Presentation SCRIPT

# Introduction (purpose, aim, user stories implemented, etc) - BEN

Good morning, we are group two from ISYS3888. The name of this project is ‘Benchmarking commercial AI products for medical imaging service providers’. As we know, AI technology has been developed rapidly over the last decade, the effectiveness and efficacy it brings to the medical imaging practice has already been proven by various scientific researches and many of the AI products are already being used commercially. For example, qXR is one of the commercial AI which has the functionality of diagnosing tuberculosis using patient’s chest CT scans. Requested by our clients, there are two main purposes of this project, one is to compare the medical imaging AI product qXR with other commercial AI products, the other is to consult on whether this product should be purchased by our clients Simon, Neysa and the ABC company. The aim for this project is to evaluate two aspects of qXR using methods accepted by the medical imaging industry. The aspects are technical performance and user experience. The technical evaluation focuses on the technical indicators like accuracy, specificity, and sensitivity. User experience evaluation targets the use cases of qXR and user experience of AI products used in medical imaging in general. The methods used are systematic review and meta-analysis because they are the two methods accepted and recognized by the industry. The stakeholders are our clients, the ABC company, radiologists and hospital managers who are interested in AI evaluation. We created comprehensive and detailed user stories in the report. I only will present three example user stories considering the time limitation. The user stories are, as a radiologist, what is the best AI medical imaging product I can find which helps the diagnosis of tuberculosis? As a radiologist, should I accept qXR into my current workflow? As an AI developer, what kind of improvement can I make to raise the user acceptance of AI as an innovation?

# Requested Changes and Reflection on the Client’s Feedback - Cailin

* “*Before we continue with our presentation, we would like to reassure you that we have taken the feedback that you have given to us from the first deployment to improve our report.*
  + *We made an outline and followed systematic review protocols,*
  + *And we have rewritten answers making sure we include analysis, and have supporting evidence for every decision that we have made*

# Testing (Bronte)

### Testing Plan

* We would also like to quickly comment on the testing we used to assess the quality of our project. Our primary method of testing was quality assurance. Each of our seven quality areas had a defined criteria to assess its quality by manually reviewing the work.
* The tests began with the smallest component of our project, the articles we collected, and progressively moved through the project to the final report as a whole, therefore, as each component of the project was added, the project was tested to determine its quality.
* In addition we completed this testing multiple times throughout the project. Overall we found that our project had an acceptable quality, however, could have been improved further with more time and more extensive knowledge of the systematic review process and analysis methods.

# Report - kiran

We will now briefly answer each of the eclair guidelines to give you an overview of the key insights each question has given us in crafting a recommendation for you.

## 2 Relevance

### 2.1 What problem is the application intended to solve, and who is the application designed for? - Kiran

The purpose of qXR is to automate the x-ray interpretation process in order to bridge the gap between the incidence of tuberculosis and the actual number of reported cases. With this application patients can be diagnosed at a much faster rate and can thus be put into treatment sooner with the process taking minutes rather than days, this can have a significant impact in mitigating the negative effects of TB and prevent further transmission within the community. qXR scans for 30 different abnormalities in chest x-rays which are then processed into a report. The end users of qXR are those in the healthcare field including radiologists and medical imaging facilities who will be able to efficiently diagnose and subsequently treat TB, further to this, as patients will be majorly impacted by choice of AI we consider them as users as well.

### 2.2 What are the potential benefits, and for whom? - Edward

There are 7 benefits in general for qXR. The first one is time-efficiency because qXR can read chest x-rays in less than one minute using deep learning algorithms. qXR also offers an automated interpretation of chest x-rays. In addition, qXR can reduce the chance of late diagnosis, under diagnosis and even misdiagnosis to improve the quality of patient care because of the time-efficiency. Since automated AI system performed much better than radiologists in accurately positioning the patients, using qXR can potentially improve image quality and reduce patient dose. qXR has also been trained to detect 29 other abnormalities which will benefit patients a lot. The workflow optimisation is also a benefit for qXR because qXR can cluster the x-rays into normal, abnormal and to be reviewed. Then prioritise cases that need immediate attention. Lastly, qXR product has an automated management platform called qTrack which can enhance the radiological practice workflow.

### 2.3 What are the risks associated with the use of the AI system?-zheyuan

Although the ai systems are very beneficial in radiology in use, but there still are many risks when using it. The AI system might have ethical and legal liability problems, since AI is a new technology, there might have weak regulations

for using AI during diagnosis, this brings an automation bias concern that the radiologists may over rely on the AI and lose their own decision.

Second, not all the AI systems for radiology are well protected, the system may be attacked by hackers and the data may be leaked to malicious people.

Finally, the data stored in AI system may be compromised, as these data is store in server or workstations,

These data might be accessed by malicious people and the data might be tampered.

## 3 Performance and Validation

We will now cover performance and validation questions

### 3.1 Are the algorithm’s design specifications clear?- Andersen

### 3.2 How was the algorithm trained?- Andersen

### 3.3 How has performance been evaluated?

#### 3.3.1 Abstract

#### 3.3.2 Method - Zichen

For the method we used in the evaluation, first we will use PRISMA-P protocol to perform the systematic review including defining the search queries and defining inclusion and exclusion criteria. By following the search queries, we then search the articles in various databases and journals to find all the relevant comparison studies with qXR. Then we will extract the necessary data from each article for the preparation of meta-analysis. Finally, we will use tableau to create some forest plots and SROC plots in terms of sensitivity, specificity and accuracy for all the data.

