



## Article

# ZTCloudGuard: Zero Trust Context-Aware Access Management Framework to Avoid Medical Errors in the Era of Generative AI and Cloud-Based Health Information Ecosystems

Khalid Al-hammuri <sup>1,\*</sup> , Fayez Gebali <sup>1</sup> and Awos Kanan <sup>2</sup> 

<sup>1</sup> Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 2Y2, Canada; fayez@uvic.ca

<sup>2</sup> Department of Computer Engineering, Princess Sumaya University for Technology, Amman 11941, Jordan; a.kanan@psut.edu.jo

\* Correspondence: khalidalhammuri@uvic.ca

**Abstract:** Managing access between large numbers of distributed medical devices has become a crucial aspect of modern healthcare systems, enabling the establishment of smart hospitals and telehealth infrastructure. However, as telehealth technology continues to evolve and Internet of Things (IoT) devices become more widely used, they are also increasingly exposed to various types of vulnerabilities and medical errors. In healthcare information systems, about 90% of vulnerabilities emerge from medical error and human error. As a result, there is a need for additional research and development of security tools to prevent such attacks. This article proposes a zero-trust-based context-aware framework for managing access to the main components of the cloud ecosystem, including users, devices, and output data. The main goal and benefit of the proposed framework is to build a scoring system to prevent or alleviate medical errors while using distributed medical devices in cloud-based healthcare information systems. The framework has two main scoring criteria to maintain the chain of trust. First, it proposes a critical trust score based on cloud-native microservices for authentication, encryption, logging, and authorizations. Second, a bond trust scoring system is created to assess the real-time semantic and syntactic analysis of attributes stored in a healthcare information system. The analysis is based on a pre-trained machine learning model that generates the semantic and syntactic scores. The framework also takes into account regulatory compliance and user consent in the creation of the scoring system. The advantage of this method is that it applies to any language and adapts to all attributes, as it relies on a language model, not just a set of predefined and limited attributes. The results show a high *F1* score of 93.5%, which proves that it is valid for detecting medical errors.

**Keywords:** access management; zero-trust; distributed medical devices; cloud; health information system; medical errors; IoT



**Citation:** Al-hammuri, K.; Gebali, F.; Kanan, A. ZTCloudGuard: Zero Trust Context-Aware Access Management Framework to Avoid Medical Errors in the Era of Generative AI and Cloud-Based Health Information Ecosystems. *AI* **2024**, *5*, 1111–1131. <https://doi.org/10.3390/ai5030055>

Academic Editors: Sheikh Tahir Bakhsh, Sabeen Tahir and Basit Shahzad

Received: 20 May 2024

Revised: 1 July 2024

Accepted: 3 July 2024

Published: 8 July 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

**Problem statement:** This article proposes a zero-trust access management framework for healthcare information systems. We also conducted a case study on medical errors to verify the proposed framework's viability.

Medical errors in health care are defined as circumstances that lead to a wrong medical decision. Such a decision may lead to prescribing the wrong drug [1,2], issuing the wrong report, or making a false diagnosis. Medical errors are very critical, as they may be caused by normal users, not just by fraud. In healthcare, 90% of vulnerabilities are due to medical error or human error. This is difficult to detect, as it is caused by authorized and authenticated users. In the era of generative AI assistance, biased, discriminated, or even wrong medical reports could be generated [3]. Medical errors could be caused by either

human or AI assistant systems. The proposed framework is designed to create a scoring system for any type of user, data, or device.

Controlling access to devices and their users (either human or AI robots) in healthcare systems, along with the associated data, represents a major challenge for any service provider. While promoting the idea of smart hospitals and telehealth, it is required to look deeply into the existing regulation and access control systems in order to ensure their validity in the context of technologies such as Internet of Things (IoT) devices, cloud [4], AI [5,6], blockchain [7,8], quantum computing [9,10], and 5G networks [11]. There are many different types of medical devices used within distributed or cloud-based healthcare information environments. Examples of these devices include patient monitoring devices, handheld and portable devices, telehealth consulting, medical imaging systems, robotics, and virtual reality.

The main challenges facing the cloud-based healthcare infrastructure involve managing access to these devices while guaranteeing that the received data are secure, clean, and clinically valid. The stressful environment and complex technology in healthcare settings means that serious attention and advanced skills are required to operate such systems. At the same time, both devices and the data should be monitored in real-time in order to intercept any abnormalities in data or wrong reports sent by healthcare practitioners. Such a system should control users, data, and output.

**Constraints:** The healthcare industry is subject to strict regulations regarding the use of patient information. In the United States, the Health Insurance Portability and Accountability Act (HIPAA) governs patient information compliance. Similarly, in Canada Bill (C-27) regulates the healthcare information system, and Health Canada also plays a role. Bill (C-27) is a new law that replaces the Personal Information Protection and Electronic Document Act (PIPEDA), and has enacted three regulations: the Consumer Privacy Protection Act (CPPA), the Personal Information and Data Protection Tribunal Act (PIDPTA), and the Artificial Intelligence and Data Act (AIDA). In Europe, compliance with the General Data Protection Regulation (GDPR) regulates the sharing of information in healthcare.

**Proposed solution:** In order to adapt to new advances in technology and strict compliance requirements, this paper proposes a zero-trust context-aware access control framework for medical IoT devices to manage the patient information system within the complex structure of healthcare systems. The proposed framework consists of a set of practices, policies, and attributes for enhancing and managing the security of access control systems for any healthcare infrastructure. The framework mainly focuses on preventing medical errors in the healthcare industry. These cases are complex and difficult to detect, as they are sent by authorized and authenticated users. The proposed context-aware system can alleviate user errors by analyzing the complex metadata of the user, device, and output. The proposed context-aware framework ensures that the data are relevant, consistent, authenticated, and only sent by authorized users to the designated destination at the endpoint devices and users.

The research goals and contributions of this article are listed in the following itemized points and linked to each related section:

- We propose a zero trust context-aware management framework to minimize medical errors by maintaining the trust cycle between the user ( $x$ ), hardware ( $y$ ), and output data ( $z$ ); see Sections 3.1 and 3.2.
- We evaluate the trust score by deriving two main trust criteria: critical trust and bond trust. Critical trust is based on a set of cloud-native microservices, while bond trust is used to evaluate the mutual relationship between the user ( $x$ ), hardware ( $y$ ), and data ( $z$ ) using syntactic and semantic analysis; see Section 3.3.
- We construct a decision-making engine to grant a final decision that considers the complex nature of the healthcare system by count for regulatory compliance, access constraints, access level, and access operations; see Sections 3.4 and 3.5.

- We validate the proposed framework by utilizing the Word2Vec language model to conduct syntactic and semantic analysis of the framework using a synthetic dataset; see Section 4.

The rest of this article is organized as follows: Section 2 provides a background of the current access control work; Section 3 explain the proposed framework in detail; Section 4 report and discuss the experiment results; and Section 5 concludes the paper.

## 2. Background and Related Work

Processing large volumes of data within the healthcare ecosystem makes medical report automation challenging and error-prone. To validate the input and output for the healthcare system data flow, it is necessary to ensure confidentiality, availability, and integrity. Implementing efficient access management is essential for data confidentiality. At the same time, an AI-based context-aware system can be used to validate data integrity [12,13]. While a resilient system [14] is important to make the system available when it is needed, a cloud-based system could be the recommended solution.

The existing healthcare information systems rely on the HL7FHIR standards [15] to define the communication and security access control system in healthcare. There are three main subsystems within the HL7FHIR:

1. Authentication: Verifies the user.
2. Access control engine: Decides which FHIR controls are allowed for the user using the CRUD method (Create, Read, Update, Delete).
3. Audit log: records actions and any suspicious system intrusions.

At the organizational level, the access control system has three main common types within the health information system:

- RBAC: Role-based access control [16,17].
- ABAC: Attribute-based access control [18,19].
- CML: Modern cloud-based machine learning access control [20–22].

**RBAC** and **ABAC** are the standard access control systems in healthcare. They are used widely in the traditional healthcare infrastructure setup for managing access control within the hospital perimeter. **RBAC** manages access based on the user's role and grants permission based on the **CRUD** or **HTTP** method. **RBAC** is complex, and considers different factors such as users (operator, patient), roles, permissions, resources objects, and context of the data access [15,23]. Please refer to Table A1 in Appendix A for more information on the role-based access system factors. The role-based access management system has limitations that make it less effective in complex modern healthcare environments. **RBAC** is time-consuming and requires manual work to adjust rules and policies, making it less effective for real-time access management, which has too many factors in a complex cloud-based environment.

**ABAC** [24,25] is based on predefined policies and conditions, and grants access to system resources or objects based on specific data attributes. For instance, in compliance with regulations, patient identification information cannot be accessed without having the patient consent attribute to process the data. On the other hand, **ABAC** has a considerable amount of challenges. It is not efficient at big-data applications, which limits its scalability. It is also time consuming and requires a considerable amount of resources, making it inefficient in dynamic and today's globally distributed healthcare systems.

