SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Innovative Cost-Efficient Cloud Computing-Based Models for Disasters Management

Ahmed Abdelaziz[1,*], Saleh Mesbah[1], Mohamed Kholief[1]

[1] College of Computing and Information Technology, Arab Academy for Science, Technology and Maritime Transport, Alexandria 21532, Egypt

ARTICLE INFO

Article history:
Received 27 October 2023
Received in revised form 13 April 2024
Accepted 6 June 2024
Available online 5 July 2024

Keywords:
Cloud computing; disaster management models; low-cost disaster management models

ABSTRACT

Cloud computing has transformed the digital landscape, offering scalable services to individuals and businesses. However, ensuring continuous cloud service availability requires robust disaster management. For instance, in case of strikes of natural disasters, a fully functioning cloud landscape will collapse, which leads to substantial loss in terms of time, effort, and monetary aspects. This research paper explores current cloud computing solutions, emphasizing the importance of disaster management, and introduces two innovative models for selecting potential backup sites. The study begins with a comprehensive review of existing cloud computing solutions and their disaster management mechanisms. Evaluating their strengths and limitations, However, the current disaster recovery (DR) solutions are costly since they demand permanent contracts with cloud providers to pre-assign constant DR locations as replicas of the primary landscapes. To minimize this cost, we stress the urgent need for cost-effective disaster recovery strategies. This is accomplished by developing two models considering the most influential factors that contribute to DR site selection. The Weighted Grid Decision Model (WGDM) combines geographical and environmental attributes to evaluate the desirability of candidate sites. This structured approach allows for informed decision-making. The second model, the Artificial Neural Network (ANN) model, leverages machine learning to analyse historical data on disaster incidents and their effects on cloud infrastructure. By identifying patterns and trends, the ANN model assists in making intelligent backup site choices. This research demonstrates the benefits of employing AI-driven decision-making tools for disaster management in cloud computing.

1. Introduction

Cloud computing has emerged as a transformative force in the field of information technology, revolutionizing the way businesses and individuals' access and manage their data and applications. The term "cloud computing" refers to the practice of storing and accessing data and programs over the internet, rather than on local servers or personal computers. This paradigm shift in computing has unlocked unprecedented levels of scalability, flexibility, and cost-efficiency, offering an array of

* *Corresponding author.*
E-mail address: abadaway@gmail.com

https://doi.org/10.37934/araset.48.1.100116

services that have reshaped the modern digital landscape. Companies like Amazon, Google, and Microsoft played pivotal roles in popularizing the cloud computing model, introducing services such as Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure, respectively [1-3].

However, cloud computing is susceptible to potential service interruptions. Therefore, disaster management is crucial in cloud computing to ensure service continuity and data protection amid various risks such as hardware failures, natural disasters, cyberattacks, and human errors. Key strategies include redundancy through data and service replication across multiple geographically dispersed data centers, guaranteeing seamless shifting in case of failure. Robust backup and recovery mechanisms involve regular automated backups stored securely to prevent data loss or corruption. Comprehensive security measures encompass encryption, access controls, and threat detection to counter cyber threats and unauthorized access.

The goal is to ensure that DR plans are adaptable and responsive to specific disaster scenarios, rather than being tied to long-term contracts. This might be applicable by automating the DR site selection process. By optimizing costs, organizations can improve the effectiveness of their DR plans and minimize unnecessary expenses.

Therefore, this research paper aims to address the challenge of automating DR site selection by developing backup selection site models to optimize disaster recovery strategies. The focus will be on identifying the most important factors that influence the choice of backup location, taking into account various considerations such as geographical proximity, data redundancy, security, and compliance requirements. Two distinct models are proposed: a grid decision model and an artificial neural network. The grid decision model will utilize a systematic approach to evaluate different backup locations based on predefined criteria, offering a structured and objective decision-making process. On the other hand, the artificial neural network will leverage machine learning techniques to identify patterns and correlations among factors, enabling data-driven backup location selection.

The remainder of this paper is organized into several sections to present a comprehensive investigation into backup site selection criteria and model development for disaster recovery in cloud computing environments. Section 2 delves into the background, providing context and relevant literature on disaster recovery and cloud computing. In Section 3, the focus shifts to the crucial aspect of backup site selection criteria, where various factors influencing the choice of backup location are explored in detail. Section 4 outlines the methodology employed for data collection and processing, highlighting the sources of data and the techniques used to ensure data accuracy and reliability. Moving forward, Section 5 elaborates on the development of two distinct models: the grid decision model and the artificial neural network, explaining their respective methodologies and approaches. Finally, Section 6 presents the conclusions drawn from the study, summarizing the findings and discussing their implications.

2. Background

This section briefly reviews the main key concepts related to disaster management, GIS, IoT, and cloud computing.

2.1 Disaster Types

Disasters can be broadly categorized into several types based on their origin and impact. These events can cause significant damage to communities, economies, and the environment, resulting in loss of life, destruction of property, and disruption of essential services. Understanding the different

types of disasters is crucial for preparedness, response, and recovery efforts. Detailed information about disaster types can be found in the studies by Shaluf *et al.,* [4] and Abdelaziz and Mesbah [5].

Natural disasters are among the most common and devastating types of disasters. They result from natural processes and forces of nature and are beyond human control. Earthquakes, for instance, occur due to the movement of tectonic plates, resulting in sudden shaking of the Earth's surface. Hurricanes, cyclones, and typhoons are intense tropical storms characterized by strong winds and heavy rainfall, often causing widespread destruction in their paths. Floods occur when water overflows onto normally dry land, and they can be triggered by heavy rainfall, storm surges, or dam failures. Wildfires, on the other hand, are uncontrolled fires that spread rapidly through forests, grasslands, or even urban areas. Tsunamis, caused by underwater earthquakes, volcanic eruptions, or landslides, lead to large ocean waves that can devastate coastal regions. Lastly, volcanic eruptions involve the outpouring of molten lava, ash, and gases from a volcano, posing risks to nearby populations [6,7].

Technological disasters are another category of disasters, resulting from human-made processes and often involving complex technological systems. Industrial accidents can lead to chemical spills, explosions, or equipment failures in industrial facilities, causing severe environmental and human health impacts. Nuclear accidents involve radioactive releases from nuclear power plants or nuclear weapons testing, with potentially long-lasting consequences for the affected areas. Cyber-attacks are also considered technological disasters, wherein malicious individuals or groups exploit computer systems, networks, or critical infrastructure, leading to significant disruptions and potential data breaches [8,9].