#### 3.3.3 Systematic Review - Zheyuan

A systematic review was conducted across four databases and three journals,

the queries used identified articles of qXR and other AI products to allow for benchmarking.

Through this method, 26 articles were identified, 8 of them were duplicated, 12 articles were excluded because of lack of data or not public.

Finally 6 articles were included in the review. Furthermore,

four of our original certified products were removed due to lack of articles supported and lack of data.

#### 3.3.4 Heterogeneity - Kiran

The next stage of our meta-analysis involved testing for heterogeneity of the data, in order to do this we created coupled forest plots for each product listed in each retained article. This allowed us to compare the results of these products and make a judgement on whether the results across studies were homogenous or heterogenous following the guidelines in Diagnostic test accuracy methods for systematic review and meta-analysis. This article suggested using an SROC curve to inform judgements on heterogeneity but it was found this could not be achieved with the data that was extracted from the selected articles, thus, we compiled the forest plots and visually concluded that the gathered results were homogenous and therefore accepted the results of the studies

#### 3.3.5 Results - Andersen

#### 3.3.6 Discussion - Andersen

### 3.4 Have the developers identified and accounted for potential sources of bias in their algorithm? - Kiran

The primary source of alleviating bias that was found to be outright addressed was in their selection of datasets to validate the product they avoided selection bias by using a combination of CT scans and an algorithmically selected dataset. Another major bias that would have been identified by the qXR developers is the underdiagnosis bias which can be especially prevalent with patients who come from underserved populations, this is an issue as it can lead to patients who are ill not receiving the necessary care, qXR did not state how this bias was accounted for in the development of the algorithm.

### 3.5 Is the algorithm fixed or adapting as new data comes in?-zheyuan

Through research, qXR was able the be repurposed to aid in the diagnosis of Covid-19, and furthermore, qXR is keep under trained with further

analysis of CXR, from 2020 to 2022 qXR has used and additional one million CXRs. Since through these evidence it can be said that qXR is capable of

adapting new data when comes in.

## 4 Usability and Integration

### 4.1 How can the product be integrated into your clinical workflow? - Cailin

* As qXR is a medical tool is used for patient diagnosis and treatment, ensuring that the radiology practice is compliant with clinical guidelines and standards is crucial before any integration begins. This includes making sure that their software is always up to date.
* qXR can be integrated into the clinical workflow of a radiology practice in the following ways:

1. API based -> simple and intuitive method of accessing qXR
2. PACS Based: -> more complex method but allows for finer-grained control

* A simple alternative is provided by Integration via AI marketplace or distribution platform. Parkh (2019) states that the AI marketplace integration requires ‘little or no support’ and ‘offering end-to-end solutions’ that solve the engineering issues (deployment, workflow integration, etc.) and commercial issues (contracting, billing, etc.) that arise when implementing AI
* Once implemented and decisions on managing the technology are completed, training the radiological practices' staff is the final step.and should be noted that qXRs implementation and training is available from qure AI
  + A large amount of time should be dedicated to improving the user’s knowledge as this improves the user’s attitudes towards qXR (this finding is further expanded on in section 3[.2.7](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.zapcmo7hu3b)) .

### 4.2 How exactly does the product impact the workflow?

The question, how exactly does the product impact the workflow? Was one of our focus questions for the evaluation. Our user experience team conducted a systematic review and thematic analysis to answer it.

#### 4.2.2 Introduction - Bronte

qXR uses artificial intelligence to interpret chest x-rays and is an example of the emerging use of AI in healthcare. There are many important considerations when evaluating whether such a product is suitable, for instance, its safety, legality, and accuracy. However, another important consideration is its impact on the existing workflow of radiologists.

It is answered by examining the impact on the workflow from the perspective of radiologists, using user experience and acceptance studies. The studies were chosen using a systematic review following the PRISMA process and synthesised using thematic analysis. The systematic review and thematic analysis was initially designed to be specific to qXR, however, no suitable studies were found, which meant the scope was broadened to AI in radiology. It is a research gap for qXR, however, the impact identified more broadly is assumed to likely be similar to qXR’s impact.

#### 4.2.3 Methods - Bronte

* Two systematic protocols were developed using PRISMA-P, the first specific to qXR and the second broadened to AI in radiology.
* The eligibility criteria was divided into seven items that the article must include. In brief these were journal articles from 2020 for qXR and 2018 for AI in radiology, the article includes primary data from a usability-related study undertaken in a standard clinical setting from the perspective of people who work in radiology, and the sample size is at least 120 for surveys and 20 for interviews and focus groups.
* The databases PubMed, ProQuest, Embase, and ScienceDirect were searched, in addition to the journal The Lancet and five articles were provided by the client. A range of search queries were defined with keywords including user experience, user acceptance, user satisfaction, and usability, alongside either qXR or AI products.