In traditional access control for healthcare information system, the user typically sends a request to the server through the REST API gateway. The server then sends a request for the REST API to verify the user information in order to grant the required **CRUD** operations based on the predefined rules and policies.

Modern cloud-based access control systems mainly rely on the zero trust principle, which analyzes everything in the network and does not grant trust to any entity for data access, either user or device, without first passing a set of conditions that are defined by the organization policy. However, the main challenge is to identify what attributes or

data context will be considered in the policy without compromising the quality of the provided service. Mitigating the risk within the network is also essential in evaluating access decisions [26].

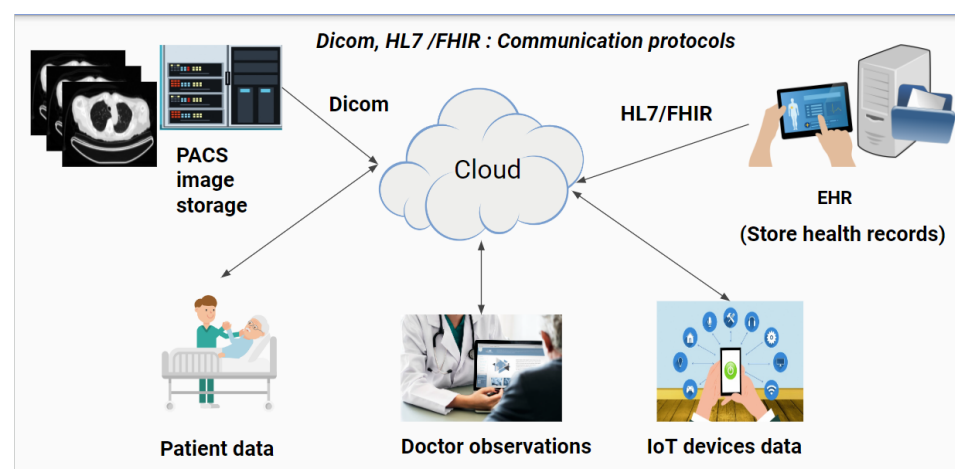
Defining attributes and enabling zero-trust features requires the utilization of advanced AI algorithms [27]. Among these algorithms, computer vision is essential for analyzing medical imaging [28,29]. It has been used for different applications, for instance, ophthalmology [30]. Natural language processing (NLP) is useful for understanding linguistic details and predicting diseases [31]. NLP is also an effective tool for medical report processing [32]. NLP is attracting more attention in the area of privacy preservation and anonymization of medical reports [33].

Voice recognition is another part of the digital transformation and automation of healthcare. Sound processing is vital for building context-aware systems to validate data integrity. Speech can be used in anomaly detection applications and emotion recognition for people with special considerations and elderly people [34,35].

Recently, large foundational models that include huge datasets for image, text, and sound have been developed. Most of the current research involves building multi-modality systems. This type of model can help in processing complex unstructured data. Using large language models for healthcare queries was evaluated in [36]. Models for ingesting and analyzing electronic health records (EHR) were investigated in [37]. Several language models used in healthcare information systems were surveyed in [38] from the standpoints of technology and ethical practice. LLaMA and GPT-4 are the two most common general-purpose language models. In the medical field, the accuracy of a specialized language model is vital. Med-PaLM [39], introduced by Google and trained on high-quality medical data, meets the expert level for answering medical questions. Large models trained on high-quality data are taking their place in the healthcare industry, as they can alleviate cost burdens while minimizing both technical and human-related error.

Within the healthcare information system, different data sources used in the decision engine act as a brain for the centralized healthcare information system. Medical images are typically stored in the Digital Imaging and Communications in Medicine (DICOM) format.

To store DICOM images, the Picture Archiving Communication System (PACS) is used to further process the raw data. Patient records are stored in Electronic Health Records (EHR). EHRs communicate with other devices through the High-Level Seven (HL7) and Fast Healthcare Interoperability Resources (FHIR) communication protocols. For other medical devices, which can be either portable, handheld, or hospital-based IoT devices, the cloud-based infrastructure processes them in real time or through batch processing mechanisms. Figure 1 illustrates the main data sources that act as the backbone of a cloud-based healthcare information system and can be utilized for minimizing medical errors.



**Figure 1.** Visualization of the main sources of healthcare-related information within the cloud-based system.

### 3. Method

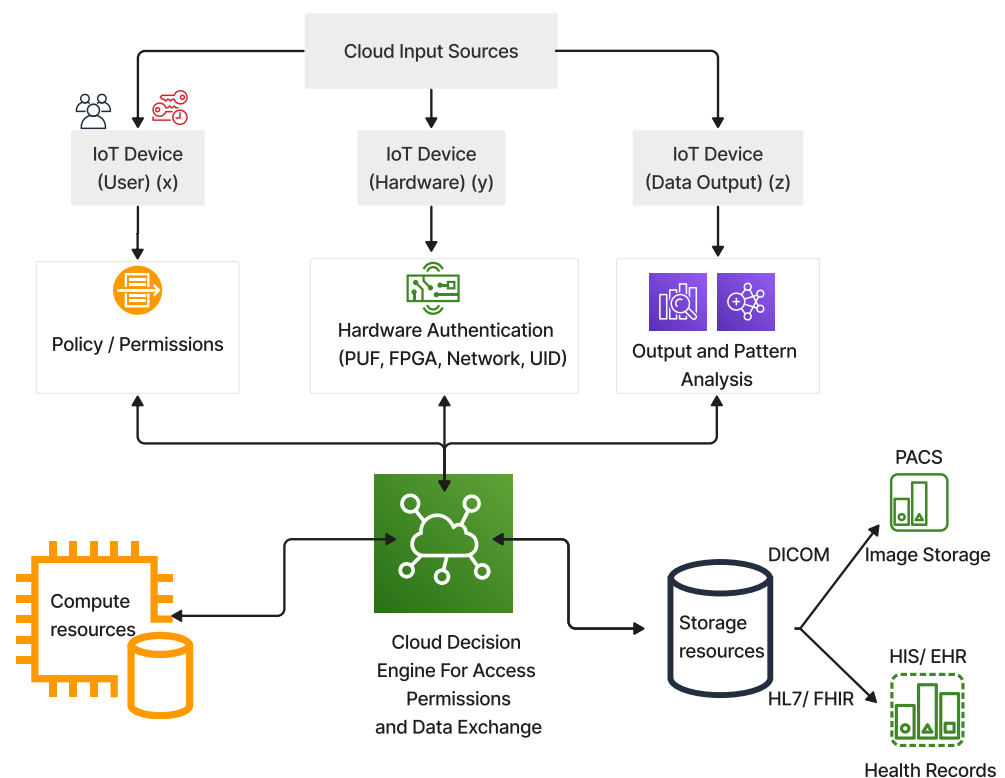
In this section, the research methodology is explained in several sub-sections: Section 3.1 provides an overview of the high-level architecture for the zero-trust context-aware system; Section 3.2 highlights the main pillars of the trust cycle for the zero-trust system; Section 3.3 evaluates the trust between different trust cycle attributes; finally, Section 3.4 explains the hierarchical process of the decision engine.

#### 3.1. Overview of the Proposed Zero-Trust Framework for Access Management

The architecture design for the proposed context-aware access management framework is depicted in Figure 2. The proposed access control system considers the zero-trust context-aware system to manage and analyze the data journey from the user of medical IoT device endpoints to the cloud resource destination.

The proposed framework is classified into three main layers, as listed below:

- **Cloud input sources:** This layer is the front-end gateway for the main input source from users, device metadata, and the context of data output either stored in the database or ingested in real-time streaming.
- **Cloud decision engine:** This is the centralized layer, acting as a brain for the decision engine. A chain of trust is built for each component based on the trust scores. There are two scores: critical trust (CT) and bond trust (BT). The engine encodes the context attributes for further analysis at a hierarchical level. In the end, it grants the final access decision, operations, and constraints based on this analysis.
- **Cloud resources:** This is the back-end layer for the zero-trust ecosystem. The main components contain the cloud computing and storage resources that are used to process and store the metadata in the healthcare database.