Biological disasters are characterized by the outbreak or spread of infectious diseases that can cause widespread illness and mortality. Pandemics and epidemics fall under this category, and they can have far-reaching global consequences. One prominent example is the COVID-19 pandemic, which led to a worldwide health crisis and major socio-economic disruptions [10].

Environmental disasters result from human activities that cause extensive damage to the environment, ecosystems, and natural resources. Deforestation, the large-scale removal of forests, leads to the loss of biodiversity and disrupts the ecological balance, impacting climate and weather patterns. Pollution, including air, water, and soil pollution, arises from industrial, agricultural, and urban activities, resulting in environmental degradation and adverse effects on human health and wildlife. Climate change, primarily driven by greenhouse gas emissions, is a long-term environmental disaster that alters global weather patterns, leading to rising temperatures, extreme weather events, and sea-level rise [11].

Man-made disasters are those that result from human actions or negligence, often with unintended consequences. Terrorism, involving deliberate acts of violence and intimidation to achieve political, religious, or ideological goals, poses significant security challenges. Civil unrest, such as riots, protests, or social conflicts, can escalate into violent situations, impacting societal stability and safety. Additionally, infrastructure failures, including the collapse of bridges, dams, or buildings due to poor maintenance or construction practices, can lead to catastrophic consequences [12].

Each type of disaster presents unique challenges and requires specific strategies for preparedness, response, and recovery. Governments, organizations, and communities worldwide work collaboratively to develop disaster management plans and implement measures to mitigate the impact of these events, protect lives, and preserve property and infrastructure.

2.2 IoT and GIS Technologies

The Internet of Things (IoT) and Geographic Information Systems (GIS) are two cutting-edge technologies that have revolutionized disaster management practices, enhancing preparedness, response, and recovery efforts.

IoT, a network of interconnected devices and sensors, plays a pivotal role in disaster management by providing real-time data and insights [13,14]. During disasters, IoT devices can be deployed in various locations to monitor critical factors such as temperature, humidity, air quality, water levels, and structural integrity. These devices can also track the movement of people and assets, allowing authorities to assess the situation and make informed decisions promptly. For instance, in the case of wildfires, IoT sensors can detect changes in temperature and smoke levels, alerting authorities to potential fire outbreaks and enabling quicker responses. Similarly, during floods, IoT-based water level sensors can monitor rising water levels in rivers and lakes, helping authorities issue timely warnings and evacuation orders. By harnessing IoT data, disaster management teams can gain a comprehensive understanding of the situation on the ground and allocate resources effectively to minimize the impact of disasters on communities and infrastructures [15].

GIS technology is instrumental in spatially analyzing and visualizing disaster-related data [16]. GIS allows disaster management professionals to integrate diverse datasets, such as topography, land use, infrastructure, population distribution, and hazard mapping, into a geospatial context. This enables them to identify vulnerable areas, assess potential risks, and create evacuation plans accordingly. GIS also facilitates the coordination of emergency responses by providing real-time maps that display critical information about the disaster's impact and the location of affected populations and resources. First responders can use GIS to optimize their routes to reach affected areas quickly and efficiently. Moreover, GIS aids in resource management during recovery operations, helping to prioritize relief efforts based on the severity and extent of damage in different regions. By using GIS technology, disaster management teams can make data-driven decisions, allocate resources strategically, and improve overall situational awareness to better respond to disasters and support affected communities [17].

The integration of IoT and GIS technologies in disaster management has paved the way for a more efficient and effective approach to handling crises. These technologies facilitate the collection and analysis of real-time data, enabling faster and more accurate decision-making during critical situations. As the IoT continues to expand and the capabilities of GIS evolve, disaster management practices will continue to benefit from the innovative solutions and insights these technologies offer, ultimately contributing to a safer and more resilient world [15,17].

2.3 Existing Cloud Solutions

Cloud computing has evolved into a diverse ecosystem of solutions, offering a range of services catering to different business needs and requirements. Among the major cloud providers, Google Cloud Platform (GCP), Microsoft Azure, and Amazon Web Services (AWS) stand out as the dominant players in the market, each offering a comprehensive set of services to cater to various industries and workloads [18-20].

GCP provides a robust and scalable infrastructure, leveraging Google's extensive global network and data centers [18]. GCP offers a variety of services, including computing resources through Google Compute Engine, managed Kubernetes with Google Kubernetes Engine (GKE), and serverless computing with Google Cloud Functions. Google's data analytics solutions, such as BigQuery, enable businesses to process and analyze vast amounts of data efficiently. Additionally, GCP provides

machine learning capabilities through its AI Platform, allowing organizations to build and deploy AI models for various applications. Google's expertise in data processing, machine learning, and analytics makes GCP particularly attractive for businesses focused on data-driven insights and advanced AI solutions.

Microsoft Azure, backed by Microsoft's vast experience in enterprise technologies, offers a comprehensive suite of cloud services for businesses of all sizes [19]. Azure's infrastructure services, such as Azure Virtual Machines and Azure App Service, provide flexible and scalable computing resources. With Azure Kubernetes Service (AKS), Microsoft enables easy deployment and management of containerized applications. Azure's strength lies in its seamless integration with Microsoft's existing software ecosystem, making it a preferred choice for organizations already using Microsoft products. Furthermore, Azure's Data and AI services, such as Azure SQL Database and Azure Cognitive Services, empower businesses with powerful data management and machine learning capabilities. Microsoft's strong focus on hybrid cloud solutions and compliance features makes Azure a popular choice for industries with stringent regulatory requirements.

AWS is the pioneer and one of the most widely adopted cloud computing platforms globally [20]. AWS offers a vast array of services, including Amazon Elastic Compute Cloud (EC2) for scalable computing, Amazon Simple Storage Service (S3) for secure and durable object storage, and AWS Lambda for serverless computing. AWS's global presence with multiple data centers enables organizations to deploy applications closer to end-users for reduced latency. The platform's extensive set of AI and machine learning services, such as Amazon SageMaker and Amazon Rekognition, empowers businesses to leverage AI capabilities easily. AWS's vast customer base, extensive partner network, and wide range of services make it a compelling choice for businesses of all sizes and industries.