To synthesise the data in the included studies, thematic analysis was chosen. It was primarily deductive, with the Unified Theory of Acceptance and Use of Technology used as the conceptual framework. The constructs from the model, performance expectancy, effort expectancy, social influence, and facilitating conditions, were used as themes, and through research and brainstorming, sub-themes were developed and defined as codes in a codebook. Each team member initially coded the data individually and then met to discuss the data and finalise the codes and themes. Then the results and discussion was written for each theme and finally the team met again to discuss the impact of the themes and their interaction to write the conclusion.

#### 4.2.4 Systematic review of qXR - Bronte

The queries were executed in all the information sources which resulted in a total of 22 articles being identified and 18 remained after duplicates were removed.

The titles and abstracts of the articles were screened and 17 articles were excluded.. The majority of articles were excluded because they were studies of diagnostic test accuracy or reviews. The remaining article was excluded after a full-text read through of the article because although it contained information about the use and implementation of qXR, however, did not contain a study on user experience.

#### 4.2.5 Systematic review of AI in radiology Benjamin

We began our Systematic review of AI in radiology from writing the protocol. The protocol explicitly regulates the criteria and database we used for this review. The detailed protocol is provided in the report and it ensures the quality of this review. The criteras for choosing articles and databases were already mentioned by Bronte in the method section. Then we followed the PRISMA process strictly to conduct the review. Initially a total of one hundred and fifty five articles were collected from five databases including pubmed, proquest, embase, science direct, and the medical journal ‘The Lancet’. In addition, our client Simon also provided 5 articles for us to review. 25 duplicated articles were removed right after collection. After that we excluded 102 articles based on the titles and abstracts. The full text of the remaining articles were then read by us so that we can exclude according to the criterias in the protocol. At last, 9 articles remained. The major reason for the exclusions was the lack of relevance to user experience evaluation for medical imaging AI products. We then analysed the characteristics of the studies based on sample size, demographics, and type of the studies. We further extracted the data into an excel spreadsheet and did a thematic review in order to integrate the results and form our conclusions.

#### 4.2.7 Results and Discussion -

##### 4.2.7.1.1 Results - PE - Bronte

Performance Expectancy is defined as “the degree to which an individual believes that using the system will help him or her attain gains in job performance”. As a construct of the UTAUT model the codes were pre-defined and no changes were required.

Code PE1 refers to the efficiency of the workflow being changed by integrating an AI product into the workflow. There were no direct questions about efficiency and only one data item labelled PE1. This was in relation to a study that looked at the impact of the AI providing explanations in addition to the highlighting of x-rays and a respondent said that explanations “save time for a clinician”.

PE2 is defined as the effectiveness of the work produced by radiologists when an AI product is introduced into the clinical workflow. The majority thought that effectiveness in a range of areas would be improved, including patient care and diagnostic practice. For example, 82.8% agreed that AI can help reduce radiation dose levels while maintaining optimal image quality. However, 42% of people were concerned about its current technical performance.

PE3 is about the extrinsic motivators and demotivators for radiologists when integrating an AI product into the workflow. This code saw significantly varying results, for example, in one study, 61.3% were concerned that AI would replace most jobs and negatively affect the profession rather than just being an assistive tool, whereas in another, 62% disagreed that radiologists’ jobs were in danger due to AI.

4.2.7.1.2 Discussion - PE - Bronte

It was anticipated that efficiency would be an important factor as it is a commonly described benefit of integrating an AI into the clinical workflow. Additionally, 77% of respondents agreed that AI can be used to help diagnostic radiology by workflow optimisation. Research into efficiency may be present in studies that were not identified through this systematic review, and a potential reason is that it may have been approached through an objective comparison, rather than user experience. It would be beneficial to conduct further research into this factor to determine whether improved efficiency has been actually reported or observed when using AI, or simply remains an assumed benefit.

Increasing the effectiveness of work output can be observed as improved patient care, less incorrect diagnoses, earlier detection, and reduced radiation doses. The results show that radiologists anticipate that increased effectiveness will be a benefit of integrating AI into their clinical workflow, however, there are current concerns about the performance of AI.That this concern will likely be addressed as AI algorithms continue to be improved through training and evidence of excellent diagnostic test accuracy emerges.

Extrinsic motivators and demotivators can have significant influence over whether radiologists support the integration of AI into their workflow. The included studies focused on the role of radiologists and their job security, with the responses varying greatly across different studies. This difference was explained by the level of AI knowledge the radiologists had. Those with less AI knowledge were much more concerned about their job security than those with intermediate or advanced knowledge.

In considering performance expectancy’s relationship with other factors, across multiple studies, it was found that performance expectancy is affected by social influence and effort expectancy. This suggests that others' opinions of AI, combined with how much effort they expect it requires impacts on how they expect it to perform. While performance expectancy affects intention to use the system and to a mixed extent, trust of the AI. This supports the earlier discussion that if people who work in radiology view the integration of AI into the workflow as something that positively impacts efficiency and effectiveness while not replacing their jobs, they would be more likely to accept its use.