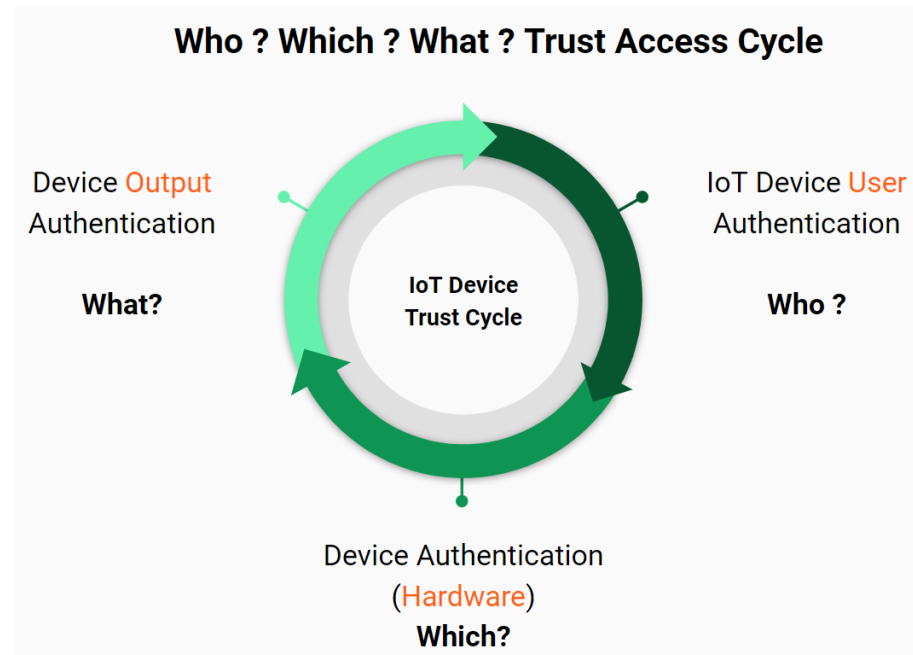


**Figure 2.** Representative image of the proposed access control functional diagram within the healthcare cloud-AI ecosystem.

The following subsections explain the components of the proposed framework in detail.

### 3.2. Trust Cycle Pillars

The proposed system harnesses the zero trust context-aware system to manage access from the cloud input sources. The zero trust principle is based on utilizing all available data points for access management, including user identity, location, device health, services, workload, and data classification. There are three main components for the context-aware cycle that consider the context of the five zero-trust elements. Figure 3 depicts the three components of the trust cycle, namely: who is the user? which device is used? and what is the output?



**Figure 3.** Trust cycle of the proposed access control framework.

The trust cycle has the following five elements, which are pillars of the proposed zero-trust principle:

1. User (identity)
2. IoT device (hardware)
3. Network (device connection)
4. Application workload (output patterns and scale)
5. Data (output transaction context).

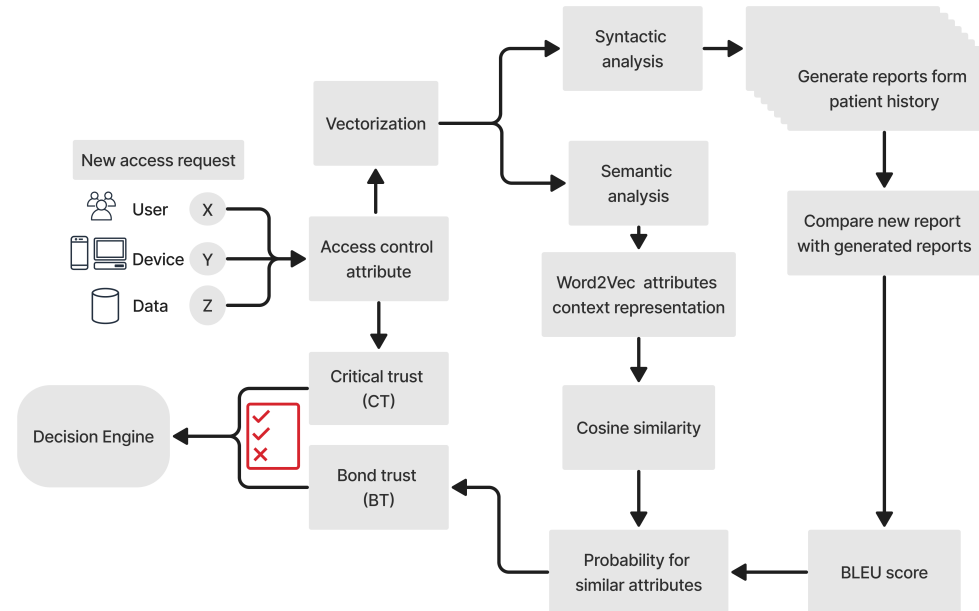
The identity relates to the user component, while the IoT device and network relate to the hardware component. Finally, the application workload and transaction context relate to the output component.

### 3.3. Trust Assessment

Building a zero-trust system requires defining a set of attributes from different categories to verify the trust cycle. The zero-trust ecosystem needs to be verified through a continuous trust cycle by implementing a series and chain of trust in order to assess semantic and syntactic relationships between the cloud input sources from users, devices, and output data. The chain of trust is important for deciding what level of access can be granted and denying access if the connection is below the threshold of an acceptable trust score.

Figure 4 illustrates the chain of trust and the assessment scoring criteria within the cloud ecosystem. The proposed framework constructs two assessment scoring criteria to manage the access of distributed medical devices. First is the critical trust (*CT*), which relies on cloud-native microservices. Second is the bond trust (*BT*), which is a proposed

scoring scheme to manage access control, as explained below. *BT* uses pre-trained machine learning models to analyze the semantic and syntactic attributes from the trusted and authorized change of the zero-trust cycle pillars see Section 3.2, related to users, devices, and data output.



**Figure 4.** Proposed framework for a continuous chain of trust based on the accumulated trust score of each zero-trust access management component.

**Critical Trust (CT):** *CT* is the initial evaluation and scoring criteria used to grant access to the cloud ecosystem. This assessment grant is preliminary and not for direct connection to the back-end resources for storage and computation. *CT* is important because it acts as an additional layer of security to separate user access control from the actual dataset resources. *CT* is evaluated using cloud-based microservices. There are four main attributes for the critical trust score. Cloud-based microservices such as authorization, authentication, logging, and encryption are digitized to derive the final *CT* score, as per Equation (1).

Each microservice attribute is assigned a logical value, i.e., 1 or 0. Then, these microservices’ logical values are multiplied by a scoring factor ( $S_i$ ) based on their importance, which can be set by the system administrator. The cloud decision engine grants access status to **allow** for trusted authority, **verify** whether more information is needed, and **deny** non-trusted access requests.

$$CT = S_1 \times A_1 + S_2 \times A_2 + S_3 \times A_3 + S_4 \times A_4 \tag{1}$$

In the above equation,  $A_1$  is the authentication and its scoring factor is  $S_1$ ;  $A_2$  is the authorization and its scoring factor is  $S_2$ ;  $A_3$  is the encryption and its scoring factor is  $S_3$ ; and  $A_4$  is the logging, with a scoring factor of  $S_4$ . Table 1 provides an example of critical trust score evaluation using different scoring factors and logical values of the micro-services.

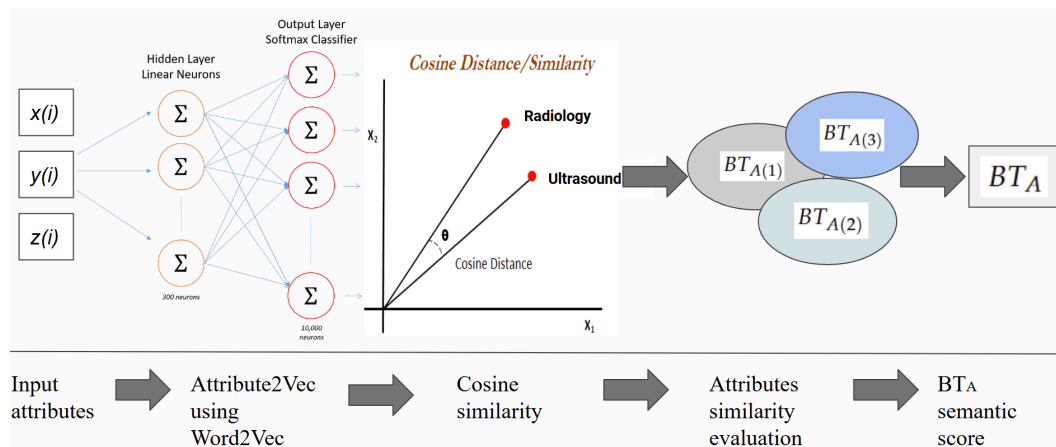
**Table 1.** Examples of critical trust score assessment.

ID	$A_1$	$S_1$	$A_2$	$S_2$	$A_3$	$S_3$	$A_4$	$S_4$	Critical Trust Score	Access Status
D1	1	0.3	1	0.4	1	0.2	1	0.1	0.9999	Allow
D2	1	0.3	0	0.4	1	0.2	1	0.1	0.6	Verify
D3	0	0.3	0	0.4	0	0.2	0	0.1	0	Deny

Bond Trust (*BT*): When a transaction passes the critical trust assessment, the bond trust is used to evaluate the relationship to other resources in order to build a trust cycle, ensuring that only authorized and highly trusted actors and designated people can access data or resources based on the organizational policy or rules. Calculating bond trust is more complex, and depends on several different aspects. *BT* has two main assessment criteria. The first is  $BT_A$ , which assesses the semantic relationship between each individual attribute stored in the health care information system. The second is  $BT_B$ , which assesses the syntactic relationship between the set of candidates in a generated health report. The reason for using these two measures is, first, that it is essential for each attribute to have meaning and to be related to similar attributes as compared to the pretrained one; second, it is essential to guarantee that the attributes in the generated report are in keeping with the context of the patient’s history to ensure that the report is highly likely to be related to the same patient, avoiding false diagnoses due to having the wrong case.