Other cloud providers, such as IBM Cloud, Oracle Cloud, and Alibaba Cloud, also offer competitive cloud solutions catering to specific industry verticals and use cases [21-23]. The cloud computing landscape continues to evolve rapidly, with providers constantly expanding their service offerings and enhancing existing features to meet the growing demands of businesses worldwide.

3. Backup Location Selection Criteria

Selecting the appropriate backup location in cloud computing is a crucial decision that can significantly impact data security, accessibility, and disaster recovery capabilities. Cloud computing backup location selection criteria revolve around a combination of factors, ranging from geographical considerations to compliance requirements and data redundancy. Organizations must carefully evaluate these criteria to ensure that their critical data is securely stored, readily accessible, and well-protected from potential disasters or disruptions. By making informed decisions based on the specific needs of their business, they can maximize the benefits of cloud-based backups while minimizing potential risks and vulnerabilities [24-26].

3.1 Location and Geographic Separation

The geographic location of the disaster recovery site is a crucial factor in mitigating risks associated with regional disasters. Research indicates that a considerable distance between the primary data center and the disaster recovery site is essential for reducing the likelihood of simultaneous disruptions due to natural calamities. Moreover, the location should also consider factors such as access to reliable utilities and transportation infrastructure, as these can impact the site's operational efficiency during recovery efforts. Disaster recovery sites with proximity to major

transportation hubs and redundant utility connections exhibited faster recovery times and reduced logistical challenges during emergencies. Therefore, geographical separation and accessibility are key considerations when selecting an optimal disaster recovery site for cloud-based data centers [24].

3.2 Redundancy and Reliability

The level of redundancy and reliability of the disaster recovery site is directly tied to its ability to provide seamless failover during a disaster. Disaster recovery sites equipped with redundant power sources, networking equipment, and data storage systems significantly reduce the risk of single points of failure. Ensuring high levels of reliability also involves robust environmental controls to prevent overheating and humidity-related issues. Furthermore, disaster recovery sites equipped with advanced cooling systems and environmental monitoring exhibited improved uptime and reduced equipment failure rates. Reliability can be strengthened by employing diversified network carriers to avoid a single carrier's network outage. By prioritizing redundancy and reliability, organizations can ensure continuous access to critical data and services during times of crisis [27].

3.3 Security Measures

The security of the disaster recovery site is of paramount importance in safeguarding sensitive data and protecting against unauthorized access. Comprehensive physical security measures, including access controls, surveillance cameras, and biometric authentication, are essential to thwart potential security breaches. Encryption plays a vital role in securing data in transit and at rest. Disaster recovery sites that employ strong encryption algorithms for data replication and storage offer an additional layer of protection against data breaches and unauthorized access. Regular security audits and vulnerability assessments are also crucial to identifying and mitigating potential weaknesses in the disaster recovery site's security infrastructure. By prioritizing security measures, organizations can bolster their data center's resilience and safeguard critical assets against cyber threats [28].

3.4 Network Connectivity and Bandwidth

A reliable and robust network connectivity is essential for seamless data replication and synchronization between the primary data center and the disaster recovery site. High-bandwidth connections with low latency ensure minimal data transmission delays during failover operations. Organizations should consider leveraging dedicated and redundant network links to ensure continuous data accessibility during disaster scenarios. Moreover, the disaster recovery site should have ample network capacity to accommodate increased workloads during recovery periods. Over-provisioning network bandwidth at the disaster recovery site can prevent bottlenecks and ensure optimal performance during the transition from the primary data center to the recovery environment.

3.5 Recovery Time Objective (RTO) and Recovery Point Objective (RPO)

The Recovery Time Objective (RTO) and Recovery Point Objective (RPO) are critical factors in determining the maximum acceptable downtime and data loss during a disaster event. Organizations with stringent RTO and RPO requirements need disaster recovery sites with near real-time data replication capabilities.

To meet aggressive RTO and RPO targets, organizations should consider employing continuous data protection (CDP) technologies. CDP solutions allow for continuous data backups, ensuring that data changes are immediately replicated to the disaster recovery site, minimizing data loss. Additionally, the use of snapshot-based replication strategies can help achieve shorter RTOs by providing point-in-time recovery options [24,25].

3.6 Scalability and Capacity

The disaster recovery site should be able to accommodate increased workloads and data volumes during recovery operations and future growth. Adopting scalable cloud-based disaster recovery solutions enables organizations to easily expand resources based on demand. Scalability is especially crucial in cloud-based environments where resource provisioning can be dynamically adjusted. Implementation of auto-scaling mechanisms allows the disaster recovery site to automatically adjust resource allocation based on predefined thresholds. This ensures that the site can handle sudden surges in demand during recovery without manual intervention. Additionally, organizations should conduct regular capacity planning exercises to assess current and future resource requirements, enabling them to make informed decisions about scaling the disaster recovery site [29].

3.7 Compatibility and Integration

Seamless integration between the primary data center and the disaster recovery site is vital for efficient data replication and failover operations. Ensuring compatibility between the hardware, software, and cloud infrastructure at both locations is critical. Deploying disaster recovery solutions from the same vendor as the primary data center can facilitate compatibility and simplify integration. Leveraging standardized data formats and protocols further enhances compatibility, enabling smooth data synchronization between environments. Interoperability testing should be conducted regularly to validate compatibility and ensure that all systems and applications function as intended during recovery efforts.

3.8 Cost and Budget Constraints

Cost considerations play a significant role in disaster recovery site selection. Organizations must strike a balance between investing in robust disaster recovery capabilities and adhering to budget constraints. While implementing high-end disaster recovery solutions might provide superior resilience, it may not always be financially viable for all organizations. Exploring cost-effective disaster recovery options, such as hybrid cloud deployments, which combine on-premises resources with cloud-based recovery capabilities. Additionally, cost optimization can be achieved by prioritizing critical workloads and data for recovery while adopting less expensive solutions for less critical assets [24].

3.9 Expertise and Support

The expertise and support provided by the disaster recovery site provider are crucial in ensuring a successful recovery process. Choosing a provider with a proven track record in disaster recovery and expertise in cloud-based solutions is essential. References from previous clients and customer testimonials can be valuable in evaluating the provider's capabilities. Furthermore, the provider

should offer 24/7 support, including trained personnel who can assist during the recovery process [24].