##### 4.2.7.2.1 Results - EE - Cailin

##### Effort expectancy is defined as the degree of ease associated with the use of the system (Venkatesh et al., 2003)

##### It is interesting to note that in several studies, it reveals that effort expectancy has no significance in influencing behavioural intention This phenomenon is demonstrated Ann Oper Res (2020) study where they found effort expectancy has an insignificant score of (β = 0.078, p > 0.05 when measuring its influence on behaviour intention. This idea's conclusion is further illustrated from the A.V. Prakash and S. Das 2021 study where they had a (β = 0.004 , p > 0.05). where they also measured the impact effort expectancy has on behavioural intention.

4.2.7.2.2 Discussion - EE - Cailin

##### A.V. Prakash and S. Das explain the reason for the insignificance of the effort expectancy metric is that ‘radiologists are used to highly complex machines in their routine clinical practice, and the ease of use associated with the system may not factor in as an influential criterion in the adoption-related decision-making. Another reason may perhaps be that radiologists did not foresee any difficulty in using ICDDSS, and they perceived the use of ICDDSS to be as easy as using their current systems/machines.’

##### However, it would be beneficial for our clients to collaborate with Qure.Ai to help improve the qXR. Our team suggests our client to complete effort expectancy-focused surveys, with questions that include “list all the reasons that made you believe the system seemed: easy to use/difficult to use”, “reasons why you felt using the system was easy to use/complex to use” that can be shared with qure.Ai where they take effort expectancy findings into account when designing their technology to make sure they improve their customer’s user experience.

##### 4.2.7.3.1 Results - SI - Ben

Social influence is defined as the degree to which an individual perceives that important others believe he or she should use the new system, and it is constructed with three factors: subject norm, social factors, and image. Subject norm is the reflection of the potential influences from the important people of the participant’s social network. We found that about 70% of the patients are excited about the use of AI in medical imaging, in this case, the opinion of the patients are considered important because they are also clients to doctors and hospitals.

The social factors of the participants is a construct that directly mirrors the individual's agreement with the general trend of the perceived trend of the society held toward the use of AI products in medical imaging practices. We found that 78% of the participants are aware that, currently there is a trend for the use of AI in medical imaging practice, only less than 2% neglects this trend.

Image reflects the degree to which use of an innovation is perceived to enhance one's image or status in one's social System. However, none of the studies raised any question relating to this factor and it leaves us with no data to analyse.

##### 4.2.7.3 .2 Discussion - SI - Ben

Some insights we drew from the results are that the majority of the radiologists thinks their patients are excited about AI used in medical imaging and this will affect their own opinion of AI. The trend of using AI products in the medical imaging practices are being recongized by doctors internationally. However, the number of hospitals and managers with an open attitude does not reflect this trend. This could be that they want to seek more evidence and assure AI is safe to use before the actual implementation. Also, less than half of the hospitals has a so called ‘local champion’ which is a person who promotes AI technology to his or her social network proactively. These phenomenas indicated that social influence is being underestimated when it comes to the promotion of AI products. More importantly, social influence is important for raising user acceptance, because it impacts performance expectancy and effort expectancy directly. It also influences the level of willingness, level of trust, and risks identification. Therefore we suggest that AI developers take social influence into account when promoting a product.

##### 4.2.7.4.1 Results - FC - Cailin

* 537 (52%) had no knowledge of informatics/statistics and 75% had no previous experience in programming (Merel Huisman, 2021)
* Therefore it should comes as no surprises that a majority of users believed they require further education and/or training to be able to embrace these emerging AI trends in medical imaging (B.O. Botwe et al, 2021. This finding was further illustrated when Botwe asked radiologists what the believed are the greatest barriers for users when using new technology in their work place which they responded with -Lack of dedicated courses and learning materials: 219 (21.5%); Lack of mentorship, guidance and support from “experts" 136 (13.3%); Lack of evidence-based material and proof of improved clinical outcomes
* Additional data findings for Facilitating conditions that Lea found was that in close to half of respondents agreed that their workplace had unstructured planning and monitoring when implementing AI. Lea also noted 33.3% of users felt that there was an absence of guidelines/best practices. Lastly, lea saw that 33% of users mentioned that there was no empirical evidence.

##### 4.2.7.4.2 Discussion - FC - Cailin

* As highlighted, a majority of users believed that the greatest barrier when working with A.I. is the lack of dedicated courses and learning materials. However there are no studies that specify what specific courses and learning materials would be the most beneficial to provide the users to help ensure a smooth integration process.
  + Therefore our team suggests further research to be done where various radiological work users are provided different resources before the technology is integrated.
    - And then they can record their findings such as: which users learnt to use qXR the quickest, which users appreciated the technology after going through all their resources etc.
    - These studies can then help identify which specific materials are crucial to provide to guarantee a positive workflow integration.
  + The supporting resources should be available throughout the entire life span of the A.I. technology in the radiological process as a large proportion of users have no experience using technology which can result in users requiring assistance for an extended amount of time.
* In Calisto’s studies, it was pointed out that there is a large proportion of respondents agreed that their workplace had unstructured planning and monitoring when implementing.
  + The importance for a structured process is that it can create a sense of order to the workplace which is essential as there is a high chance of users are worried with this unfamiliar technology that will change their original workflow.
  + Monitoring technology usage is a crucial element too as the metrics from it can be used as feedback to see how well the technology is actually being used instead of just relying on user’s verbal responses which has risk of not being entirely accurate.