The proposed assessment of  $BT_A$  uses an Attribute2Vec representation based on a pretrained Word2Vec model [40,41]. Attribute2Vec is used to map the attributes and their synonyms to words that have the same context from the user ( $x$ ), hardware ( $y$ ), and output ( $z$ ) attributes stored in their electronic health records. The skip-gram methodology [42] is used to derive the attributes with the same context; in this framework, we suggest using the first three words with the highest context probability. The advantage of using this assessment technique is to generalize the model by accepting a wide variety of attribute descriptions in a global context. Word2Vec is valid for different languages and dialects; for example, it was used by Altibbi.com [43] to train 1.5 million medical consultation questions in the Arabic language. We recommend using a matching engine on the Vertex AI platform at Google Cloud to ensure that the word embedding and vector similarity matching processes are efficient and reliable.

Figure 5 depicts the process of assessing bond trust. The input has three attributes: users, devices, and output. The hidden layer extracts features and the *SoftMax* layer is used to predict the probability and extract the set of similar attributes that has the highest probability. In this research, we selected the three highest attributes. Eventually, the cosine similarity is used to predict the relationship between attributes from different categories ( $x$ ,  $y$ , and  $z$ ). Then, bond trust scoring is used to derive the final score to decide whether to accept or reject the attributes based on the predefined threshold.



**Figure 5.** Semantic trust assessment using Attribute2Vec, based on the Word2Vec model; here,  $BT_{A(1)}$ ,  $BT_{A(2)}$ , and  $BT_{A(3)}$  are the set of bond trust between the three input sources, respectively  $x$ ,  $y$ , and  $z$ , while  $BT$  is the final bond trust score.

The cosine distance is used in Equation (2) to predict the similarity probability of the context of attributes of  $x$ ,  $y$ , and  $z$ :



$$\text{Similarity}(A,B) = \cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{\|\vec{A}\| \|\vec{B}\|} \quad (2)$$

where  $\vec{A} \cdot \vec{B} = \sum_{i=1}^N (A_i \times B_i)$  is the dot product between two vector attributes  $\vec{A}$  and  $\vec{B}$ , while  $\|\vec{A}\| = \sqrt{\sum_{i=1}^N (A_i)^2}$ ,  $\|\vec{B}\| = \sqrt{\sum_{i=1}^N (B_i)^2}$  are the respective L2-norms of attributes  $\vec{A}$  and  $\vec{B}$  and  $\theta$  is the angle between the two vectors.

The attribute vectors with the highest probability between  $x$ ,  $y$ , and  $z$  are then used to derive the bond or semantic mutual relationship in three bond trust scores sets:  $BT_{A(1)}$ ,  $BT_{A(2)}$ , and  $BT_{A(3)}$  for the relationships between  $xy$ ,  $xz$ , and  $yz$ , where  $BT_{A(i)}$  is the bond trust set, which is derived using the two inputs described below.

A. Cosine similarity logical evaluation: Algorithm 1 is used to assign a logical value to the cosine similarity between two attributes, taking an assigned value of either one or zero based on the relationship between attributes  $x$ ,  $y$ , and  $z$ . The value is assigned based on the threshold of the angle  $\theta$  between the two attributes. Equation (2) is used to derive  $\theta$  using the cosine similarity between the attribute vector product for the given index  $i$  or position for similar context attributes. The algorithm produces a set of three logical values  $Sim_A(\vec{x}_i, \vec{y}_i)$ ,  $Sim_B(\vec{x}_i, \vec{z}_i)$ , and  $Sim_C(\vec{y}_i, \vec{z}_i)$  for each given index  $i$ .

---

**Algorithm 1** Algorithm for the proposed cosine similarity logical evaluation process

---

**Input:** User ( $x$ ), Device ( $y$ ), Output data ( $z$ ), Angle threshold ( $Th_\theta$ )

```

1: if  $\theta_{xy} \geq Th_\theta$  then
2:    $Sim_A(\vec{x}_i, \vec{y}_i) = 1$ 
3: else if  $\theta_{xy} < Th_\theta$  then
4:    $Sim_A(\vec{x}_i, \vec{y}_i) = 0$ 
5: end if
6: if  $\theta_{xz} \geq Th_\theta$  then
7:    $Sim_B(\vec{x}_i, \vec{z}_i) = 1$ 
8: else if  $\theta_{xz} < Th_\theta$  then
9:    $Sim_B(\vec{x}_i, \vec{z}_i) = 0$ 
10: end if
11: if  $\theta_{yz} \geq Th_\theta$  then
12:    $Sim_C(\vec{y}_i, \vec{z}_i) = 1$ 
13: else if  $\theta_{yz} < Th_\theta$  then
14:    $Sim_C(\vec{y}_i, \vec{z}_i) = 0$ 
15: end if
16: Output:  $Sim_A(\vec{x}_i, \vec{y}_i)$ ,  $Sim_B(\vec{x}_i, \vec{z}_i)$ ,  $Sim_C(\vec{y}_i, \vec{z}_i)$ 

```

---

B. Weight: The weight is calculated using the GloVe word embedding model [44] to consider the co-occurrence of the attributes in a global representation context of the healthcare database. The weight is based on the conditional probability of attribute occurrence or importance, as shown in Equation (3):

$$w_i = \frac{P_{BA}}{P_B} \quad (3)$$

where  $w_i$  is the probability of word  $B$  occurring in the context of word  $A$  in a given index  $i$  of two semantic or syntactically similar attributes.

The three scaler values of  $BT_{A(1)}$ ,  $BT_{A(2)}$ , and  $BT_{A(3)}$  are stored in  $BT_A$ , as shown in Equation (4), where  $BT_A$  is a  $1 \times 3$  vector:

$$BT_A = [BT_{A(1)}, BT_{A(2)}, BT_{A(3)}]. \quad (4)$$

In the above equation,  $BT_1$  is the relationship score between the user ( $x$ ) and hardware ( $y$ ), and is derived using Equation (5);  $BT_2$  is the relationship score between the user ( $x$ ) and

output ( $z$ ), and is derived using Equation (6); and  $BT_3$  is the relationship score between the output ( $z$ ) and hardware ( $y$ ), and is derived using Equation (7):

$$BT_{A(1)} = \sum_{i=1}^N (w_i)_{xy} \cdot Sim_A(\vec{x}_i, \vec{y}_i) \quad (5)$$

$$BT_{A(2)} = \sum_{i=1}^N (w_i)_{xz} \cdot Sim_B(\vec{x}_i, \vec{z}_i) \quad (6)$$

$$BT_{A(3)} = \sum_{i=1}^N (w_i)_{yz} \cdot Sim_C(\vec{y}_i, \vec{z}_i) \quad (7)$$

where  $w_i$  is a scalar weight that is used to scale the bond score for each attribute based on the importance of the feature at given  $i$  and derived by Equation (3) and  $N$  is the sequence number of attributes, which are numbered based on the probability of their context relationship. Only each similar class attribute of user, devices, and output is multiplied by each other; if they belong to the same category, the algorithm assigns them a similarity score of either 0 or 1, then multiplies them by the scalar weight for that attribute. This step is repeated for all attributes. The final multiplication is then aggregated to obtain a final scalar number that resembles the combined similarity score for  $BT_{A(i)}$ .

The  $BT_{A(i)}$  vector is normalized in Equation (8) using the *SoftMax* function. The normalization process produces a new vector  $BTN$  of dimension  $1 \times 3$ .

$$BTN_i = SoftMax(BT_{A(i)}) = \frac{\exp(BT_{A(i)})}{\sum_j \exp(BT_{A(j)})} \quad (8)$$

The result is stored in Equation (9), and has three scalar values that are between zero and one.

$$BTN_i = [BTN_1, BTN_2, BTN_3] \quad (9)$$

The first part of the bond score is calculated in Equation (10) by aggregating the three normalized scores,  $BTN_1$ ,  $BTN_2$ , and  $BTN_3$ :

$$BT_A = BTN_1 + BTN_2 + BTN_3 \quad (10)$$

where  $BT_A$  takes a value between zero and one, where zero indicates completely non-matched attributes and one indicates the highest attribute similarity match. Any number between zero and one requires an additional trust verification and reassessment.