3.10 Compliance and Regulatory Requirements

Adherence to industry regulations and data protection laws is paramount when selecting a disaster recovery site. Different industries may have specific compliance requirements, such as HIPAA for healthcare or GDPR for businesses dealing with EU citizens' data. Organizations must ensure that the disaster recovery site complies with all relevant regulations and meets the necessary data security and privacy standards [26].

3.11 Testing and Monitoring Capabilities

Regular testing and monitoring of the disaster recovery site are essential to validate its effectiveness and identify any potential issues proactively. Conducting periodic disaster recovery drills and simulations enables organizations to identify weaknesses in their recovery plans and make necessary improvements. Implementing comprehensive monitoring tools helps track the health and performance of the disaster recovery site continuously. This ensures that any deviations from expected performance are detected promptly and addressed [24].

4. Data Collection and Processing

4.1 Data Source

In this research, the data were collected from the three major cloud providers, namely GCP, Microsoft Azure, and AWS [18-20]. For each provider, the geographic distribution and location of the data centers were collected on the regional level. The collected dataset contains 134 data centers, divided among the cloud provides as follows: 23 for GCP [30], 52 for AWS [20], and 58 for Microsoft Azure [31]. Subsequently, the geographic coordinates (latitude and longitude) of each data center were found using the address of each data center with the help of Google Sheets [32] and ezGeocode extension [33]. These geographic coordinates will then be used in the model development to estimate the distances between data centers. The geographic distribution of all considered data centers of the three cloud providers was visualized using ArcMap [34] as presented in Figure 1.

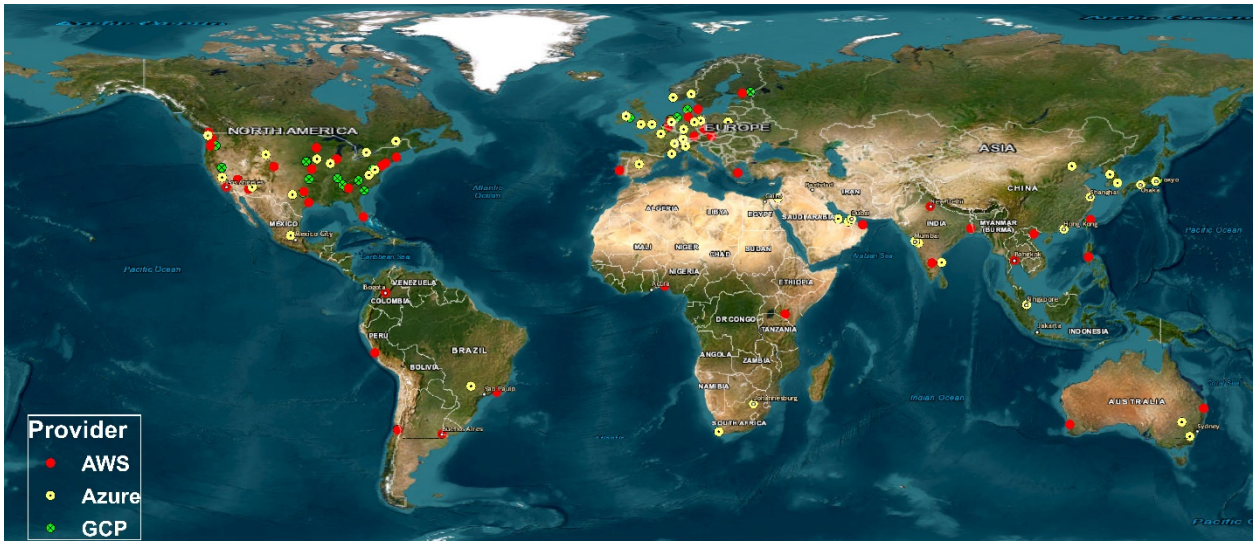

Fig. 1. Geographical distribution of cloud data centers

In this study, an unconventional approach was undertaken by merging regional data centers from the three leading cloud providers, namely GCP, Microsoft Azure, and AWS, into a unified entity. This unconventional method aimed to construct an expansive and diverse dataset, treating the collective data centers as a single cloud provider. The rationale behind this unconventional consolidation was to curate a dataset that encompassed a substantial number of data centers, thus facilitating the development of robust models designed to meticulously select the optimal backup location. By treating these major cloud providers as a cohesive unit, the study sought to enhance the accuracy and effectiveness of the models, ensuring they are equipped to make informed decisions that align with the complex demands of data backup and recovery strategies.

4.2 Data Processing

Disaster recovery (DR) site selection depends on several factors, of which we consider the most important factors to be distance, time of data transfer, Network latency, predictions of natural disasters, cyber security, and physical security as described in Section 3. These factors formulate our disaster recovery model, which allows us to specify potential disaster recovery locations. In other words, the mentioned factors are the independent variables that affect our dependent variable, disaster recovery site selection. To generate the values corresponding to each independent variable, we assumed that Beijing is the original data center and the availability of all data centers of the three cloud providers as part of a hypothesized collaboration protocol in case of threats of natural disasters. the following approach was adopted as follows:

(a) Distance: we assumed that the original data center location is in Beijing (located in the Far East), from which we are seeking the optimal DR location across the globe and among the three cloud providers, assuming that the farthest location is the safest one. Since we know the geographic coordinates of all data centers, we calculated the spatial distances from Beijing to the remaining 133 data centers using MATLAB mapping toolbox [35,36]. Subsequently, we specified the maximum distance between Beijing and the farthest data center which was 19263 km. We then normalized all calculated distances in kilometres on a scale from 0-100 as a percentage of the maximum distance as shown by Eq. (1).

$$D_N = 100 \times \frac{D_i}{D_{\max}} \quad (1)$$

Where: D_N: the normalized value of the distance; D_i: the distance value in km; and $D_{\max}$: the maximum distance value in km (19263 km).

(b) Time of data transfer: the time to transfer the vulnerable data from the original data center (Beijing) to the DR location primarily depends on data size and internet bandwidth. To account for all potential data sizes of various landscapes or architectures, we calculated the time of data transfer per terra byte (TB) of data. As a result, the time of data transfer will solely rely on the bandwidth, assuming symmetric download and upload speeds, which were calculated using Eq. (2).

$$t_N = 100 \times \frac{1/t_i}{1/t_{\min}} \quad (2)$$

Where t_N: the normalized time of transfer for a data record; t_i: the time of transfer in seconds; and $t_{\min}$: the minimum value of transfer in seconds (10 seconds).