##### 4.2.7.5.1 Results - PA - Ben

As the research advanced to the data extraction process, it is to the team’s recognition that the four initial factors adapted from the UTAUT model are not sufficient. The team discovered that a significant percentage of the data extracted points toward the subjective opinion of the participants. Thus, a new factor ‘personal attitude’ is added after group discussion. This factor directly reflects the subjective attitude of the participants toward AI technology used in medical imaging from four aspects: acceptance, willingness, trust, and risk.

The factor ‘Acceptance’, intentionally examines the level of acceptance of the participants. For this factor, we found that 82% of the participants are excited about the use of AI in medical imaging, only 4% are not excited. But 67.5% of the are concerned about the use of AI, only 11% are not concerned. 76% think AI will bring more benefit than harm, 4% disagree with this statement. A significant 85% look for tools and support when using AI at work.

The factor ‘willingness’ is the degree of the intention the participant holds on using AI in medical imaging. We found that the willingness to use AI rose significantly from 4% to 58% after aiding the participant with a special kind of support. The level of confidence rose from 30% to 67% after receiving the support. The general willingness to use AI in medical imaging practice is high, reaching 85%.

The factor trust mainly reflects the level of trust on the accuracy of AI in diagnoses. Some minor aspects related to implementation and job security. Some of the findings of trust are intriguing, one study reveals that the doctors prefer AI results with strong explanations rather than weak explainations. Even no explanation is better than weak explanation. 58% of the participants showed inconsistency in trusting AI products and only 25% showed strong trust in these products.

Risk is the factor reflecting the risks subjectively identified by radiologists when it comes to the use of AI in medical imaging. Almost all of the studies related to this topic are able to point out risks that the participant thinks he or she is facing when using AI at work. We found that 42% of the participants showed doubt in quality and safety in AI products. 58% have concern regarding funding, and one-thrid of them believes there is a legal and regulatory issue.

##### 4.2.7.5.2 Discussion - PA - Ben

Personal attitude is the most important construct in our current user experience model

because is a direct reflection of the intention of use. Interestingly, controversy is detected while integrating our results. On one hand, the acceptance and willingness of using AI is high. On the other hand, the level of trust is low and it is to the radiologists' belief there are unresolved risks.

Obviously, the radiologists are excited about AI and they are willing to embrace this innovation into their current workflow. However, the majority of doctors think that accuracy in diagnosing diseases is the same with or without AI. They believe they are as accurate as AI when it comes to diagnoses.

We also found that tools and support are essential for users to accept AI products. Weak explanation is worse than not explaining at all because it will confuse the radiologists.

It is surprising how low the level of trust radiologists are showing and it is against how high the willingness we identified. Something must be wrong and lead this to controversy. Also, funding and legal risks are the most concerning risk factors according to participants. We suggest that the developer of AI products tackle these risks because risks directly impact trust. More importantly, the developer should provide the users with more knowledge and more support and make sure they trust the product.

Personal attitude is so important because it is proven to be related to so many factors, including social influence, effort expectancy, facilitating conditions, and performance expectancy. It also relates to some factors like motivation, security, and effectiveness.

##### 

#### 4.2.9 Conclusion - Bronte

Overall, this evaluation took the view that by considering the perceived impact of AI on the workflow and how radiologists become accepting of AI in the workflow, would reveal valuable information about workflow impact concerns and how to address them to support the adoption of AI. When investigating the themes identified through the thematic analysis, they were found to be interconnected and explained impacts of AI on the workflow alongside the user experience and acceptance of people who work in radiology.

Facilitating conditions had the least inter-connectedness as it was only found to impact on effort expectancy and intention to use AI, although both of these had very strong correlations. This indicates if an organisation provides good infrastructure that is compatible with the AI, and provides resources and support to its staff, it has significantly decreased the expected amount of effort required and encourages them to use the system. It was found that performance expectancy, social influence, and personal attitude directly impacted on a person’s intention to use AI in radiology. Additionally, social influence and personal attitude influenced performance expectancy and effort expectancy.

Based on the earlier discussion of the social influence and personal attitude themes and in seeking to further understand how someone develops a positive attitude towards AI and would influence others to adopt AI, we propose that it fundamentally comes down to knowledge. When a person gains knowledge of AI, such as an understanding of how it works, its benefits, and its risks, they are more likely to become more accepting of AI in their profession, willing to use it, and trust it.

We propose further research into the importance of AI knowledge as an important factor impacting on the acceptance of AI in radiology, rather than merely a demographic or moderating factor. This could include more information about what participants know about AI to determine what specific knowledge and what level of knowledge is broadly needed for people to become more accepting of AI.

### 

### 4.3 What are the requirements in terms of IT infrastructure? - Cailin

if you decide to not go the AI Marketplace route, common hosting option are:

* Cloud-Based using Qure.AI’s servers which are predominantly through AWS and Microsoft Azure

Or

* Clients Own Services - No specific requirements are listed on provided from the Qure.ai representative. However, common hardware requirements are High-performance servers, AI-enabled GPU’s and large storage systems with backup drives in case of failure.
  + API Based-Private Cloud
  + API Based-Local IT Infrastructure

We suggest using Qure.AI’s cloud-based services. AWS and Microsoft Azure are highly secure reputable cloud hosting options and do not require any specific requirements for the Radiology practice and do all the maintenance for the program. By having the computer server, data storage, firewall, load balancer managed externally, it means radiology practices can focus on important tasks instead of doing mundane tasks such as patching and updates.

