

[Neoteric Health Care Company]

BUSINESS REPORT

Prepared by: [Rokshana Pervin]

Contents

[Neoteric Health Care Company]	1
Prepared by: [Rokshana Pervin].....	1
Executive summary	3
Part A: Current storage and security strategies and risks.....	3
Part B: Data quality ethical analysis.....	4
Part C: Towards a solution	5
Conclusion	8
References	9

Executive summary

“Nowadays, many organisations such as health care organisations are confronting many problems to organise their internal and external data and its infrastructure. Here, Neoteric Healthcare is one of the Australia’s leading provider medical diagnostic services. Neoteric is planning many infrastructure regarding their internal and external data and struggling with many issues such as current data storage strategies and mechanisms for maintaining the security of big data, national and global regulatory limitations as well as data quality issues. In this business report, both current storage and data quality issues will be explained with translucent solutions as well as applied ethical lens by analysing the actual data breaches of regulation in the identified solutions”(as shown in appendix A).

Part A: Current storage and security strategies and risks

There are some risks associated with the Neoteric’s Healthcare Company in relation to current storage strategies and mechanisms for maintaining the security of big data. These are:

Ranking of risk	Type of Risks names	Details of risks
1	Data Security	one of the European newspapers claimed the security software of this hospital detected several unauthorised access requests to the hospital data base. However, there is no evidence whether the data breach was successful.
2	Data Inaccuracy	Inaccurate diagnostic results due to actual or alleged mistakes or errors could result in financial loss and/or reputational damage.
3	Financial risk	Neoteric’s reported revenue and earnings will fluctuate with changes in the currency exchange rates between the Australian dollar (Neoteric’s reporting currency) and the currencies of Neoteric’s offshore operations.
4	Managing relationship	If for any reason, Neoteric failed to maintain strong relationships with these parties (including physicians, hospital groups and third party) or damaged its reputation with them, there would be a risk that it could lose businesses to competitors.
5	IT	Technological changes in diagnostic testing tend to happen more slowly than in industries.
6	Regulatory risk	Changes in regulation can have the impact of increasing costs or reducing revenue (through volume reductions). Loss of a licence or accreditation required to operate one or more of Neoteric’s businesses could impact revenue both directly and through damage to Neoteric’s reputation.

According to Australian Government Office of the Australian Information Commissioner (2022) Privacy regulatory action policy, National regulatory constraints around data privacy are:

- Data Breaches
- Losing customer trust.
- Ambiguity about information
- Inconsistency

Apart from this, global industry-standard regulatory constraints around data privacy are:

- Data dominion
- Data localization
- Privacy Acts (PROTEGRITY 2023).

Part B: Data quality ethical analysis

There are some ethical risks associated in handling with Neoteric's data qualities including source, time and how it's created that are represented with the table structure:

Level of Risks	Type of Risk name	Risks details
1	Ethical risk about timeliness	The timeliness of information is playing a vital role in data quality characteristics. Poor data quality extends functional cost because time and other resources are spent detecting and correcting errors. The incomplete, inaccurate, and delayed data and action sometimes make patients' life and death consequences which are completely unethical on timely manner guidelines and can cause also reputational damage for any organisation.
2	Ethical risk on source	<p>As Neoteric is using patients' medical records and their own data, in this case, Neoteric might struggle with patient data ownership. As this information are personally identifiable (PII), it will be unlawful and unethical while collecting this information without consents and might bring the large amount of penalties for this organisation. For instance, Google and YouTube has to pay \$136 million to the FTC (Federal Trade Commission) and New York Attorney General allege where YouTube violated the COPPA (Children's Online Privacy Protection Act) Rule by collecting children's personal information without their parents' consents.</p> <p>Apart from this, it is found that Neoteric handles vast amount of data which known as big data as well as trying to embark their digital transformation to migrate and integrate with the newly infrastructure such as cloud. In this case, Neoteric might confront with various technologies and tools' challenges in adoption in the cloud such as Hadoop and spark system where needs to take decision appropriately in which technology or algorithm will be suitable for the organisation. As machine learning learn from the input data, it is considered to focus on the data source or where the data come from? If the algorithms draw an unethical factor based on race, gender, socioeconomic status, then Neoteric has to conclude the algorithm's activities immediately which will be harmful to its reputation and assets.</p>