We randomly generated bandwidth speeds that varied from 1 to 100 GB/sec, which takes into consideration the variance of bandwidth among sites globally. An example bandwidth of GCP that can reach up to 100 GB/sec can be found via Google Cloud [37].

(c) Network latency: we assumed the benchmark network latency to be 20 ms, which is optimally convenient for real-time applications. Therefore, we randomly generated latency values varying from 20 ms to 200 ms to account for the distances from the origin data center to potential DR locations. It is crucial to mention that while distance is correlated to network latency, the spatial distance, the first independent variable, is not physically correlated to the network latency because it was considered as a measure of safety. All latency values were normalized on a scale of 1-100 using Eq. (3).

$$L_N = 100 \times \frac{1/L_i}{1/L_{\min}} \quad (3)$$

Where L_N: the normalized value of the network latency; L_i: the value of network latency in milliseconds; and $L_{\min}$: the minimum value of network latency (20 milliseconds).

(d) Predictions of natural disasters: it was assumed that predictions that reach 50% or above are considered an imminent threat, in which case all DR sites under this prediction threshold will be ignored. On the contrary, predictions that below 50% are considered safe. As a result, values from 0% to 49% were randomly generated for all potential DR locations.

(e) Cyber and physical security: values for these independent variables were assumed and randomly generated between 1 and 100 for all DR sites.

The final result after the data processing is a set of independent variables whose values vary from 0 to 100. A sample of the final dataset is presented in Table 1.

Table 1
A sample of the cloud data centers dataset

Data center	Distance (x_1)	Time of data transfer (x_2)	Network latency (x_3)	Prediction of natural disaster (x_4)	Cyber security (x_5)	Physical security (x_6)
Berkeley County, South Carolina	61	7	10	48	28	82
Council Bluffs, Iowa	54	55	11	14	15	19
The Dalles, Oregon	46	62	11	43	20	83
Douglas County, Georgia	60	77	20	4	69	15
Henderson, Nevada	52	68	42	49	39	50

5. Development of Low-cost Disaster Recovery Models

The proposed DR models in this section are optimally suitable for large-scale business corporations that possess a large landscape or architecture, where such architecture requires heavy computational resources (large capital expenditure). In this scenario, the disaster recovery service will be significantly expensive because a similar architecture of a productive system must be replicated in the DR site. These types of contracts are billed either monthly or yearly. To minimize the cost, the proposed models will drastically reduce the cost. In other words, payments will be required only at the time of a potential disaster.

5.1 Grid Analysis Decision Model

After completing the data processing stage, a prerequisite for developing a grid analysis decision model is determining the weights of the independent variables. In this research, these weights are estimated as shown in Table 2.

Table 2
Weights of the factors considered for disaster recovery location selection

Independent variable	Distance (x_1)	Time of data transfer (x_2)	Network latency (x_3)	Prediction of natural disaster (x_4)	Cyber security (x_5)	Physical security (x_6)
Weight (w)	2	3	1.5	1	1	1

It is important to mention that the assigned variable weights reflect the importance of each variable. This weight estimation is completely subjective, and such subjectivity is common when developing grid analysis decision models. However, the weight estimations can be less susceptible to subjectivity should more data records be available (i.e., big data from major cloud providers).

The final step of the grid decision model is to calculate the scores of all potential DR locations, which is accomplished using Eq. (4).

$$Score = \sum_{i=1}^{6} w_i \times x_i \tag{4}$$

5.2 Artificial Neural Network Model

Artificial neural networks (ANNs) are computational models designed to imitate the intricate workings of the human brain. Just like the human brain's interconnected web of neurons, ANNs consist of layers of artificial neurons, each performing simple operations and passing signals to the next layer. These connections enable them to process vast amounts of data, recognize patterns, and make decisions, much like the human brain's synaptic connections facilitate learning and cognition. Moreover, just as the brain refines its connections through learning, ANNs utilize algorithms to adjust the strengths of their connections based on input data, allowing them to adapt and improve their performance over time. While ANNs are still simplified representations of the brain's complexity, they demonstrate the remarkable capacity to process information and perform tasks, fueling ongoing research and advancements in artificial intelligence and machine learning. These networks excel in modelling complex relationships and patterns in data, making them versatile for various applications. ANNs have found extensive use in modelling tasks, such as image recognition, natural language processing, speech recognition, and recommendation systems. Additionally, ANNs have also proven to be highly effective in modelling regression problems, where the goal is to predict a continuous output variable based on input features. ANNs can capture complex nonlinear relationships between input variables and the target, making them well-suited for regression tasks with intricate data patterns. By employing various activation functions and multiple hidden layers, ANNs can learn to approximate complex functions, enabling accurate regression predictions [38-43].

For instance, the effectiveness of demonstrated the effectiveness of deep convolutional neural networks like VGG and ResNet in image classification and object detection tasks was demonstrated by Krizhevsky *et al.,* [44]. In the domain of natural language processing, recurrent neural networks (RNNs) and transformer-based models like BERT and GPT have achieved state-of-the-art results in tasks such as machine translation, sentiment analysis, and language generation [45]. Moreover, ANN applications extend beyond these domains, encompassing diverse fields like finance, healthcare, and robotics.

In this study, the ANN model developed using Keras was tailored to address a specific problem with input data containing six features [46]. The architecture was thoughtfully crafted to extract meaningful representations from the input data through the use of multiple layers. The initial layer serves as the input layer with six neurons, each corresponding to one feature in the dataset. This layer acts as the gateway for the data, passing it forward for further processing. Subsequently, the model comprises two hidden layers, the first with 12 neurons and the second with eight neurons. These hidden layers are essential for the network's ability to learn intricate and non-linear relationships within the data. The neurons in these layers apply weighted transformations to the input data, capturing higher-level features and patterns. The choice of 12 and eight neurons in the hidden layers was carefully determined through experimentation and tuning to strike a balance between model complexity and generalization. Moving forward in the architecture, we reach the output layer with a single neuron. This neuron is responsible for producing the final prediction based on the learned representations from the previous layers. Depending on the nature of the problem being addressed, this prediction could be a regression value, binary classification, or a probability score, among others. The architecture of the developed network is presented in Figure 2, generated by the NN-SVG tool [47,48].