### 4.4 Interoperability - How can the data be exported for research and other purposes? - Benjamin

The product qXR is capable of identifying different types of abnormality such as opacity, cavity, nodule, and so on. It classifies the chest X-rays into three categories, normal, abnormal, and undetermined. If abnormalities are detected, qXR will produce the disease it diagnosed for the patient on the top left corner. It can diagnose up to 24 diseases and the accuracy of such diagnoses were provided in the technical performance evaluation. The results of qXR are easy to extract, the user could simply store the patient’s chest X-ray report generated by qXR to local devices or a cloud platform which qure.ai provides. The benefit of using cloud service is it won’t take storage from the hospital’s devices. To export the patients’ report, the hospital just has to set up access to enter the cloud platform. If the results are stored locally, the doctors can just simply download for further analysis. More importantly, the areas of abnormality identified are highlighted so the doctors can easily spot them.

### 4.5 Will the data be accessible to non-radiologists? - Cailin

Pushpendra Rawat revealed: ‘For cloud based & on-premise (intranet) deployments all images can be made available to all physicians with their logins. These can be shared with external physicians as well with just a web-link as well. An original image is also given to remove any biases for the readers.’

### 4.6 Are the AI model’s results interpretable? - Cailin

‘Secondary capture has AI annotations on top of the original image. A DICOM viewer also comes along with it with tools such as measurements, levels, zoom etc.’

The generated report is easily interpretable for all medical healthcare professionals, clearly stating if there is or is not a detection of a certain abnormality. This simple reporting structure decreases the time for a practitioner to interpret the results, hence improving efficiency in the workplace.

The structured report involves:

* Probability score as well as dichotomous output indicating whether each abnormality is present or absent
* Probability score for TB as well as dichotomous output indicating whether TB is likely present or likely absent
* A box indicating the location of the abnormalities
* Abnormalities detected by the product for which a separate abnormality score is given include: abnormal, TB etc.

### 

## 5 Regulatory and Legal Aspects - Bronte

Now we will move onto regulatory and legal aspects that require consideration.

### 5.1 Australian Medical Device Regulation

Medical devices are regulated in Australia under the Therapeutic Goods Act which requires medical devices to be included on the Australian Register of Therapeutic Goods for it to be legal to use them. qXR is a medical device because it is a software used for human beings for the purpose of diagnosing a disease.

However, an issue arises, qXR is not included on the Register which means that it is currently not legal to be used in Australia.

#### 5.1.1 Australian medical device regulation

For qXR to be approved for use in Australia, there are two applications that must be submitted and approved in sequential order. First is the manufacturer’s evidence application which is a TGA certification or evidence from a comparable overseas operator. qXR has evidence from a comparable overseas regulator because Qure.ai has obtained a class IIa CE-mark certification for qXR. This means qXR is a medical device that has been approved for use in the European Economic Area. It is under the European Union Directives which is a comparable overseas regulator recognised by Australia.

Once that is approved, the second application for inclusion on the register can be submitted which includes a range of information, such as, the device’s purpose, classification, and compliance with the essential principles. Qure.ai states that qXR’s intended purpose is to aid in the detection of abnormal findings on a chest x-ray and it will likely have the same classification as it does in Europe. There are 6 essential principles which are mostly concerned about the safety of the medical device. qXR has minimal risks because it is not intrusive and is not affected by transport or storage. Additionally, it has been thoroughly tested and shown to be highly accurate. Although, a risk that qXR poses is that it fails to identify an abnormality on a chest x-ray. This risk is mitigated because it is intended to be used by a radiologist who is expected to also review the x-ray themselves.

#### 5.1.2 Recommendation

Based on reviewing the essential principles and the fact that qXR is approved for use in Europe, it is highly likely to be approved in Australia and simply requires a sponsor to complete the applications for it to be included on the Register. The sponsor is the person or company that imports the goods and is an Australian resident. We recommend that ABC seeks a sponsor from qXR, rather than being the sponsor themselves to avoid the start-up and ongoing regulatory burden of being responsible for qXR itself.

### 5.2 Data protection regulations

Another important legal aspect is data protection, and is particularly significant for qXR because it relies on access to personal and health information to function. Australia’s data protection regulation is contained in the Privacy Act, most importantly, the Australian Privacy Principles that many companies must comply with. Companies that store health information and provide a health service must comply with the principles, regardless of their size, therefore, it is assumed that ABC is already complying with them.

#### 5.2.1 Australian Privacy Principles

There are 13 principles which relate to the management of personal information in an open and transparent way, collection of personal information, use and disclosure of personal information, integrity of personal information, and access to and correction of personal information. Much of this is about the way that ABC manages personal information, however, there are also some points that have an application directly to qXR. For example, reasonable steps must be taken to ensure the security of personal information. Qure.ai states that its products are certified under two international standards, a general information security standard and a standard specific to medical devices and includes the protection of confidential data. In addition it has a cybersecurity team and submits cybersecurity audit reports to US and European regulatory bodies.