At the same time,  $BT_B$  is used to assess the similarity in the generated report text by evaluating the syntactic performance of the candidate report generated from the stored data in the healthcare information system. Unlike semantic analysis, syntactic analysis is effective for evaluating a full report, not just the meaning of a single word; on the other hand, semantic analysis provides a wider contextual analysis using various probabilistic-related attributes.  $BT_B$  is inspired by the *BLEU* score [45], which was originally designed by *IBM* for scoring machine translation evaluations, as shown in Equation (11):

$$BT_B = \min(1, \exp(1 - \frac{reference - length}{output - length})) (\prod_{i=1}^n precision_i)^{1/n} \quad (11)$$

where it can be seen that  $BT_B$  has two parts; the first is the brevity penalty, which compensates for the length of a short generated report, while the second is the precision for the  $n$ -gram candidates. Here,  $n$  refers to the number of candidates used to evaluate the score; the notation  $n$  is typically 4 and can be increased to include more restrictions around identifying medical errors. In the case of  $n = 4$ , the *BLEU* score requires the candidate report to match the reference template by at least four attributes.

In a case with no patient history, the  $BT_B$  score is zero, making it less effective for syntactic analysis. This scoring evaluation is more meaningful when the patient has a

previous history in the EHR. The final bond trust normalizes the summation of  $BT_A$  and  $BT_B$  to keep the value between zero and one in Equation (12).

$$BT = \frac{BT_A + BT_B}{2} \tag{12}$$

### 3.4. Decision Engine Encoding and Hierarchy

The decision engine for similarity scoring is built through encoding and a hierarchical process. There are two main stages for the hierarchical process. These stages are vital to ensure that the final decision follows logical flow based on a set of constraints. The encoding shown in Figure 6 visualizes the two stages of encoding and hierarchy process for access control management.

Stage one: The decision engine performs the initial critical check for the end point device or user requesting access from the server. The critical check is essential to guarantee that the endpoint components have passed the regulatory compliance and critical trust score thresholds; see Table 1. An example of one of the main regulatory compliance factors that need to be considered is HIPAA, which lists 18 patient information identifiers [46] that are restricted from being shared without consent from patients and meeting all security guidelines within the healthcare information system, as shown in Table A2.

Stage two: The decision engine encodes the attributes from devices, output, users, and critical trust in a 32-digit hexadecimal array. This 32-digit array is then analyzed to make the final decision. The final decision is encoded to resemble the access level, operations, access resources, and constraints; see Section 3.5.

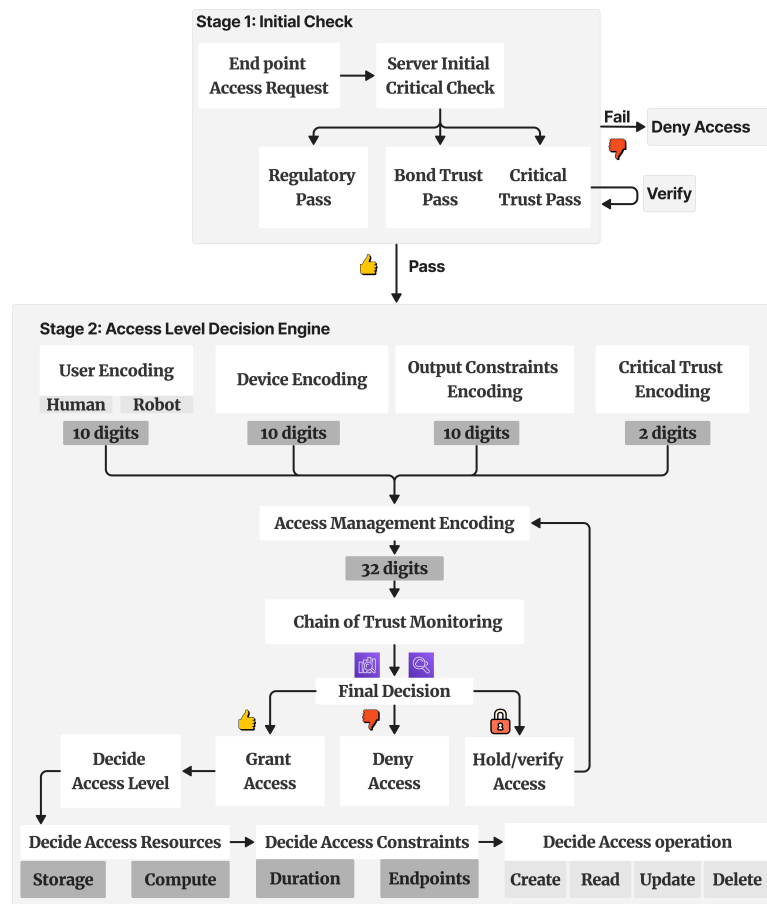


Figure 6. Representative image of the access control engine’s decision hierarchy and the related encoding.

Figure 7 illustrates an example of the encoding criteria for the three proposed zero-trust components of user, device, and output based on different attributes. Each component has a ten-digit hexadecimal value and a two-digit value for each one of the five attributes. The importance of these attributes is to ensure that the access request belongs to the designated group, has a predefined access level and type, and passes the bond trust threshold.

(A)	<b>User Encoding</b>	0	2	0	1	0	5	9	9	5	0
	<b>User encoding description</b>	<b>User type</b> 1.Human 2.Robot		<b>User role access level</b>		<b>User group access level</b>		<b>(User / Device) Bond Trust Score</b>		<b>(User / Data) Bond Trust Score</b>	
(B)	<b>Device Encoding</b>	0	2	0	3	0	4	9	3	5	0
	<b>Device encoding description</b>	<b>Device type</b> 1.Imaging 2.Laboratory		<b>Device category access level</b>		<b>Device group access level</b>		<b>(Device/ User) Bond Trust Score</b>		<b>(Device/ Data) Bond Trust Score</b>	
(C)	<b>Output Encoding</b>	0	2	0	2	1	1	9	0	9	9
	<b>Output Encoding Description</b>	<b>Output type</b> 1.Imaging 2.Laboratory		<b>Output Category access level</b>		<b>Output constraints access level (HIPAA, (PC))</b>		<b>Data/ Device (Bond Trust Score)</b>		<b>Data/ user (Bond Trust Score)</b>	

**Figure 7.** Example of access control encoding: (A) user encoding, (B) device encoding, and (C) output encoding. PC stands for patient consent. The different colors used in the tables are only used for arbitrary categories classes but are not scaled for measurable assessment.

### 3.5. Final Decision and Access Operations

The final decision has four main criteria and information, as listed below. Table 2 depicts the final decision encoding information using hexadecimal digits.

- Access level: Decides the access level for each transaction.
- Access resources: Decides which storage and computation resources will be used.
- Access constraints : Decide what the access constraints are, such as duration, location, number of access trials, and size of data transferred.
- Access operations: Grants access based on the CRUD or HTTP method.

**Table 2.** Final decision encoding.

F	Decision	Hex	Encoding Description
F1	Access Level	2 digits	To specify the five access levels. Ex. 10 for access level L0.
F2	Access resources	6 digits	The first three digits are for compute resources and the rest are for storage resources metadata.
F3	Access constraints	16 digits	8 digits for time, and the other eight digits for other constraints. Ex. 6421EC5F for 2023 Y, 03 M, 27 D, 21 h, 19 mm, 59 ss.
F4	Access operations	1 digit	For example, F is in hexadecimal to represent admin access of all operations

Algorithm 2 shows the logical process of the proposed framework. The framework has three inputs:  $x$ ,  $y$ , and  $z$ . The initial step requires passing the threshold for  $CT$  and  $BT$  that is specified by the system admin. Typically,  $CT \geq 99.99\%$ ,  $BT \geq 0.7$ , where each attribute in  $BT_i \geq \theta$ . If the score of  $CT$  and  $BT$  is zero, then access is denied; for any value between zero and the threshold, the access request should be verified again within a given time interval. The final access decision is granted based on the assessment of the trust scores.

---

**Algorithm 2** Logical process of the proposed access management framework.
 

---

**Input:** User ( $x$ ), IoT hardware ( $y$ ), IoT output data ( $z$ )

**Trust assessment:** Critical trust ( $CT$ ), Bond trust ( $BT$ ), Trust threshold ( $Th$ )

```

1: if  $CT = 0, BT = 0$  then
2:   Deny access
3: end if
4: while  $Th \neq 0$  do
5:   if  $CT \geq Th, BT \geq Th$  then
6:     Initially accepts access
7:   else if  $CT < Th, BT < Th$  then
8:     Verify access again
9:   end if
10: Output: Grant final access decision
11: end while

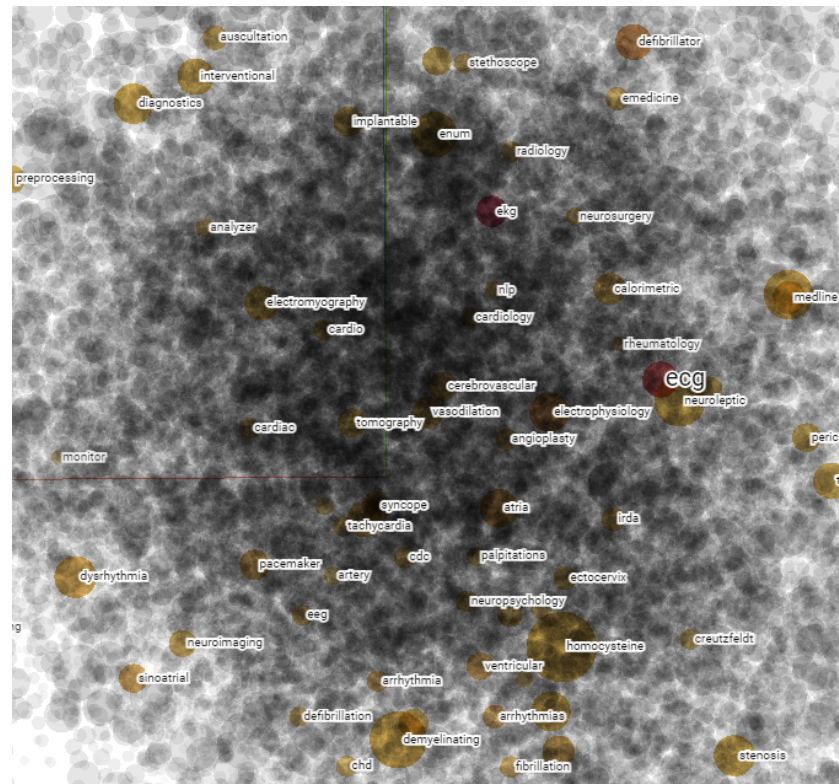
```