To ensure the model's ability to generalize well on unseen data and avoid overfitting, the dataset was thoughtfully divided into three sets: the training set, responsible for training the model; the cross-validation set, used for tuning hyperparameters and early stopping; and the testing set, employed to assess the model's performance on entirely new data.

For optimization during training, I opted to use the Adam optimizer, which is an adaptive learning rate algorithm. The Adam optimizer combines the advantages of both the stochastic gradient descent (SGD) and root mean square propagation (RMSprop) optimizers. It efficiently adapts the learning rate based on each parameter's past gradients, allowing the model to converge faster and potentially find better solutions for the given task. The use of Adam further enhances the training process and contributes to the model's overall performance and accuracy.

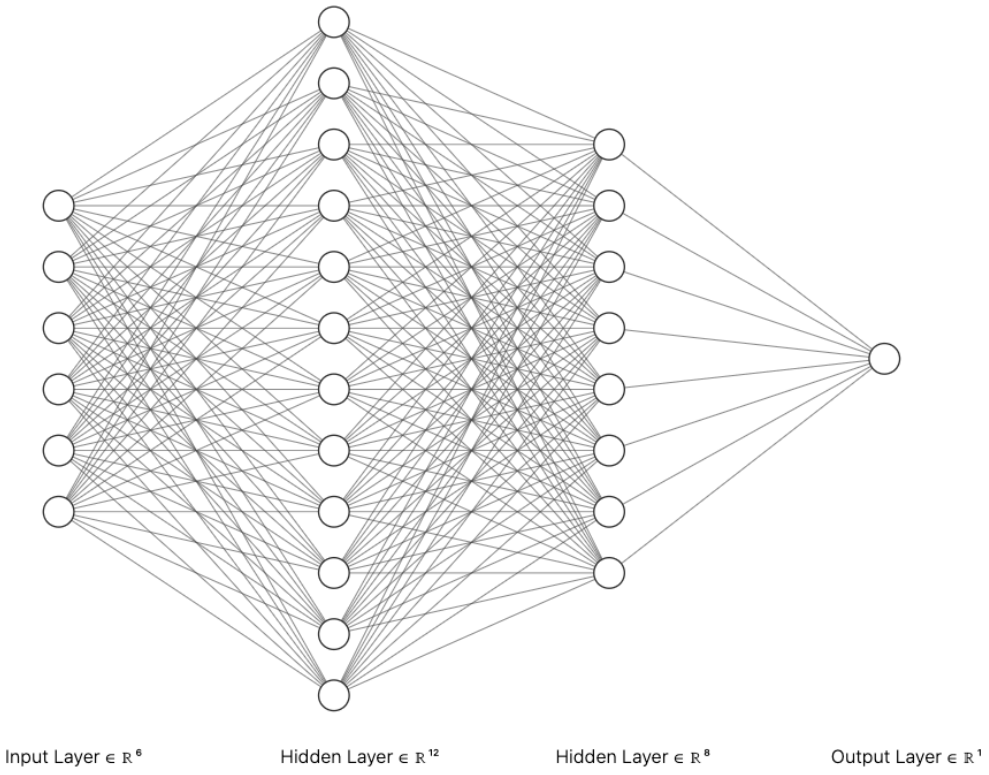

Fig. 2. A graphical visualization of the developed ANN

As shown in Figure 3, the Root Mean Squared Error (RMSE) curve provides valuable insights into the performance of the ANN during the training and validation phases. The RMSE is a widely used metric for regression tasks, measuring the average difference between the predicted values and the actual target values. As the training progresses, the RMSE curve illustrates how effectively the model is learning from the data. A decreasing RMSE indicates that the network is successfully reducing prediction errors and honing its ability to approximate the target values. In the context of this ANN, it is encouraging to observe a decreasing trend in the RMSE curve for both the training and validation sets, signifying that the model is making accurate predictions for the unseen validation data as well. This suggests that the ANN is effectively capturing underlying patterns and generalizing well to new samples.

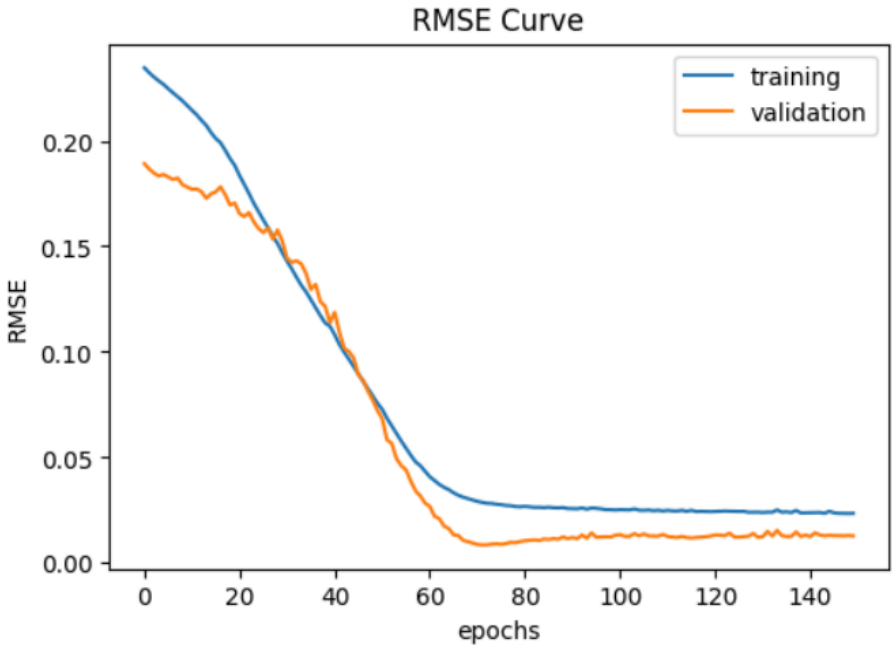

Fig. 3. ANN performance expressed by the RMSE for training and validation

The loss curve is presented in Figure 4. The figure showcases the performance of the ANN from a computational perspective, depicting how the loss function evolves during the training and validation processes. The loss function represents the discrepancy between the predicted outputs and the actual targets. A lower value of the loss function signifies that the model's predictions are closer to the ground truth, indicating better convergence and improved model performance. Throughout the training and validation stages, observing a decreasing trend in the loss curve is a positive indicator, indicating that the network is effectively minimizing the error and fitting the data well. The alignment of the training and validation loss curves is crucial; if both curves follow a similar trajectory, it implies that the model is not overfitting and maintains a stable generalization ability. The presence of consistent and decreasing loss values in this ANN's training and validation curves signifies its accurate learning, strong generalization, and overall effectiveness in handling the given task.

