AI Risks and rai framework for elsba insurance

**MIS716 – Artificial Intelligence for Business - Trimester 1 2021 Assessment Task 2 – Business Report**

MOHITKUMAR PARESHKUMAR KANJIYA

220077053

GULBAZ KHAN

219296087

MALAWARA ARACHCHIGE SANDUNI PIUMALI

219052007

# **Executive summary**

This report is based on a case study by Elsba car Insurance. Elsba car Insurance provides the development and deployment of an intelligent, voice-activated assistant system for automating and customising customer interactions. However, insurance clients in the affected areas have stated that the activation process creates management-based discrimination and racial impact. The purpose of this report is to adopt a responsible artificial intelligence (RAI) framework and ethical principles to minimise risks associated with a broader range of institutional and social situations are to recommend.

The first part of the report applies to the Elsba car Insurance AI Risk Framework for identifying, analysing and assessing the risks of various Elsba's stakeholders. Analysis shows that clients and customers are at increased risk because the proposed AI system uses data, including Elsba identification data, to produce results that can lead to customer privacy and security breaches. The second part of the report applies to Elsba's RAI framework for identifying potential problems and solutions, while the final section of the report focuses on identifying risk areas during the AI ​​development series. The report also describes where the risks occurred at each stage of the AI ​​development and provides recommendations to minimise the problem.

Contents

[Executive summary 1](#_Toc71651580)

[Introduction 3](#_Toc71651581)

[Background of the Case Study 4](#_Toc71651582)

[AI risks: 5](#_Toc71651583)

[**RAI Framework:** **13**](#_Toc71651584)

[Recommendations and limitations: 20](#_Toc71651585)

[References 22](#_Toc71651586)

**List of Tables:**

[**Table 1 - List of Stakeholders 5**](#_Toc71652547)

[**Table 2 - Stakeholders and their risks 12**](#_Toc71652548)

[**Table 3 - RAI Framework (Problems & Solutions) 19**](#_Toc71652549)

[**Table 4 - Risk Mapping 20**](#_Toc71652550)

[**Table 5 - Recommendation Prioritization 22**](#_Toc71652551)

**List of Figures:**

[**Figure 1 - RAI Framework (Source: PwC) 14**](#_Toc71652559)

[**Figure 2 - AI Predictions for 2019 (Source:PwC) 14**](#_Toc71652560)

# **Introduction**

Increasing AI capabilities are changing the way an organisation interacts and works together with internal and external users. Currently, the insurance industry uses AI in several ways: Almost every company uses AI, which opens up numerous opportunities for significant growth in all industries.

The findings of the following research paper show the biased approach of AI during its usage. There are many risks, including privacy and security issues. According to research articles, “85% of CEOs agree that decisions should be made by AI **(Ramnath Balasubramanian, 2018)**. The four emerging Business Models(BM) show: In the first model, the insurer takes a small part of the value chain and allows others with advanced AI and data to acquire a more significant part. In the second model, the insurer keeps the same format and value chain but uses AI to improve efficiency. In the third mode, the insurer uses AI and adapts its layout to find new data sources and customers. Finally, in the fourth model, a technology-focused company uses its existing AI capabilities, advanced data, and a broader customer base and adds insurance funding **(Alex Zarifis, 2019)**.

This case study evaluated the AI ​​implementation of Elsba’s insurance companies. Its purpose is to define the ability to define ethically responsible strategies to reduce the risk to an organised and broader social situation by a Responsible Artificial Intelligence (RAI) framework and ethical principles.

# **Background of the Case Study**

Elsba Insurance is an Australian car insurance providing firm that mainly operates in Eastern Australia. Apart from their corporate website, the company also has a call centre to promote and sell insurance policies. Through this insurance company, personal and business customers can obtain information, purchase, manage and renew an insurance policy annually. Recently, an active, intelligent assistant system was introduced to automate and personalise customer transactions. It was developed jointly by a local software vendor, Mr Misha Hansen and the company's Business Intelligence and Analytics Team. Some executive team members feel that the system lacks transparency in decision-making and has some issues with the customers' data privacy. So the company recruited external consultants to identify and assess the risks associated with insurance. The company reported total revenue of $700 million in 2019.