---

## 4. Experiments and Results

### 4.1. Dataset Information

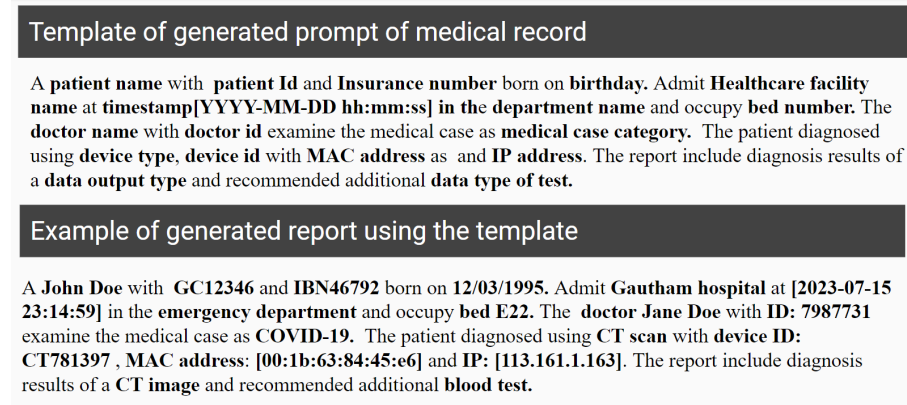
The dataset used in this experiment contains a synthetically generated set of attributes for users, medical IoT devices, and data. The generated data were synthesized using Synthea [47]. The synthetic dataset was then used to fine-tune the Word2Vec model in order to enhance the results within the healthcare information system. Figure 8 shows some examples from the generated dataset, including attributes of the users, devices, and data used to prevent medical errors.



**Figure 8.** Example of selected attributes from the generated data using Synthea and fine-tuned Word2Vec pre-trained model. The figure is a snapshot from a multi-dimensional representation of a large language model data that is represented in latent space. Each attribute is defined as a scalar vector and the distance between each vector is measured by the cosine function.

The synthetic dataset is not limited to only a small subset of attributes. These attributes were extracted from large language models, and include all attributes related to the healthcare information system for the three main categories of user, device, and data. Some examples of these categories are listed in Table A3.

In order to evaluate the syntactic information, we generated a sample of different patient reports based on a predefined template from the same set of attributes mentioned before. The templates were used to generate all possible syntactically and semantically similar reports that could be related to the patient based on their medical history. Figure 9 depicts the template used to generate the final report that includes information about users, devices, and data.



**Figure 9.** Arbitrary example of generated text prompt from patient history record. The mentioned names are arbitrary examples, and do not refer to any true identities.

#### 4.2. Experimental Results and Discussion

The results of our ablation study were examined to evaluate the accuracy of identifying medical errors using the proposed model by examining the relationship between different attributes based on the critical and bond trust scoring. The study was conducted using 17,625 attributes for the user, medical IoT device, and data output categories. The F1-score was 93.5%, which means that the proposed methodology is valid for identifying the relationship between different attributes within the healthcare information system and alleviating any medical errors that may produce false medical reports. Figure 10 depicts the confusion matrix of the experiment results.

	Positive	Negative	Evaluation category	Result
Positive	TP 8868	FP 763	F1-score	93.5%
Negative	FN 471	TN 7523	Precision	92%
			Recall	95%

**Figure 10.** Confusion matrix for the ablation study on the accuracy of detecting medical errors by identifying the relationship between selected attributes. TP is true positive, FP is false positive, FN is false negative, and TN is true negative.

The evaluation of the final results uses different criteria for semantic and syntactic information. Semantic information analysis is used to evaluate the relationship of each word in the context of medical-related data, while syntactic information analysis is used to

evaluate the corpus context of the newly generated report and compare it with the stored data in a healthcare information system. Table 3 lists some examples of the evaluation of the  $BT_A$  for extracting semantic relationships from the healthcare information system for different medical specialty classes.

**Table 3.** Precision, recall, and f1-score for evaluating  $BT_A$  on a selected variety of specialty cases.

Speciality Class	Precision	Recall	F1-Score
Radiology	0.880	0.880	0.880
Gynecology	0.773	0.840	0.805
Oncology	0.793	0.772	0.782
Dermatology	0.712	0.740	0.726
Cardiology	0.833	0.871	0.851
Urology	0.765	0.724	0.744
Emergency	0.865	0.834	0.849
Dentistry	0.79	0.77	0.779
Psychology	0.766	0.784	0.774

Table 4 shows the effect of syntactic analysis on the final decision when using different measures. The table shows the 1-gram measure of the accuracy when detecting one word compared to the context of reference length from the stored data in the healthcare system. While it has high precision, it is not accurate for making decisions, as it does not account for the relationship with other attributes. The decision confidence increases gradually with reference to the n-gram rank, as it has a more meaningful meaning.

The BLEU score can be used for judging a corpus of attributes, but performs badly on single entries. In the case of syntactic analysis, it is more efficient for scoring the generated reports; however, it is not efficient for judging semantic information or detecting sentences with grammatical errors.

The proposed method obtained the best results, as it accounts for both semantic and syntactic information. The decision engine in the cloud generates different reports from the stored data that account for different synonyms, words, or attributes related to the stored data. It can also fix any grammatical errors in the entry and suggest an attribute within the same context. This provides the method with a generalized capability to assess any new report or data entry within the healthcare information system through distributed users, devices, and sorted data. At the same time, the proposed method accounts for security measures that require authentication, authorization, encryption, and logging. Table 4 compares the confidence score of the proposed method with other scoring metrics that are used for syntactic analysis.

**Table 4.** Comparison of the proposed scoring method with other metrics for syntactic analysis.

Metric	Decision Confidence Score
1-g	26%
2-g	33%
3-g	47%
4-g	66%
BLEU	71%
Proposed method	89%

The proposed framework focuses on the cloud–AI access control system by managing access to the cloud resources for users, devices, and data. This is accomplished by

implementing a zero-trust context-aware system that always analyzes the data for each transaction and for the uses of each of the users, devices, network, workload, and data. The framework considers an information security model with three pillars of data confidentiality availability. Regulatory compliance is another part of the proposed context-aware access control. HIPAA is the most important compliance factor that identifies protected health information data. Protected data cannot be used without following a series of security and privacy protection guidelines, including patient consent, disclosure agreement, de-identifications, data encryption, and a well-managed access control system.

The framework also takes advantage of cloud-native microservices to implement critical trust assessment criteria. To build a chain of trust between different attributes for each component, the framework proposes a bond trust evaluation approach inspired by large language models.

Table 5 shows a sample of results for access management decisions based on evaluation of the critical trust (*CT*) and bond trust (*BT*) assessments.

**Table 5.** Example of an access management decision based on scoring evaluation.

<i>N</i> (Samples)	<i>CT</i> (avg.)	<i>BT</i> (avg.)	Decision
402	0	0	Decline
267	0.99	0	Decline
224	0.99	0.5	Verify
259	0.99	0.9	Accept
110	0.99	0.83	Accept
479	0.99	0.79	Accept

While the zero-trust context-aware system is robust against different situations, there are different challenges and limitations that apply when implementing it in the healthcare industry:

- **Data privacy and security:** ML acts as a backbone of the zero trust access control system, which requires training on a considerably large dataset; the size required to obtain efficient results may be in the millions or even billions of parameters. In addition, obtaining sensitive and accurate data is challenging due to privacy concerns around health information regulatory compliance, which may limit the accuracy of the system.
- **Complexity:** The complex healthcare IT infrastructure makes it difficult to implement and manage a zero-trust context-aware access control system. These systems need to be able to integrate with existing systems and applications, and they need to be able to handle the large volume of data generated in healthcare settings.
- **Cost:** Implementing a zero-trust access control system requires an enormous investment in back-end infrastructure. The costs of implementing and maintaining these systems need to be balanced against the potential benefits, such as improved data security and reduced risk of data breaches, as compared to the cost of investment.
- **Skills:** Zero-trust principles rely on many factors. These factors should be aligned with the current and most advanced technologies; this requires highly skilled professionals, who are always in demand due to the absence of these skills in most employees.