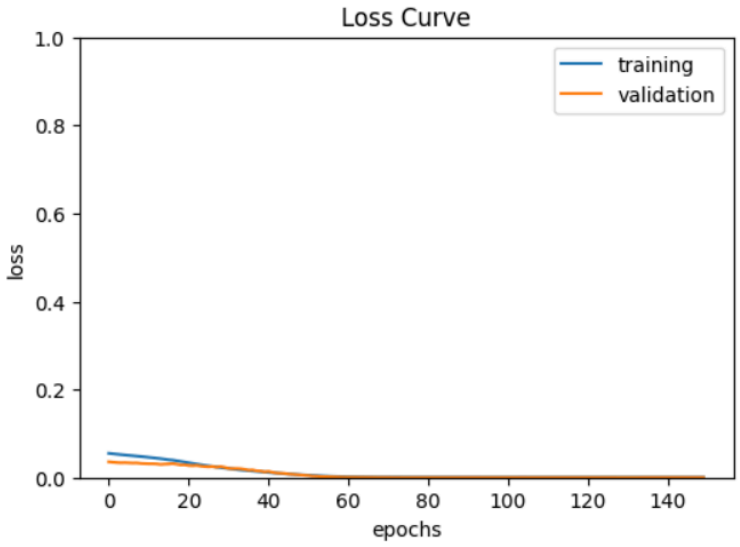

Fig. 4. ANN loss curve for training and validation

6. Conclusions

In this research, we have explored various factors that play a significant role in this decision-making process. Among these factors, redundancy, data center location, and proximity to users have emerged as the most important considerations for choosing a backup site. To aid organizations in making informed decisions, we have developed two models: the grid decision analysis model and a neural network model. These models offer valuable insights and a systematic approach to evaluating and selecting the optimal backup site from available alternatives. The power of helping, through the Internet of Things (IoT), represents a significant advancement in disaster preparedness and mitigation strategies that help propose a cost-efficient disaster recovery approach. By leveraging IoT-enabled sensors, data collection, and predictive analytics, organizations can enhance their ability to anticipate and respond to potential disasters effectively. This approach involves the real-time monitoring of environmental conditions, infrastructure health, and critical systems. By combining theoretical analysis with practical application, our research contributes to enhancing disaster preparedness and ensuring service continuity in cloud computing environments. It is crucial to mention that the developed models require more real-world datasets collected by cloud providers in order to increase the generalization and robustness of the models. As cloud technology continues to evolve, the findings of this study can serve as a foundation for further research and improvements in disaster management strategies.

Acknowledgement
This research was not funded by any grant.

References

[1] Hayes, Brian. "Cloud computing." *Communications of the ACM* 51, no. 7 (2008): 9-11. https://doi.org/10.1145/1364782.1364786
[2] Wang, Lizhe, Gregor Von Laszewski, Andrew Younge, Xi He, Marcel Kunze, Jie Tao, and Cheng Fu. "Cloud computing: a perspective study." *New Generation Computing* 28 (2010): 137-146. https://doi.org/10.1007/s00354-008-0081-5
[3] Qian, Ling, Zhiguo Luo, Yujian Du, and Leitao Guo. "Cloud computing: An overview." In *Cloud Computing: First International Conference, CloudCom 2009*, Beijing, China, December 1-4, 2009. Proceedings 1, pp. 626-631. Springer Berlin Heidelberg, 2009. https://doi.org/10.1007/978-3-642-10665-1_63
[4] Shaluf, Ibrahim M., Fakharu'l-razi Ahmadun, and Aini Mat Said. "A review of disaster and crisis." *Disaster Prevention and Management: An International Journal* 12, no. 1 (2003): 24-32. https://doi.org/10.1108/09653560310463829
[5] Abdelaziz, Ahmed, and Saleh Mesbah. "A review of disaster management frameworks." *Journal of Management Information and Decision Sciences* 24 (2021): 1-10.
[6] Alexander, David. *Natural disasters*. Routledge, 2018. https://doi.org/10.4324/9781315859149
[7] Ritchie, Hannah, and Max Roser. "Natural disasters." *Our World in Data* (2014).
[8] Mohamed Shaluf, Ibrahim. "An overview on the technological disasters." *Disaster Prevention and Management: An International Journal* 16, no. 3 (2007): 380-390. https://doi.org/10.1108/09653560710758332
[9] Mohamed Shaluf, Ibrahim. "Technological disaster stages and management." *Disaster Prevention and Management: An International Journal* 17, no. 1 (2008): 114-126. https://doi.org/10.1108/09653560810855928
[10] Artik, Yakup, Nevra Cesur, Levent Kenar, and Mesut Ortatatli. "Biological disasters: an overview of the covid-19 pandemic in the first quarter of 2021." *Afet ve Risk Dergisi* 4, no. 2 (2021): 163-182. https://doi.org/10.35341/afet.977488
[11] Vallero, Daniel A. *Unraveling environmental disasters*. Newnes, 2012.
[12] Pidgeon, Nick, and Mike O'Leary. "Man-made disasters: why technology and organizations (sometimes) fail." *Safety Science* 34, no. 1-3 (2000): 15-30. https://doi.org/10.1016/S0925-7535(00)00004-7
[13] Madakam, Somayya, Ramya Ramaswamy, and Siddharth Tripathi. "Internet of Things (IoT): A literature review." *Journal of Computer and Communications* 3, no. 5 (2015): 164-173. https://doi.org/10.4236/jcc.2015.35021

