundant and environmentally friendly. Approaches could include nanotechnology, biotechnology, or synthetic alternatives.
- ♣ These recommendations aim to address "the impact of big data on the environment and the exploration of alternative materials," offering insights that may prove valuable for further discussion and implementation.

G. CONCLUSION

The paper delves into the environmental consequences of big data and the pursuit of sustainable alternatives within the realm of Big Data Environment. It outlines significant challenges such as the demand for energy, pollution from emissions, and the need for resource preservation, while advocating for inventive approaches



to address these issues. Central to its recommendations are the adoption of eco-friendly, recycled, nanotechnology-based, and intelligent materials as substitutes for conventional ones, aimed at waste minimization, energy conservation, pollution reduction, and enhanced operational efficiency. It suggests implementing scalable distributed architectures, effective data storage solutions, and fostering an organizational ethos that prioritizes environmental stewardship. The overarching theme of the paper is the imperative for adopting a sustainable and strategic framework to navigate the complexities of an evergrowing data-centric world.

H. REFERENCE

- Secondary data in research methodology (with examples) | indeed.com
 Available at: Secondary Data In Research Methodology (With Examples) | Indeed.com India (Accessed: 01 March 2024).
- 2. A beginner's guide to starting the research process Scriber.

 Available at: A Beginner's Guide to Starting the Research Process (scribbr.com) (Accessed: 02 March 2024).
- 3. Bhandari, P. (2023) What is qualitative research? Methods & examples, Scriber.

 Available at: What Is Qualitative Research? | Methods & Examples (scribbr.com) (Accessed: 02 March 2024).
- 4. Big data defined: Examples and benefits | google cloud Google.

 Available at: What Is Qualitative Research? | Methods & Examples (google.com) (Accessed: 02 March 2024).
- 5. George, T. (2024a) Primary research: Definition, types, & examples, Scriber.

 Available at: Primary Research | Definition, Types, & Examples (scribbr.com) (Accessed: 02 March 2024).
- 6. George, T. (2024b) What is secondary research? Definition, types, & examples, Scriber.

 Available at: What is Secondary Research? | Definition, Types, & Examples (scribbr.com) (Accessed: 02 March 2024).
- 7. Li, C. and Huang, M. (2023) Environmental sustainability in the age of big data: Opportunities and challenges for business and industry environmental science and pollution research, SpringerLink.

 Available at: Environmental Sustainability in the Age of Big Data: Opportunities and Challenges for Business and Industry | Environmental Science and Pollution Research (springer.com) (Accessed: 02 March 2024).