The objectives of the topic are:

* AI risks analysis with applying AI in a case study presented from the perspective of key stakeholders
* Risk solutions can be explored using the RAI framework and ethical principles by assembling and discussing AI risks with various risk partners
* A series of recommendations to reduce risk

As follows, we analyse the following risks of stakeholders in this case study.

# **AI risks:**

While analysing the proposed AI system, it is crucial to identify the stakeholders who are going to be affected by the system. The different types of stakeholders can be divided into three categories:

1. **Individuals**
2. **Organisation**
3. **Society**

The main three can be further subdivided into:

|  |  |  |
| --- | --- | --- |
| Individuals | Organisation | Society |
| * Clients * Families of Clients * Customer Service Staff | * **Business Intelligence and Analytics Team** * **Elsba Executive Team** | * **Public** |

Table 1 - List of Stakeholders

AI Risks can be identified with the help of the AI Risk Framework proposed by PwC. The framework compares the different types of risks that each stakeholder may face. The risks can then be compared and dealt with according to the severity level of the risk.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CLIENTS** | | | | | | | | | | | | | | | | |
| **AI Risk Domains** | | **Risk Description** | | **AI design/use factor** | | | | | | **Likelihood** | | **Consequence** | | | **Overall Rating** | |
| **Performance** | | **Risk of Error -**  Any error in decision-making leading to the system generating a false positive or true negative decision | | The training data is unsupervised, which may lead to errors in decision making in future insurance claims. | | | | | | **Possible** | | **Major** | | | **High** | |
| **Risk of Bias –**  AI system is making Biased Decision towards any society, caste or racial profile based on the data provided by a particular client during the onboarding process. | | The AI system makes a general biasness towards a particular society or caste because the source of training data can be biased towards some communities. | | | | | | **Almost Certain** | | **Catastrophic** | | | **Very High** | |
| **Risk of Stability of Performance –** Systems can malfunction, affecting the customer service time and causing delays in the claim process | | Electricity fluctuation, improper maintenance service, untimely system updates can cause the system to shut down for sometime | | | | | | **Unlikely** | | **Minor** | | | **Low** | |
| **Security** | | **Risk of Misuse of Data –** Client data can be misused for other marketing purposes or sold to other companies for profits | | Whether it's selling personal information to a third party or producing targeted advertising, some businesses have developed their entire business model around customer data. **(Freedman, 2020)** | | | | | | **Likely** | | **Major** | | | **Very High** | |
| **Risk of Data Stealing –** System can be hacked, and the client data can be stolen | | One of the major concerns for any AI system is Data Security, as clients data is the most critical asset for any company. Any loss in data will lead to a loss in credibility and reputation of the company | | | | | | **Rare** | | **Major** | | | **Low** | |
| **Control** | | **Risk of Data Privacy –** The client data collected by the system can breach some data privacy rights | | Personal information extracted from various websites may lack controls resulting in privacy breach of client data **(Stefan Larsson, 2020)** | | | | | | **Possible** | | **Major** | | | **High** | |
| **FAMILIES OF CLIENTS** | | | | | | | | | | | | | | | | |
| **AI Risk Domains** | **Risk Description** | | | | **AI design/use factor** | | | | **Likelihood** | | | | **Consequence** | | | **Overall Rating** |
| **Performance** | **Risk of Error -**  Any error in decision-making leading to the system generating a false positive or true negative decision for a client after a major car accident | | | | After an untimely death of a client, the client's family member may try to claim the insurance. But the system can make an error by not considering the family members and refusing the claim | | | | **Rare** | | | | **Minor** | | | **Very Low** |
| **Legal And Compliance** | **Risk of Opaqueness –**  The system can operate in a certain way and decide on its own to accept or reject a claim | | | | The algorithm can sometimes refuse an insurance claim based on the relation between the client and the family member trying to claim the insurance | | | | **Likely** | | | | **Moderate** | | | **High** |
| **CUSTOMER SERVICE STAFF TEAM** | | | | | | | | | | | | | | | | |
| **AI Risk Domains** | | | **Risk Description** | | | | **AI design/use factor** | | **Likelihood** | | | **Consequence** | | | **Overall Rating** | |
| **Performance** | | | **Risk of Error -**  Any error in decision-making leading to the system generating a false positive or true negative decision | | | | The decision for a complex case would be firstly taken by the Elsba agent and transferred to the service team if necessary. The agents decision may influence the conclusion of a service staff member, which can be misleading. | | **Likely** | | | **Moderate** | | | **High** | |