[14] Cing, Hong Cing, and Nur Syaimasyaza Mansor. "Internet of Things (IoT): Real-Time Monitoring for Decision Making Among The Malaysian Contractors." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 32, no. 3 (2023): 455-470. https://doi.org/10.37934/araset.32.3.455470
[15] Ray, Partha Pratim, Mithun Mukherjee, and Lei Shu. "Internet of things for disaster management: State-of-the-art and prospects." *IEEE Access* 5 (2017): 18818-18835. https://doi.org/10.1109/ACCESS.2017.2752174
[16] Schuurman, Nadine. *GIS: A short introduction*. Oxford: Blackwell, 2004.
[17] Tomaszewski, Brian. *Geographic information systems (GIS) for disaster management*. Routledge, 2020. https://doi.org/10.4324/9781351034869
[18] Google. "Google Cloud." *Google Cloud*. Accessed April 1, 2023. https://cloud.google.com/.
[19] Microsoft Azure. "Azure. Limitless Innovation. " *Azure*. Accessed April 1, 2023. https://azure.microsoft.com/en-us/.
[20] Amazon. "Amazon Web Services." *AWS*. Accessed April 1, 2023. https://aws.amazon.com/about-aws/global-infrastructure/localzones/locations/.
[21] IBM. "IBM Cloud: AI-ready, secure, and hybrid by design." *IBM*. Accessed April 1, 2023. https://www.ibm.com/cloud.
[22] Oracle. "Oracle Cloud Infrastructure." *Oracle*. Accessed April 1, 2023. https://www.oracle.com/eg/cloud/.
[23] Alibaba. "Alibaba Cloud." *Alibaba*. Accessed April 1, 2023. https://www.alibabacloud.com/en?_p_lc=5&utm_key=se_1012047735&utm_content=se_1012047735.
[24] Abualkishik, Abedallah Zaid, Ali A. Alwan, and Yonis Gulzar. "Disaster recovery in cloud computing systems: An overview." *International Journal of Advanced Computer Science and Applications* 11, no. 9 (2020). https://doi.org/10.14569/IJACSA.2020.0110984
[25] Alhazmi, Omar H., and Yashwant K. Malaiya. "Evaluating disaster recovery plans using the cloud." In *2013 Proceedings Annual Reliability and Maintainability Symposium (RAMS)*, pp. 1-6. IEEE, 2013. https://doi.org/10.1109/RAMS.2013.6517700
[26] Yimam, Dereje, and Eduardo B. Fernandez. "A survey of compliance issues in cloud computing." *Journal of Internet Services and Applications* 7 (2016): 1-12. https://doi.org/10.1186/s13174-016-0046-8
[27] Andrade, Ermeson, and Bruno Nogueira. "Dependability evaluation of a disaster recovery solution for IoT infrastructures." *The Journal of Supercomputing* 76, no. 3 (2020): 1828-1849. https://doi.org/10.1007/s11227-018-2290-0
[28] Carlin, Sean, and Kevin Curran. "Cloud computing security." In *Pervasive and Ubiquitous Technology Innovations for Ambient Intelligence Environments*, pp. 12-17. IGI Global, 2013. https://doi.org/10.4018/978-1-4666-2041-4.ch002
[29] Lehrig, Sebastian, Hendrik Eikerling, and Steffen Becker. "Scalability, elasticity, and efficiency in cloud computing: A systematic literature review of definitions and metrics." In *Proceedings of the 11th International ACM SIGSOFT Conference on Quality of Software Architectures*, pp. 83-92. 2015. https://doi.org/10.1145/2737182.2737185
[30] Google. "Cloud Locations." *Google Cloud*. Accessed April 1, 2023. https://cloud.google.com/about/locations.
[31] Microsoft Azure. "Microsoft Azure Data Center Locations." *Data Center Locations*. Accessed April 1, 2023. https://datacenterlocations.com/microsoft-azure/.
[32] Google. "Create, connect and collaborate with the power of AI." *Google Workspace*. Accessed April 1, 2023. https://workspace.google.com/.
[33] Google. "ezGeocode." *Google Workspace Marketplace*. February 28, 2023. https://workspace.google.com/marketplace/app/ezgeocode/163876376905.
[34] ArcMap. "Resources for ArcMap." *Esri*. Accessed April 1, 2023. https://www.esri.com/en-us/arcgis/products/arcgis-desktop/resources.
[35] MathWorks. "Matlab. Math. Graphics. Programming." *MathWorks*. Accessed April 1, 2023. https://www.mathworks.com/products/matlab.html.
[36] MathWorks. "Mapping Toolbox." *MathWorks*. Accessed April 1, 2023. https://www.mathworks.com/products/mapping.html.
[37] Google. "Network bandwidth." *Google Cloud*. Accessed April 1, 2023. https://cloud.google.com/compute/docs/network-bandwidth.
[38] Hopfield, John J. "Artificial neural networks." *IEEE Circuits Devices Magazine* 4, no. 5 (1988): 3-10. https://doi.org/10.1109/101.8118
[39] Jain, Anil K., Jianchang Mao, and K. Moidin Mohiuddin. "Artificial neural networks: A tutorial." *Computer* 29, no. 3 (1996): 31-44. https://doi.org/10.1109/2.485891
[40] Abraham, Ajith. "Artificial neural networks." *Handbook of Measuring System Design* (2005). https://doi.org/10.1002/0471497398.mm421
[41] Krogh, Anders. "What are artificial neural networks?." *Nature Biotechnology* 26, no. 2 (2008): 195-197. https://doi.org/10.1038/nbt1386

[42] Yegnanarayana, Bayya. *Artificial neural networks*. PHI Learning Pvt. Ltd., 2009.
[43] Aquil, Mohammad Amimul Ihsan, and Wan Hussain Wan Ishak. "Comparison of Machine Learning Models in Forecasting Reservoir Water Level." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 31, no. 3 (2023): 137-144. https://doi.org/10.37934/araset.31.3.137144
[44] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in Neural Information Processing Systems* 25 (2012).
[45] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention is all you need." *Advances in Neural Information Processing Systems* 30 (2017).
[46] KERAS. "Simple. Flexible. Powerful." *Keras*. Accessed April 1, 2023. https://keras.io/.
[47] LeNail, Alexander. "NN-SVG: Publication-Ready Neural Network Architecture Schematics." *Journal of Open Source Software* 4, no. 33 (2019): 747. https://doi.org/10.21105/joss.00747
[48] LeNail, Alexander. "NN-SVG: Tool." *AlexLeNail*. Accessed April 1, 2023. http://alexlenail.me/NN-SVG/index.html.