3	Ethical risk on creation data	<p>As Neoteric uses the artificial intelligence algorithms to facilitate their fitness app and will be included lots of unknown data, there is some probability to confront with the ethical dilemmas such as:</p> <ul style="list-style-type: none"> • Users may program biased AI systems without noticing the data. • Watch-owner can enter additional biased data which can be erred for AI algorithm. <p>Here, the ethics of practice enlightened the concerning question regarding responsibilities, liabilities of people and organisations in charge or data processes.</p>
---	-------------------------------	---

Part C: Towards a solution

Data Architecture Solutions for Part A aligned with the Neoteris's data characteristics:

Strategy for financial risk:

As Neoteric is planning to make broader business globally with various countries, it's revenue would be fluctuate. Hence, to make Neoteric's revenue stable at least in nationally, it is considered to use machine learning approach by following the data science and BI (Business Intelligence) technologies(Ghosh 2021). In this case, BI or business intelligence is technology-driven process which helps data to be interpreted to the users by applying BI dashboard (Knight 2021). On the other hand, data science will help to use the machine learning approach where ROI (Return on Investment) term will support to maximise the revenue (Linna 2020). Here, if experts or data scientists use the following features in the right algorithms, they will boost the revenue in every year.

$$returns = value - (1 - accuracy) * cost\ of\ a\ mistake$$

Here, the variables are,

- **returns:** Generated net value or profit per prediction.
- **value:** The new value generated by every prediction (e.g. assigning a document to the right category now takes 0.01 seconds instead of 5 minutes, so the value is 5 minutes saved)
- **accuracy:** The accuracy of predictions made by the algorithm
- **cost of a mistake:** The additional costs incurred by a wrong prediction (e.g., it takes 20 minutes for someone to correct the mistake in the system)

Figure: Linna (2020), Return on Investment for Machine Learning, accessed 11 May 2023. <https://towardsdatascience.com/return-on-investment-for-machine-learning-1a0c431509e>

Strategy for regulatory risk:

As changing regulation will impact on Neoteric's revenue (governentrisk360 n.d.), they should use some response to mitigate the organisation's disruption. In this case, according to Australian Government Office of the Australian Information Commissioner (2022), Neoteric's should apply the data breach response plan or framework as it's changing rules or regulations which can impact on it's customer and experts.

Strategy for IT (Information technology) risk:

According to The Framework for Information Quality (2008), Neoteric healthcare needs to use POSMAD framework to find the right technology to find out the appropriate information.

In this case, as Neoteric is planning to transfer its technology to the cloud they can use data lakes that offers to store the structured and unstructured data. In this case, `Faster Amazon Redshift ETL` service will help Neoteric to provide rapid diagnostic test to the patients (aws 2023).

Strategy to manage relationship:

To manage relationship or partnerships with stakeholders, it is considered to follow the TOGAF (The Open Group Framework) approach as a framework where enterprise architecture will help Neoteric healthcare to form the strong bond with stakeholders (The TOGAF Standard 2022)

Apart from this, according to data governance model framework, Neoteric should use the non-invasive approach to manage the relationships with individuals (The Data Administrator Newsletter 1997).

Strategy to increase the data security:

As Neoteric is handling PII (Personally Identifiable information) that use with other relevant data individually, Neoteric can use 'data masking' technique which also known as `data obfuscation` that hides the actual data using modified content to protect their sensitive data.