| **Risk of Bias –**  AI system is making Biased Decision towards any society, caste or racial profile based on the data provided by a particular client during the onboarding process. | | | | The AI system makes a general biased decision towards a particular society or caste because of the source of training data. The staff member can also agree to the decision due to their behaviour. | | **Almost Certain** | | | **Catastrophic** | | | **Very High** | |
| **Risk of stability of Performance –** The system can cause delays in feeding data to the customer service team leading to slow customer services | | | | While handling complex situations, the customer service team staff would rely on the data provided by the agents. Thus, any fluctuation in the system may slow down the process | | **Rare** | | | **Insignificant** | | | **Very Low** | |
| **Economics and Financial Performance** | | | **Risk of Job Displacement –** The use of an AI system may lead to job loss for some customer service staff members | | | | The Elsba agents will replace many staff members as most of the work, including onboarding, till insurance claims have diverged to the AI system | | **Unlikely** | | | **Minor** | | | **Low** | |
| **Societal** | | | **Change of Power Distribution –** AI system would replace the decision-making process, thus transferring power from a human to a machine | | | | The risk of AI making a false decision can happen when the decision making power is given to a machine that will not consider any human logic. | | **Rare** | | | **Moderate** | | | **Low** | |
| **BUSINESS INTELLIGENCE AND ANALYTICS TEAM** | | | | | | | | | | | | | | | | |
| **AI Risk Domains** | | | **Risk Description** | | | **AI design/use factor** | | **Likelihood** | | | **Consequence** | | | **Overall Rating** | | |
| **Performance** | | | **Risk of Error -**  The developers of the AI system can make some error like the errors which can occur using the wrong algorithms in Natural Language Processing and Machine learning. | | | There is a high chance in software and AI development to make errors if the system being developed is not tested or due to the ignorance of developers. | | **Rare** | | | **Catastrophic** | | | **Tolerable** | | |
| **Risk of Bias –**  The data used for development can be biased, so a whole community can be ignored if diverse data is not considered. | | | The developing team is responsible for the collection of fair data which does not favour any particular community. The chances are very high as data is not readily available, affecting the voice command system. | | **Likely** | | | **Major** | | | **Very High** | | |
| **Security** | | | **Privacy risks –**For AI risk assessment, the networks and the security design designed by the team cannot be neglected. | | | While developing and implementing this system, it is imperative to consider secure networks to secure the data within the company as intruders can steal the company data if insecure networks are used. | | **Possible** | | | **Moderate** | | | **Tolerable** | | |
| **Control** | | | **Risk of AI going Rouge –** This is a part of the testing phase for compliance which is the responsibility of developers and the quality assurance team. | | | It is necessary to test the AI system on every stage of development as it will cause a severe issue if the team ignores it at an early stage. This risk should be assessed as a part of our risk assessment strategy. | | **Rare** | | | **Minor** | | | **Very Low** | | |
| **Liability and Reputation risk –** This risk is associated with the developing and analytics team representing the Elsba insurance. | | | If the analysts and developers make any mistake in the development and the users, find it so the company's reputation will be destroyed, resulting in losing clients. This risk should be assessed. | | **Unlikely** | | | **Insignificant** | | | **Very Low** | | |
| **Societal** | | | **Risk of Intelligence Divide –** The collection of data is the responsibility of the developing team. Training the model using the information is a crucial part of developing. | | | Suppose the data is collected from one neighbourhood for training purposes. This will favour one community in voice control and also in decision making, so there is a high chance other community may not like it. | | **Rare** | | | **Major** | | | **Low** | | |
| **ELSBA EXECUTIVE TEAM** | | | | | | | | | | | | | | | | |
| **AI Risk Domains** | | | **Risk Description** | | | **AI design/use factor** | | **Likelihood** | | | **Consequence** | | | **Overall Rating** | | |
| **Performance** | | | **Risk of Error -**  The errors which can occur due to the wrong implementation of the AI system | | | There is a possibility of implementing the wrong AI system at the wrong place, such that the executive team can mix up the onboarding and renewal. That’s why the risk assessment of this error is mandatory. | | **Possible** | | | **Catastrophic** | | | **Very High** | | |