**Future considerations:** The current research implements a zero-trust context-aware system to minimize medical errors by analyzing the data context among users, devices, and data. The current model utilizes a fine-tuned Word2Vec model to analyze different attributes. To improve the current algorithm, it is recommended that future work should consider more secure protocols such as data encryption, blockchain technology, and the use of larger language models to improve accuracy. Using larger models such as GPT4, Gemini, Mistral, Llama, or Claude could improve the accuracy of the current implementation. In addition, utilizing these models could increase generalization capabilities as well, as



these models are trained on larger data sources. Additional recommended future work for securing medical IoT devices involves employing the Physical Unclonable Function (PUF) to authenticate the hardware used for telehealth distributed devices [48]. PUF is attracting more legitimate attention and has even been adopted by the US Presidential Administration as a recommended technology for securing IoT devices.

## 5. Conclusions

In this research, a zero-trust framework has been designed to alleviate medical errors within healthcare information systems. The proposed framework implements a theoretical scoring assessment criteria and uses a synthetic dataset derived from a de-identified dataset. The scoring assessment uses critical trust based on cloud microservices and a derived bond trust that assesses the trust cycle between the mutual relationship among users, devices, and output data. The healthcare-related attributes were derived using a pre-trained Word2Vec model. The language model allows for processing multilingual datasets and provides a generalized capability to process healthcare information systems. The designed zero-trust framework can be used in practical applications to enhance the security of the healthcare system, and can also help to ensure that the generated medical reports are consistent and safe for patients. The semantic and syntactic analyses also help to protect healthcare professionals against medical errors, which can reduce the pressure, time, and legal consequences for healthcare service providers. Future work should focus on utilizing large language models trained on larger datasets.

**Author Contributions:** Conceptualization, K.A.-h., A.K. and F.G.; methodology, K.A.-h.; software, K.A.-h.; validation, K.A.-h.; formal analysis, K.A.-h.; investigation, K.A.-h.; resources, K.A.-h.; data curation, K.A.-h.; original draft preparation, K.A.-h.; review and editing, F.G., A.K. and K.A.-h.; visualization, K.A.-h.; supervision, F.G. and A.K.; funding acquisition, F.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by a grant from the National Research Council of Canada (NRC) through the Collaborative Research and Development Initiative.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data is available upon request from the correspondence author or cited work in the article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
AIDA	Artificial Intelligence and Data Act
ABC	Attribute-based Access Control
BT	Bond Trust
CT	Critical Trust
CPPA	Consumer Privacy Protection Act
CML	Cloud and Machine Learning
CRUD	Create, Read, Update, Delete
DDoS	Distributed Denial-of-Service
GDPR	General Data Protection Regulation
IoT	Internet of Things
ML	Machine Learning
PIPEDA	Personal Information Protection and Electronic Document Act
PIDPTA	Personal Information and Data Protection Tribunal Act
RBC	Role-Based Control
SQL	Structured Query Language

## Appendix A. Access Control System Information

### Appendix A.1. Role-Based Access Control System: Important Factors

**Table A1.** Examples of role-based access control system data source information.

Information Source	Example of Data Source Information Factors
Client or operator	User id, role, department, level of access, geographic location
Patient	Patient ID, clinical condition, department, family doctor name, patient consent policy
Resources	Confidentiality, sensitivity, type of data, date ranges covered by the data, author of the data
Data context	System identity, transaction time, the expiration time of token data, the scope and purpose of the token, security of transaction

### Appendix A.2. HIPAA 18-Restricted Patient Identifiers

**Table A2.** HIPAA 18-protected patient identifiers [46].

Iden.	Description	Iden.	Description
1	Names	10	IP address
2	SN (Social security number)	11	Medical records number
3	Geographic locations smaller than states	12	Biometrics identifiers
4	Telephone numbers	13	Health plan beneficiary numbers
5	Fax numbers	14	Full face photographs
6	Devices IDs and serial numbers	15	Account numbers
7	Email address	16	Any other unique identifying numbers
8	Web URLs	17	Certificate; license numbers
9	Vehicle identifiers (e.g., license plate)	18	All element of dates (Except years)

### Appendix A.3. Data Category Examples

**Table A3.** Examples of some attributes from the generated dataset categories.

Attribute Example	Category
Position	User
Department	User
Speciality	User
ID	User
Insurance number	User
User access level	User
User consent	User
Password	User
Manager ID	User
MAC address	Device
IP address	Device
Device model	Device
Device type	Device
Device location	Device
Device manufacture	Device
Data category	Data
Data encryption	Data

Table A3. Cont.

Attribute Example	Category
Data storage location	Data
Storage type (Ex. Container, SSD, VM...)	Data
Data sensitivity level	Data
Data compliance	Data

## References

- Cousins, G.; Durand, L.; O’Kane, A.; Tierney, J.; Maguire, R.; Stokes, S.; O’Reilly, D.; Arensman, E.; Bennett, K.E.; Vázquez, M.O.; et al. Prescription drugs with potential for misuse: Protocol for a multi-indicator analysis of supply, detection and the associated health burden in Ireland between 2010 and 2020. *BMJ Open* **2023**, *13*, e069665. [CrossRef] [PubMed]
- Islam, A.R.; Khan, K.M.; Scarbrough, A.; Zimpfer, M.J.; Makkena, N.; Omogunwa, A.; Ahamed, S.I. An Artificial Intelligence-Based Smartphone App for Assessing the Risk of Opioid Misuse in Working Populations Using Synthetic Data: Pilot Development Study. *JMIR Form. Res.* **2023**, *7*, e45434. [CrossRef]
- Volovici, V.; Syn, N.L.; Ercole, A.; Zhao, J.J.; Liu, N. Steps to avoid overuse and misuse of machine learning in clinical research. *Nat. Med.* **2022**, *28*, 1996–1999. [CrossRef] [PubMed]
- Nancy, A.A.; Ravindran, D.; Raj Vincent, P.D.; Srinivasan, K.; Gutierrez Reina, D. Iot-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning. *Electronics* **2022**, *11*, 2292. [CrossRef]
- Valizadeh, M.; Parde, N. The AI doctor is in: A survey of task-oriented dialogue systems for healthcare applications. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, Dublin, Ireland, 22–27 May 2022; Volume 1: Long Papers; pp. 6638–6660.
- Loh, H.W.; Ooi, C.P.; Seoni, S.; Barua, P.D.; Molinari, F.; Acharya, U.R. Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022). *Comput. Methods Programs Biomed.* **2022**, *226*, 107161. [CrossRef] [PubMed]
- Chauhan, S.; Tanwar, H.K.S. Application of Blockchain Technology in Healthcare: A Systematic Review. In Proceedings of the 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAC), Salem, India, 9–11 May 2022. [CrossRef]
- Lakhan, A.; Mohammed, M.A.; Nedoma, J.; Martinek, R.; Tiwari, P.; Vidyarthi, A.; Alkhayyat, A.; Wang, W. Federated-learning based privacy preservation and fraud-enabled blockchain IoMT system for healthcare. *IEEE J. Biomed. Health Inform.* **2022**, *27*, 664–672. [CrossRef] [PubMed]
- Rasool, R.U.; Ahmad, H.F.; Rafique, W.; Qayyum, A.; Qadir, J. Quantum computing for healthcare: A review. *Future Internet* **2023**, *15*, 94. [CrossRef]
- Kumar, A.; Bhushan, B.; Shriti, S.; Nand, P. Quantum computing for health care: A review on implementation trends and recent advances. In *Multimedia Technologies in the Internet of Things Environment*; Springer: Berlin/Heidelberg, Germany, 2022; Volume 3, pp. 23–40.
- Chen, B.; Qiao, S.; Zhao, J.; Liu, D.; Shi, X.; Lyu, M.; Chen, H.; Lu, H.; Zhai, Y. A security awareness and protection system for 5G smart healthcare based on zero-trust architecture. *IEEE Internet Things J.* **2020**, *8*, 10248–10263. [CrossRef] [PubMed]
- Nandagopal, M.; Seerangan, K.; Govindaraju, T.; Abi, N.E.; Balusamy, B.; Selvarajan, S. A Deep Auto-Optimized Collaborative Learning (DACL) model for disease prognosis using AI-IoMT systems. *Sci. Rep.* **2024**, *14*, 10280. [CrossRef]
- Kernberg, A.; Gold, J.A.; Mohan, V. Using ChatGPT-4 to Create Structured Medical Notes From Audio Recordings of Physician-Patient Encounters: Comparative Study. *J. Med. Internet Res.* **2024**, *26*, e54419. [CrossRef]
- Seyghalani Talab, F.; Ahadinezhad, B.; Khosravizadeh, O.; Amerzadeh, M. A model of the organizational resilience of hospitals in emergencies and disasters. *BMC Emerg. Med.* **2024**, *24*, 105. [CrossRef] [PubMed]
- HL7FHIR. 6.1.0 FHIR Security. 2023. Available online: <https://www.hl7.org/fhir/security.html> (accessed on 30 June 2024).
- Zhang, S.; Yang, S.; Zhu, G.; Luo, E.; Zhang, J.; Xiang, D. A Fine-Grained Access Control Scheme for Electronic Health Records Based on Roles and Attributes. In Proceedings of the Ubiquitous Security: First International Conference, UbiSec 2021, Guangzhou, China, 28–31 December 2021; Revised Selected Papers; Springer: Berlin/Heidelberg, Germany, 2022; pp. 25–37.
- Rashid, M.; Parah, S.A.; Wani, A.R.; Gupta, S.K. Securing E-Health IoT data on cloud systems using novel extended role based access control model. In *Internet of Things (IoT) Concepts and Applications*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 473–489.
- Khan, S.; Iqbal, W.; Waheed, A.; Mehmood, G.; Khan, S.; Zareei, M.; Biswal, R.R. An efficient and secure revocation-enabled attribute-based access control for eHealth in smart society. *Sensors* **2022**, *22*, 336. [CrossRef] [PubMed]
- Sanders, M.W.; Yue, C. Mining Least Privilege Attribute Based Access Control Policies. In Proceedings of the 35th Annual Computer Security Applications Conference, New York, NY, USA, 9–13 December 2019; ACSAC’19; pp. 404–416. [CrossRef]
- Nobi, M.N.; Krishnan, R.; Huang, Y.; Sandhu, R. Administration of Machine Learning Based Access Control. In Proceedings of the Computer Security–ESORICS 2022: 27th European Symposium on Research in Computer Security, Copenhagen, Denmark, 26–30 September 2022; Proceedings, Part II; Springer: Berlin/Heidelberg, Germany, 2022; pp. 189–210.