#### 5.2.3 Recommendation

Qure.ai does not state whether qXR complies with the APPs, however, internationally it complies with the United States Health Insurance Portability and Accountability Act and the European Union’s General Data Protection Regulations.

Based on this and our assessment of the principles, we recommend that qXR will comply with Australia’s data protection regulations, however, ABC will likely need to take steps to ensure their ongoing compliance, such as updating their privacy policy.

### 

## 6 Financial and Support Services Considerations

### 6.1 What is the licensing model? -Benjamin

A demo is available on qure.ai's official website for the potential clients to test the functionality of the product qXR. Also the user can get familiar with the user interface of the product before actually using it. The payment method is pay-per-use, and the pricing differs based on different geographical locations. We highly recommend that the client use AI marketplace to purchase qXR because it can help from installation, deployment, to implementation. Although we are not able to get an exact pricing of qXR, the price of qXR is described as ‘competitive’ according to the feedback from the AI marketplace. The factors affecting price are number of annual studies, number of concurrent users, number of sites, and some other factors.

### 6.2 How are user training and follow-up handled? Benjamin

There is a specific four-steps workflow that the company qure.ai follows when it comes to user training. It is considered to be comprehensive. The first step is the Orientation, the orientation works as a general introduction of the product and its functionality. The second step is technical readiness which is to guide the clients on the specific technical operations for qXR. The company is expecting the clients to fully understand how to operate the product and how to interpret the results at clinical level. More importantly, the clients will learn to use the interface of the qXR application. The SLA, service level agreement, will be demonstrated and comprehended by the clients after the first two steps. Then the company will help the clients to deploy the system to their own devices including computers and chest CT scanner. The full deployment process takes approximately 6 hours, consisting of 2 hours for gateway deployment and 4 hours of on-precise deployments depending on the speed of the internet. Then the last step is volunteer training and support. If the clients encounter any problem while using the product, the support team will respond as quickly as possible.

As for the follow-up, there are two teams responsible for the communication with the clients, the client engagement team and project delivery team. The clients can contact the client engagement team directly if they encounter any problem. The project delivery team can help the clients to design the best technical approach of using this product for a specific task according to the nature of the task.

### 6.3 How is the maintenance of the product ensured? Cailin

a dedicated team which provides support for all our clients in running the services. There is no AMC for cloud deployments.

### 6.4 How will potential malfunctions or erroneous results be handled? Cailin

The process for identifying potential malfunctions and erroneous results from qXR scans may vary depending on the particular implementation However, some possible methods for identifying such errors could include comparing the results of multiple qXR scans of the same patient to look for discrepancies and reviewing error logs,

If a potential malfunction is detected from qXR, the system will automatically generate a notification to the user. The qure.ai team will then investigate and determine the root cause.

There are a few potential consequences of erroneous results:

1. Inaccurate results – If qXR malfunctions, it could provide inaccurate results. This could lead to patients not getting the proper treatment or diagnosis.
   1. False positives – This could lead to patients being unnecessarily sent for expensive and invasive molecular testing for conditions they do not have
   2. False negatives – This could lead to patients not getting the treatment or diagnosis they need and can then go on to infect others
2. Delayed results – Another potential consequence of a qXR malfunction is delayed results. This could cause a delay in treatment or diagnosis, which could be detrimental to the patient.

Medical healthcare readers must be aware of the consequences of potential malfunctions and erroneous results from qXR by qure.ai. Having the clinic’s physicians consistently check the software can prevent incorrect diagnoses and time delays, as they can temporarily stop the upcoming chest x-rays that need to be processed. This will allow time for the software to be fixed and in the meantime getting a radiologist to do the entire CXR reading .

## 7 Conclusion

### 7.1 Recommendation - Ed

Based on the previous evaluation of qXR, we recommend that ABC proceeds with a pilot of qXR. As AI technology is an emerging technology in medical imaging, ABC should consider using AI products to help with diagnosis. In addition, qXR performed the best in all certified AI products in terms of DTA studies. Using the qTrack will also potentially be beneficial for ABC in efficiently diagnosing and managing databases. Moreover, since there are not many studies related to UX of qXR, having a pilot study for qXR can help ABC identify the performance of qXR in their unique workflow, a pilot program will also help in changing the attitude and thinking of using qXR for radiologists.

Neysa’s Final presentation expectations:

- ECLAIR

- Everyone present, cameras on

- Send report the day before

- How DTA compares

- Why choose qXR

- Considerations when investing

# 

# 

# 

# MARKING CRITERIA

## Goals

The final product delivery and deployment to the client is a major client milestone that should help each group to deploy and demonstrate a working version of the complete system/software in the client's desired environment. The aim is to ensure that the project requirements and features agreed in the project scope:

* Are correctly implemented as per the client's perspective;
* Run well at the client's deployment/testing/production environment (esp., when appropriate, integrates well with the client's existing infrastructure);
* Meet the client's expectations;
* Handed over all the project deliverables (source code, documentation/reports, user guides, installation/deployment script, and/or anything else agreed with the client as a deliverable).