| **Risk of Stability of performance**  This risk judges the credibility of executing the AI systems in the real world scenario. The whole organisation will take part, including the executive team, to test the performance. | | | Once executed successfully in the company, the AI system is hard to manage, and we will have to assess the performance of the system time to time to check the stability when the number of user increase or the voice command is activated in a different accent. We will assess the risk of stability of performance with these aspects. | | **Possible** | | | **Moderate** | | | **Tolerable** | | |
| **Risk of Opaqueness**  This risk is associated with the transparency of the functions within the organisation. The Elsba claim and renewable are highly at a greater risk to be a hunt of this opaqueness about the tasks after the AI has done its work. | | | As many machine learning algorithms are being used in the AI system, Blackbox risk can occur as these systems can learn a different pattern than expected one to predict the likelihood for the clients. On the other hand, The executive team can make different decisions after investigation, doubting the AI systems' transparency. | | **Unlikely** | | | **Major** | | | **Tolerable** | | |
| **Economics and Finance** | | | **Risk of job displacement**  These risk are the ones which are an impact on the executive team which are mostly likely to be ignored by many companies while implementing the AI. | | | If the system is executed successfully, the AI system will do a job that many people in the customer service already did. We have to keep in mind about the staff to see if their work is needed anymore or not so many people can lose their job. | | **Rare** | | | **Catastrophic** | | | **Tolerable** | | |
| **Reputation Risk**  This risk is associated with implementing the AI system for customer interactions. The reputation of the company is at stake over here. | | | The organisation's reputation will be questioned if the AI system cannot provide the desired accuracy of the results. This error can result from the negligence of the testing phase by the executives of the company. | | **Almost Certain** | | | **Catastrophic** | | | **Very High** | | |
| **Ethical** | | | **Lack of values risk**  This risk aligns with the company's values, which it promises to deliver to its customers. | | | Previously, people were able to talk, and the emotions of the people were being delivered; with AI, it will be hard to understand people's emotions. There are many who need interpretation on the phone. This risk is essential to investigate to cater the disable people as the company’s values can affected. | | **Likely** | | | **Moderate** | | | **High** | | |
| **GENERAL PUBLIC** | | | | | | | | | | | | | | | | |
| **AI Risk Domains** | | | **Risk Description** | | | **AI design/use factor** | | **Likelihood** | | | **Consequence** | | | **Overall Rating** | | |
| **Performance** | | | **Risk of Bias**  The risk of biasness effect the general public if the data collection tool is already biased and favouring a specific community. | | | The algorithms cannot be fully accurate as we will need a very diverse data which is very costly to get. This risk can also occur if the general public is checking their likelihood of insurance or accident. They may expect some positive result but due to the biasness the Ai system gives them different result than expected. | | **Likely** | | | **Catastrophic** | | | **Very High** | | |
| **Security** | | | **Privacy risks**  The data which is collected from the public is open to intrusions attack leading to the leakage of data which is a privacy breach when it comes to privacy rights. | | | The AI system will be collecting the data of general public which will be used by the algorithms to generate a specific output and store the results in the company servers. The servers can be hacked and data can be leaked if any attempt to breach the security is made. | | **Unlikely** | | | **Major** | | | **Tolerable** | | |
| **Control** | | | **Risk of AI going Rouge**  The data of public is stored in the AI system which can be misused if the system crashes. | | | The system can crash due to the overload or the data can. Be misused without the permission of public if there is some problem is AI system. Control measures should be taken place to assess this risk. | | **Rare** | | | **Major** | | | **Low** | | |
| **Economic** | | | **Job displacement risk**  Many people which are a part of this organisation also and the related institutions can lose their job due to this AI. | | | This risk can be analysed if proper investigation is done by the company for the impact of this system on the jobs of the people already working in this department.  The institutions may also be affected like the replacement car company as the system will not give everyone false claim so they will have a loss or increase of income due to the incompetency of the Elsba Claims. | | **Likely** | | | **Catastrophic** | | | **Very High** | | |