21. Nobi, M.N.; Krishnan, R.; Huang, Y.; Shakarami, M.; Sandhu, R. Toward Deep Learning Based Access Control. In Proceedings of the Twelfth ACM Conference on Data and Application Security and Privacy, Washington, DC, USA, 25–27 April 2022. [CrossRef]
22. Jin, Z.; Xing, L.; Fang, Y.; Jia, Y.; Yuan, B.; Liu, Q. P-Verifier. In Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, Los Angeles, CA, USA, 7–11 November 2022. [CrossRef]
23. Outchakoucht, A.; Hamza, E.S.; Leroy, J.P. Dynamic access control policy based on blockchain and machine learning for the internet of things. *Int. J. Adv. Comput. Sci. Appl.* **2017**, *8*, 7. [CrossRef]
24. Chiquito, A.; Bodin, U.; Schelén, O. Attribute-Based Approaches for Secure Data Sharing in Industrial Contexts. *IEEE Access* **2023**, *11*, 10180–10195. [CrossRef]
25. Hu, V.C.; Kuhn, D.R.; Ferraiolo, D.F.; Voas, J. Attribute-based access control. *Computer* **2015**, *48*, 85–88. [CrossRef]
26. Ghorbani, A.; Lashkari, A.H.; Mamun, M.S.I.; Gil, G.D. Systems and Methods for Cybersecurity Risk Assessment of Users of a Computer Network. U.S. Patent App. 16/753,301, 30 July 2020.
27. Wu, Y.; Li, L.; Xin, B.; Hu, Q.; Dong, X.; Li, Z. Application of machine learning in personalized medicine. *Intell. Pharm.* **2023**, *1*, 152–156. [CrossRef]
28. Al-hammuri, K.; Gebali, F.; Kanan, A.; Chelvan, I.T. Vision transformer architecture and applications in digital health: A tutorial and survey. In *Visual Computing for Industry, Biomedicine, and Art*; Springer: Berlin/Heidelberg, Germany, 2023; Volume 6. [CrossRef]
29. Guo, D. Applying Medical Language Models to Medical Image Analysis. Ph.D. Thesis, UCLA, Los Angeles, CA, USA, 2024.
30. Lu, Z. Multimodal Large Language Models in Vision and Ophthalmology. *Investig. Ophthalmol. Vis. Sci.* **2024**, *65*, 3876.
31. Shapiro, J.; Baum, S.; Pavlotzky, F.; Mordechai, Y.B.; Barzilai, A.; Freud, T.; Gershon, R. Application of an NLP AI Tool in Psoriasis: A Cross-Sectional Comparative Study on Identifying Affected Areas in Patients' Data. *Clin. Dermatol.* **2024**; ISSN 0738-081X. [CrossRef]
32. He, D.; Prabhakaran, T.J.; Wang, E.; Chung, S.T. Analyzing Electronic Medical Records of Low Vision Patients using a Natural Language Processing Framework. *Investig. Ophthalmol. Vis. Sci.* **2024**, *65*, 5472.
33. Wiest, I.C.; Lessmann, M.E.; Wolf, F.; Ferber, D.; Van Treeck, M.; Zhu, J.; Ebert, M.P.; Westphalen, C.B.; Wermke, M.; Kather, J.N. Anonymizing medical documents with local, privacy preserving large language models: The LLM-Anonymizer. *medRxiv* **2024**. [CrossRef]
34. Gismelbari, M.A.; Vixnin, I.I.; Kovalev, G.M.; Gogolev, E.E. Speech Emotion Recognition Using Deep Learning. In Proceedings of the 2024 XXVII International Conference on Soft Computing and Measurements (SCM), Saint Petersburg, Russian, 22–24 May 2024; pp. 380–384.
35. Jiang, H. Research on emotion management for elderly based on speech signal analysis technology. In Proceedings of the Fourth International Conference on Sensors and Information Technology (ICSI 2024), Sydney, Australia, 9–11 December 2024; Volume 13107, pp. 1026–1033.
36. Jin, Y.; Chandra, M.; Verma, G.; Hu, Y.; De Choudhury, M.; Kumar, S. Ask Me in English Instead: Cross-Lingual Evaluation of Large Language Models for Healthcare Queries. In Proceedings of the The Web Conference 2024, Singapore, 13–17 May 2024.
37. Yang, X.; Chen, A.; PourNejatian, N.; Shin, H.C.; Smith, K.E.; Parisien, C.; Compas, C.; Martin, C.; Costa, A.B.; Flores, M.G.; et al. A large language model for electronic health records. *NPJ Digit. Med.* **2022**, *5*, 194. [CrossRef]
38. He, K.; Mao, R.; Lin, Q.; Ruan, Y.; Lan, X.; Feng, M.; Cambria, E. A survey of large language models for healthcare: From data, technology, and applications to accountability and ethics. *arXiv* **2023**, arXiv:2310.05694.
39. Singhal, K.; Azizi, S.; Tu, T.; Mahdavi, S.S.; Wei, J.; Chung, H.W.; Scales, N.; Tanwani, A.; Cole-Lewis, H.; Pfohl, S.; et al. Large language models encode clinical knowledge. *Nature* **2023**, *620*, 172–180. [CrossRef] [PubMed]
40. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.
41. Church, K.W. Word2Vec. *Nat. Lang. Eng.* **2017**, *23*, 155–162. [CrossRef]
42. Hung, P.T.; Yamanishi, K. Word2vec skip-gram dimensionality selection via sequential normalized maximum likelihood. *Entropy* **2021**, *23*, 997. [CrossRef]
43. Habib, M.; Faris, M.; Alomari, A.; Faris, H. Altibbivec: A word embedding model for medical and health applications in the Arabic language. *IEEE Access* **2021**, *9*, 133875–133888. [CrossRef]
44. Pennington, J.; Socher, R.; Manning, C.D. Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
45. Papineni, K.; Roukos, S.; Ward, T.; Zhu, W.J. Bleu: A method for automatic evaluation of machine translation. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, PA, USA, 6–12 July 2002; pp. 311–318.
46. Portability, Insurance, and Accountability Act. *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*; Human Health Services: Washington DC, USA, 2023. Available online: <https://www.hhs.gov/guidance/document/guidance-regarding-methods-de-identification-protected-health-information-accordance-0> (accessed on 5 March 2023).

- 
47. Walonoski, J.; Klaus, S.; Granger, E.; Hall, D.; Gregorowicz, A.; Neyarapally, G.; Watson, A.; Eastman, J. Synthea™ Novel coronavirus (COVID-19) model and synthetic data set. *Intell.-Based Med.* **2020**, *1*, 100007. [[CrossRef](#)]
  48. Gebali, F.; Mamun, M. SRAM Physically Unclonable Functions for Smart Home IoT Telehealth Environments. In *Cybersecurity in Smart Homes: Architectures, Solutions and Technologies*; Wiley Data and Cybersecurity; ISTE Ltd.: London, UK, 2022; pp. 125–154.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.