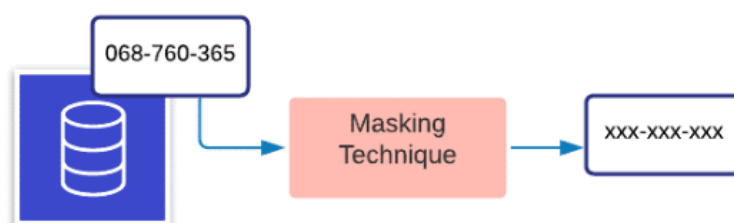


Figure: Wickramasinghe (2021), What's Data Masking? Types, Techniques & Best Practices, <https://www.bmc.com/blogs/data-masking/>

Here, the most secure data masking technique can be used as `data encryption` that masks the data and requires key to decrypt the data (Wickramasinghe 2021).

Strategy to increase data accuracy:

To get the accuracy and completeness of the diagnostic results as well as to prevent duplicate patients' records, Neoteric can use `Patient Demographic Data Quality (PDDQ) Framework` (The Office of the National Coordinator for Health Information Technology n.d.). Because, using this framework, Neoteric will be facilitated in using accurate data.

“Apart from this, as Neoteric is transferring its technology to the cloud, it can use open-source tool by following the machine learning approach to improve the data quality such as ‘Holoclean’ tool which will help to detect the automatic error (O'Reilly 2019).

Strategy to mitigate the Neo's fitness app risk:

As healthcare has hundreds of thousands of languages, Neoteric healthcare company can use machine learning approach by following Amazon Comprehend Medical process to add contextual data points to individual's profiles. In the following process, unstructured medical text such as doctors note, patients' health records have been provided to the Amazon comprehend medical service that automatically extracted as medications and medical conditions. Then this service links with the extracted entities or features to medical ontologies” (as shown in appendix B).



Figure: Aws (2023) *Amazon Comprehend Medical*, <https://aws.amazon.com/comprehend/medical/>

Here, machine learning approach using natural language processing (NLP) algorithm that supports medical knowledge categories.

Supported medical knowledge categories

The Healthcare Natural Language API supports the following medical knowledge categories:

Medical knowledge category	Description
ANATOMICAL_STRUCTURE	Complex part of the human body

Figure: GoogleCloud (2007) Healthcare Natural Language API, <https://cloud.google.com/healthcare-api/docs/concepts/nlp>

Solutions for identified risks in Part B:

To mitigate the ethical risk about timeliness, Neoteric needs to monitor regularly on upcoming data and infrastructure to collect the precise information from the patients' records as well as to follow the real-time processing method by following 'data pipeline management' (tdwi 2018) to consolidate intensive care insights with broader data that establishes trends and helps to anticipate future outcomes (Bresnick 2016).

Apart from this, according to The Data Administrator Newsletter (1997), Neoteric needs to follow data governance rules and policies to minimise the ethical risk on data source or accountability and responsibility.

Ethically, it is considered that the actual breaches of regulation and trust in the above-identified solutions can cause potential harm such as big amounts of penalties as well as reputational damage for the organization.

Conclusion

"Healthcare organisations are sensitive in regards to collect, store and processing their internal and external data. If any risk associate with the certain data, it will bring potential harm and damage orgainsation's reputaions within milliseconds. In this article, Neoteric's potential risks and strategies have been discussed elaborately to retain the reputation and prevent from any damage that may occur from collecting and processing information. Moreover, ethical consideration has been applied in analysing actual breaches of regulation in the identified solutions so that stakeholders and providers can take decisions appropriate for their future events" (as shown in appendix C).

References:

- Australian Government Office of the Australian Information Commissioner (2022) *Privacy regulatory action policy*, oaic website, accessed 19 May 2023. <https://www.oaic.gov.au/about-the-OAIC/our-regulatory-approach/privacy-regulatory-action-policy>
- Australian Government Office of the Australian Information Commissioner (2022) *Preparing a data breach response plan*, oaic website, accessed 19 May 2023. <https://www.oaic.gov.au/about-the-OAIC/our-regulatory-approach/privacy-regulatory-action-policy>
- Aws (2023) *Analytics on AWS*, aws website, accessed 11 May 2023. <https://aws.amazon.com/big-data/datalakes-and-analytics/>
- Aws (2023) *Amazon Comprehend Medical*, aws website, accessed 11 May 2023. <https://aws.amazon.com/comprehend/medical/>
- Bresnick J (2016) *How Big Data Velocity Informs Population Health, Patient Safety*, healthitanalytics websites, accessed 06 May 2023. <https://healthitanalytics.com/news/how-big-datas-velocity-informs-population-health-patient-safety>
- Ghosh P (2021) Data Topics, *Data Science vs. Business Intelligence*, dataversity website, accessed 08 May 2023. <https://www.dataversity.net/data-science-vs-business-intelligence/>
- GovernmentRisk360 (n.d.) *Regulatory Risk*, governmentrisk360 website, accessed 12 May 2023, <https://governmentrisk360.com.au/about-gr360/regulatory-risk/>
- GoogleCloud (2007) *Healthcare Natural Language API*, google website, accessed 11 May 2023. <https://cloud.google.com/healthcare-api/docs/concepts/nlp>
- Knight M (2021) Data Topics, *Data Science vs. Business Intelligence*, dataversity website, accessed 08 May 2023. <https://www.dataversity.net/what-is-business-intelligence/>
- Linna E (2020) *Return on Investment for Machine Learning*, towardsdatascience website, accessed 11 May 2023. <https://towardsdatascience.com/return-on-investment-for-machine-learning-1a0c431509e>
- O'Reilly 2019, *Recent trends in data and machine learning technologies - Ben Lorica* (O'Reilly Media), YouTube website, accessed 14 May 2023. <https://www.youtube.com/watch?v=xnf4Dr8regw>
- PROTEGRITY (2023) *EMPOWER GLOBAL BUSINESS WITH BORDERLESS DATA*, borderlessdata website, accessed 21 May 2023. https://www.borderlessdata.org/?utm_term=global%20data%20privacy%20compliance&utm_campaign=US/UK/AUS++Cross-

[Border&utm_source=adwords&utm_medium=ppc&hsa_acc=4329975922&hsa_cam=19914755275&hsa_grp=145224739821&hsa_ad=656248489883&hsa_src=q&hsa_tgt=kwd-2006295333786&hsa_kw=global%20data%20privacy%20compliance&hsa_mt=b&hsa_net=adwords](https://www.google.com/adsense/adsense?border&utm_source=adwords&utm_medium=ppc&hsa_acc=4329975922&hsa_cam=19914755275&hsa_grp=145224739821&hsa_ad=656248489883&hsa_src=q&hsa_tgt=kwd-2006295333786&hsa_kw=global%20data%20privacy%20compliance&hsa_mt=b&hsa_net=adwords)

The Framework for Information Quality (2008) accessed 20 May 2023.

https://static1.1.sgspcdn.com/static/f/739370/15242245/1321966972997/10steps_AllQuickRef.pdf?token=C9IBTgHWQOvqyNeZYEneh5jzFVI%3D

The TOGAF Standard (2022) *Welcome to the Digital Edition of the TOGAF Standard*, opengroup website, accessed 13 May 2023. <https://pubs.opengroup.org/togaf-standard/index.html>

The Data Administrator Newsletter (1997) *Comparing Approaches to Data Governance*, tdan website, accessed 11 May 2023. <https://tdan.com/defining-data-governance-core-principles/17087>

The Office of the National Coordinator for Health Information Technology (n.d.) *Data Quality Assessment*, healthit website, accessed 13 May 2023. <https://www.healthit.gov/playbook/pddq-framework/data-quality/data-quality->

[Tdw \(2018\) *Data Cataloging Comes of Age*, tdwi website, accessed 09 May 2023, https://tdwi.org/Articles/2018/09/17/DIQ-ALL-Data-Cataloging-Comes-of-Age.aspx?Page=2](https://www.healthit.gov/playbook/pddq-framework/data-quality/data-quality-)

The Data Administrator Newsletter (1997) *Defining Data Governance Core Principles*, tdan website, accessed 11 May 2023. <https://tdan.com/defining-data-governance-core-principles/17087>

Wickramasinghe s (2021), *What's Data Masking? Types, Techniques & Best Practices*, bmc website, accessed 12 May 2023. <https://www.bmc.com/blogs/data-masking/>