Table 2 - Stakeholders and their risks

Levels of severity of the overall rating can be expressed through different colours:

|  |  |
| --- | --- |
| **Severity Level** | **Representing Color** |
| Very High |  |
| High |  |
| Tolerable |  |
| Low |  |
| Very Low |  |

## **RAI Framework:**

Stakeholders such as Elsba Executive Team and the developers know a lot about how AI and data are used, developed, and managed to mitigate the risks. You need to apply the RAI framework to your AI system. In addition, the use of artificial intelligence technology (RAI) is the use of ethical, transparent, and responsible AI technology in accordance with the values, rules, and social norms desired by users.

The potential for harm intended or unintended arising from algorithmic decision-making indicates that an ethical dimension is needed. **(Richard Vidgen, 2019)**

The Responsive Artificial Intelligence(RAI) Toolkit is a set of frameworks, tools, and processes designed to harness the power of AI, from ethics and strategy to implementation. There are use RAI framework solutions that organisation's needs benefit in and mitigate risk such as against AI algorithms to ensure the confidentiality and security of accurate and interpretable automated diagnostic data.

The following RAI framework, as shown in the figure developed by PWC, applies to the trolley and analyses its risks, and seeks solutions:

Text, application, whiteboard

Description automatically generated

**Figure 1 - RAI Framework (Source: PwC)**

Organisations in the global aspect need to adopt RAI for a few collection of expectations to be addressed and fulfilled. Verification and validation of security aspects will be vital in the scenarios to be considered related to automation. Data protection and legality of data usages are considered as 64% of the expectations to be addressed. Globally more than half of the business expected reliable and adaptable algorithms to be exercised. These systems to be ethically correct and transparent. AI always focus on maintaining good governance practices by example. The algorithms are to be, and machine learnings also help to keep standardisation as expected.