## Contact Your Client

Each group should contact your client to arrange a 1-hour (or longer) meeting between the client, all team members and other essential stakeholders (if available). This deployment should be scheduled ahead and completed during week 12. Remember to use dates instead of weeks when making appointments with your client as most clients are not familiar with the academic calendar and terms. When contact your client, you are supposed to check the following points with your client to prepare your final product delivery:

* The formal group name and project name (used in our course) so it can be used in the assessment form by your client.
* Agenda for this meeting with a timeline
* All user stories/requirements that will be deployed on the client's desired infrastructure and/or demonstrated in the final product.
  + User stories/requirements that you both agreed to demonstrate in the final delivery of the product.
* Brief about the format of the deployment and demo you are planning to run:
  + Check with your client for their preferences; e.g., platform or specification of the deployment/testing infrastructure environment, requirements to integrate integration with other systems/tools
  + Check with your client what they would like to see in the final deployment and demo.
  + Check with your client what project documentation and deliverables they would like to get as part of the hand-over of the project outcomes (source code, documentation/reports, user guides, installation/deployment script, or the files you submitted via canvas).
* A brief summary of the set of tests you have already carried out (as a report) with results (including acceptance criteria) and the ones you plan to perform during deployment and their results (including full system testing, user acceptance testing, UI)
* Essential project stakeholders who should attend this meeting as well and their expectations.
* Q&A discussion in week 3 and week 6 to collect feedback from the client and other project stakeholders.
* Any other thing that the client expects you to present in the final delivery.
* Inform your client the whole session of the deployment/demo will be recorded and the recording will be uploaded to canvas.

**Note:** The deployment/demo will be assessed by your client.

## Requirements and Suggestions

* You are supposed to prepare a professional presentation that is consistent with the above contents and includes any other information that your group and the client agree to present in the final delivery.
* You may reuse some of the slides from your ‘final project demo and presentation’ slides but keep in mind that the client is likely to be more interested in results rather than processes and wiki documentation. So, you may need to present a list of user stories that you developed, success/acceptance criteria for results, a live demo of the implementation of the user stories (functional and non-functional requirements if applicable), system design and architecture, quality of technical work, a few live tests for each user story, and live deployment.
* Prepare a list of items involved in the project delivery including:
  + Documentation such as reports, presentations, demos/videos
  + Source code hand over and code documentation
  + Guidelines for how to use the system/software, deploy and run the software/system
  + Platform or specification of the deployment/testing infrastructure environment, requirements to integrate integration with other systems/tools/accounts
  + Any other items that should be part of the project hand-over
* You can prepare a summary slide of the things that cannot be covered in this delivery due to the time limit (e.g., report of the extensive testing you carried out, how did you address non-functional requirements, high-level system/software architecture with external components/APIs you integrate the internal system with).
* During the deployment/demo, collect feedback from the client about what you have demoed and presented and think of ways to respond or provide suggestions to handle it.
* Record the entire session and upload it to canvas.
* Update all the feedback on your project wiki page so that it can be accessed by your tutor.
* Rehearse, rehearse, rehearse! Please well prepare and rehearse your deployment/demo and presentation as a group. Before the meeting with the client, test and run what you're planning to deploy/integrate/run/test during the demo.

**Note:** You need to keep all the records of client deployment and email correspondence as evidence on your Bitbucket site. Also, remember to cc your tutors every time when you contact your client.

## Submission

* Each group has to submit all the artifacts involved in the final product delivery (e.g., slides, video demo, reports, recorded meeting) as a zip folder through the submission link provided in this Canvas page by 23:59 pm, 28th October.
* Make sure your client aware that they are supposed to submit his feedback by 23:59 pm, 4th November.
* There should be one submission per group.
* Add a text file into your zip folder including the information of team members (names, SID, and Unikeys), formal group name (e.g., SOFT3888\_TU17\_03), formal project name (e.g., P01 - Autonomous Car Using Neural Networks and Computer Vision), client name, your tutor name and tutorial time. (For group members who are doing multiple capstone units/projects, add a note to indicate the student name, and the other group name and the project name.)
* Name your submission using the group name and project number (e.g., COMP3888\_TU17\_02\_P02.pdf).
* Make sure your client records the correct course's project name, group name and team members’ name (as those in Canvas). The clients are asked to record absentees whose marks will be scaled down.

## Client Assessment

Your client will assess your final deployment/demo following the below criteria:

* Quality of deployed/demoed user stories; the final product is running effectively and efficiently, and implements all user stories/requirements that agreed to be in the final product (30%)
* Quality of the testing (functional and non-functional tests run successfully; reasonable testing coverage reports; satisfactory testing results) (10%)
* Quality of the group’s ability to respond to requested changes and reflect on the client’s feedback from the first deployment/demonstration and weekly meetings (15%)
* Client's satisfaction with the deployment of the final product/system and the hand-over of the project's artifacts as per the client's requirements (20%)
* Quality of the deployment/demo presentation (e.g., presentation, live demonstration, live tests) (10%)
* Client's satisfaction with the group’s problem solving, making initiatives and their creativity to solve problems and deliver quality outcomes/results (10%)
* Client's satisfaction with the group's communication, ability to complete assigned work, and member's contribution to the overall work (5%)

**Note:** No-contributing and/or absent students will not get the group mark.