A picture containing application

Description automatically generated

**Figure 2 - AI Predictions for 2019 (Source:PwC)**

|  |  |  |
| --- | --- | --- |
| Identified Risk | RAI Framework Alignment | Problem Statement and Potential Alignment |
| Risk of Error | **Performance (Robustness)** | **Problem:**  The AI Algorithm can be implemented incorrectly or can make errors due to false or unsupervised data |
| **Solutions:**  Implementation of AI can be reviewed and closely monitored with a constant feedback process.  The quality of data can be improved or bought from a verified source. |
| Risk of Bias | **Performance (Bias and Fairness)**  **Ethics and Regulations** | **Problem:**  AI algorithm makes false positive or true negative decisions based on biased unsupervised data towards specific communities |
| **Solutions:**  The training data needs to inspected by a neutral body before used in the pilot phase of training.  Identify the training data (e.g., biased historical decisions) that are most responsible for the unequal decisions taken by the AI system, and then delete the biassed training points using impact functions. **(Sahil Verma, 2021)**  A team of experts should judge the AI system before final implementation to remove any biased judgements. |
| Risk of Stability of Performance | **Performance (Robustness & Security)**  **Performance (Interpretability & Explainability)**  **Governance** | **Problem:**  AI system can make unexplainable decisions leading to performance issues. Faulty governance can lead to terrible decisions. |
| **Solutions:**  All stakeholders to be taken into consideration before final implementation  Governance body should be set up to look over the AI system at all times  Ethical Policies to be drafted and strictly followed |
| Risk of Opaqueness | **Ethics and Regulation** | **Problem:**  AI system lacks transparency in decision making and leads to conflicts during the final insurance claim decisions |
| **Solutions:**  All departments and executive team should have a common understanding of how the final decisions are going to be taken by the AI system  Constant monitoring of the system with regular OTA updates |
| Risk of Misuse of Data | **Performance (Robustness and Security)**  **Ethics & Regulation** | **Problem:**  The collected data can be misused for other marketing purposes leading to ethical issues with public |
| **Solutions:**  Agreement notice signed by the customer for how the data will be used by the company  Building or adopting an ethical framework for data use and AI is, therefore, critical in trying to stop poor use of data hitting the headlines **(Armstrong, 2016)** |
| Privacy Risks | **Performance (Robustness & Security)** | **Problem:**  The AI system can be attacked by the intruders to steal and misuse the data. |
| **Solutions:**  The solution can be seen in this paper Technical control **(Australia Gov, n.d.)**;   * Information Privacy and Data Protection Act as stated should be used. * Encryption of data to avoid third-party interference. * Regular update Privacy and security policies.   The company can also provide limited and restricted access to the employees and have a firewall for extra protection of the servers. |
| Risk of AI going rogue | **Control (Robustness & Security)** | **Problem:**  The AI system can become uncontrollable, stop working or does not self-learn so resulting in harmful activities with the user and their data. |
| **Solution**   * Elsba insurance can add an additional AI application to check the accuracy of the output **Accenture (2018)** * Quality assurance team of Elsba insurance should test the system time to time and defince its compatibility to the users.   Use of third-party services for ectra protection of the AI system |
| Liability and Reputation Risk | **Economic (Ethics & Regulation)** | **Problem**  AI system can have a mistake in the development and it is found by the users so the reputation of company will be destroyed and legal actions can be taken against the company.  It can aslo have reputation risk of racism in the community. |
| **Solution**   * Data should take into account different criteria, geographical groups, genders etc. **PwC (2019)**. * Do not give all the power to AI to avoid the legal actions against it. * Avoid using one source of data to rely on the results. Multiples data souces should be used with diversity. |
| Risk of Job Displacement | **Economic (Ethics & Regulation)** | **Problem**  The people working in the organisation can lose their jobs as AI will reduce the work or if there is an error in the AI. The institutions which are related to the Elsba Insurance can have an effect on their income. |
| **Solution**   * Train employees to work and develop new AI technologies. **Accenture (2018)** * Workforce Reskilling **(Taulli, 2019)** * Company should be prepared to address concerns raised by small businesses which are affected. * Conduct workshops |
| Lack of Values Risk | **Ethical (Bias and Fairness/Interpretability and Explainability/Ethics & Regulation)** | **Problem**  AI system can make biased outcome based on ethnicity, location and voice recognition. |
| **Solution**   * Data should take into account different criteria, geographical groups, genders etc. **PwC (2019)**.   Develop AI ethic policies incorporating **(Australia Gov, 2020)**;   * Technology and human rights. * OECD Principal Act * Data rights of consumers   Company can also :   * Assign AI ethical executive. * Incorporate legal, privacy, fairness, transparency and accountability into ethical framework. * Determine a level of acceptable bias and then a level that will not be acceptable to its key stakeholders. |
| Change of Power Distribution Risk | **Societal (Governance)** | **Problem**  Elsba claims and onboarding would replace the decision-making process, thus transferring power from a human to a machine |
| **Solution**   * The company should make a new department to look over the decisions made by the AI system. * Control measures should be taken into account to restrict the power of AI system. For example for small claims AI algorithms can be given full power but for major claims, a case officer should be assigened to overlook the claim. * The range of the amount of claim can be decided for the company and an alert system can be used if case officer is needed thus efficiently using the work force. |

Table 3 - RAI Framework (Problems & Solutions)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **RAI FRAMEWORK DIMENSION** | | | | | | | |
| **AI RISK DIMENSION/TYPE** |  | | **GOVERNANCE** | **PERFORMANCE** | | | **ETHICS & REGULATION** |
| **Interpretability & Explainability** | **Bias & Fairness** | **Robustness & Security** |
| **Performance** | **Risk of Errors** |  |  |  | **Checkmark** |  |
| **Risk of Opaqueness** |  |  |  |  | **Checkmark** |
| **Risk of Instability of Performance** | **Checkmark** | **Checkmark** |  | **Checkmark** |  |
| **Risk of Bias** |  |  | **Checkmark** |  | **Checkmark** |
| **Security** | **Risk of Misuse of Data** |  |  |  | **Checkmark** | **Checkmark** |
| **Privacy Risk** |  |  |  | **Checkmark** |  |
| **Control** | **Risk of AI going "Rogue"** |  |  |  | **Checkmark** |  |
| **Economic** | **Liability Risk** |  |  |  |  | **Checkmark** |
| **Risk of Job Displacement** |  |  |  |  | **Checkmark** |
| **Societal** | **Change of Power Distribution** | **Checkmark** |  |  |  |  |
| **Ethical** | **Lack of Values Risk** |  | **Checkmark** | **Checkmark** |  | **Checkmark** |

**Table 4 - Risk Mapping**

# **Recommendations and limitations:**

The following recommendations were created assuming that the initial AI system implementation of virtual assistants of Elsba (3 agents) are some of the best practices used to ensure virtual assistance service security. Simple use of virtual assistant or chatbot is a software that uses machine learning (ML) to communicate with other users over the Internet (also known as online chat).

While virtual assistants service is undoubtedly an attractive and innovative development of numeral interaction among customers and companies, it also gives hackers another chance to find access for the website to personal and confidential info. Therefore, like all aspects of security when accessing a website, virtual assistant (chatbot) security is in your hands. The more layers of website security you enable, the harder it is for cyber-criminals to sacrifice your site and visitors.

The best practise of virtual assistant services is given two security developments for virtual assistants are authentication and authorisation. The first is to verify the identity of the user and the second is to allow specific users to perform specific functions and tasks or access the portal.

**Authentication timeout:** This security application specifies the length of time an authentic user can wait to "sign-in". May you have seen this on your Elsba car insurance website. A popup agency will ask you to sign in again, asking you to confirm that you are still active, or that timed out. This will prevent cybercriminals from having time to access someone's secure account.

**Two-factor Authentication:** The time tested of security system requires identify in uses of two different ways. For example, Elsba car insurance customer’s using a username, password, and then responding to feedback query referred to the user by a unique email.

**Use a Web Application Firewall (WAF):** WAF defends websites from harmful data- traffic and malicious requirements. Thus, WAF can help prevent malicious code from insurance details entering the virtual assistants service of customer iframe.

**User IDs and Passwords:** As a replacement for allowing any customer to use virtual assistants service, ask them to turn out to be an insured record user. Easy target of criminals for data hacking system. Therefore, an extra step like record-keeping on a website can prevent potential cyber-criminal for customer data.

**End-to-End Encryption:** Anyone prevents this, other than the sender and management from seeing any part of business communication.

In addition, general recommended points are face-to-face chat with tools that will assistance you get quick responses to difficult queries. For instance of internal and Elsba allow to representatives ask to each other for assistance. Without interrupting the flow of server conversation sessions. This means that sub-administrators will be able to respond to delegates at the beginning of the session. When you need to go to second-level support for chats, chat switching is always an option without voice activation.

This means that sub-administrators will be able to respond to delegates at the beginning of the session. When you need to go to second-level support for chats, chat switching is always an option without voice activation. Therefore, customers can always turn off the voice-activated option and have a text-based conversation with Elsba. The Elsba agents guide customers through their respective standard processes and transfer the call to the customer service staff when recognising certain complex conditions.

Mitigating the AI-risk, regardless of when AI tools were used to validate new job applications such as virtual assistant and business analysis tools for the first interview, all companies should be familiar with AI and secure human resources to provide robust results. So, conduct periodic risk assessment reviews (meetings), collect customer feedback after they interact with the AI system (user experience, ratings for service), assign a risk manager to oversee the AI systems security and risks, and run a pilot system implementation before the actual implementation of the system.

|  |  |
| --- | --- |
| **Recommendations** | **Priority** |
| **Authentication timeout** | **Moderate** |
| **Two-factor Authentication** | **High** |
| **Use a Web Application Firewall (WAF)** | **Very High** |
| **User IDs and Passwords** | **Low** |
| **End-to-End Encryption** | **High** |

Table 5 - Recommendation Prioritization

# References

Alex Zarifis, C. P. H. A. M., 2019. *Evaluating the impact of AI on insurance: The four emerging AI- and data-driven business models,* s.l.: Emerald Open Research.

Armstrong, H., 2016. *We need to work out the rules for AI before data misuse hits the headlines.* [Online]   
Available at: https://tech.newstatesman.com/guest-opinion/ethical-artificial-intelligence

Chu, S. K., 2020. *How to Stop Data Theft.* [Online]   
Available at: https://www.currentware.com/how-to-stop-data-theft/

Freedman, M., 2020. *How Businesses Are Collecting Data (And What They're Doing With It).* [Online]   
Available at: https://www.businessnewsdaily.com/10625-businesses-collecting-data.html  
[Accessed April 2021].

Ramnath Balasubramanian, A. L. a. D. M., 2018. *Insurance 2030—The impact ofAI,* s.l.: McKinsey & Company.

Richard Vidgen, G. H. I. R., 2019. *Exploring the ethical implications of business analytics with a business ethics canvas,* Sydney: Elsevier.

Sahil Verma, M. E. R. J., 2021. *Removing biased data to improve fairness and accuracy,* Seattle: Cornell University.

Stefan Larsson, F. H., 2020. *Transparency in artificial intelligence,* Sweden: s.n.

Australia Gov (Australian Government) (2020). [*Understanding AI ethics in context*](https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/understanding-ai-ethics-in-context)*. Australian Government,* accessed 5 May. 2021.

Accenture (2018). [*ACCENTURE FEDERAL SERVICES RESPONSIBLE AI: A Framework for Building Trust in Your AI Solutions*](https://www.accenture.com/_acnmedia/pdf-92/accenture-afs-responsible-ai.pdf.)*. Accenture,* accessed 5 May. 2021.

Taulli, T. (2019). [*How To Reskill Your Workforce For AI (Artificial Intelligence*](https://www.forbes.com/sites/tomtaulli/2019/05/04/how-to-reskill-your-workforce-for-ai-artificial-intelligence/?sh=63ab2f943eb9)*), Forbes,* accessed 5 May. 2021

Australia Gov (Australian Government) (n.d.). [*Undertaking a Privacy Impact Assessment — a free eLearning course from the Office of the Australian Information Commissioner*](https://www.oaic.gov.au/s/elearning/pia/topic6.html)*,Office of the Australian Information Commissioner,* accessed 5 May. 2021.