



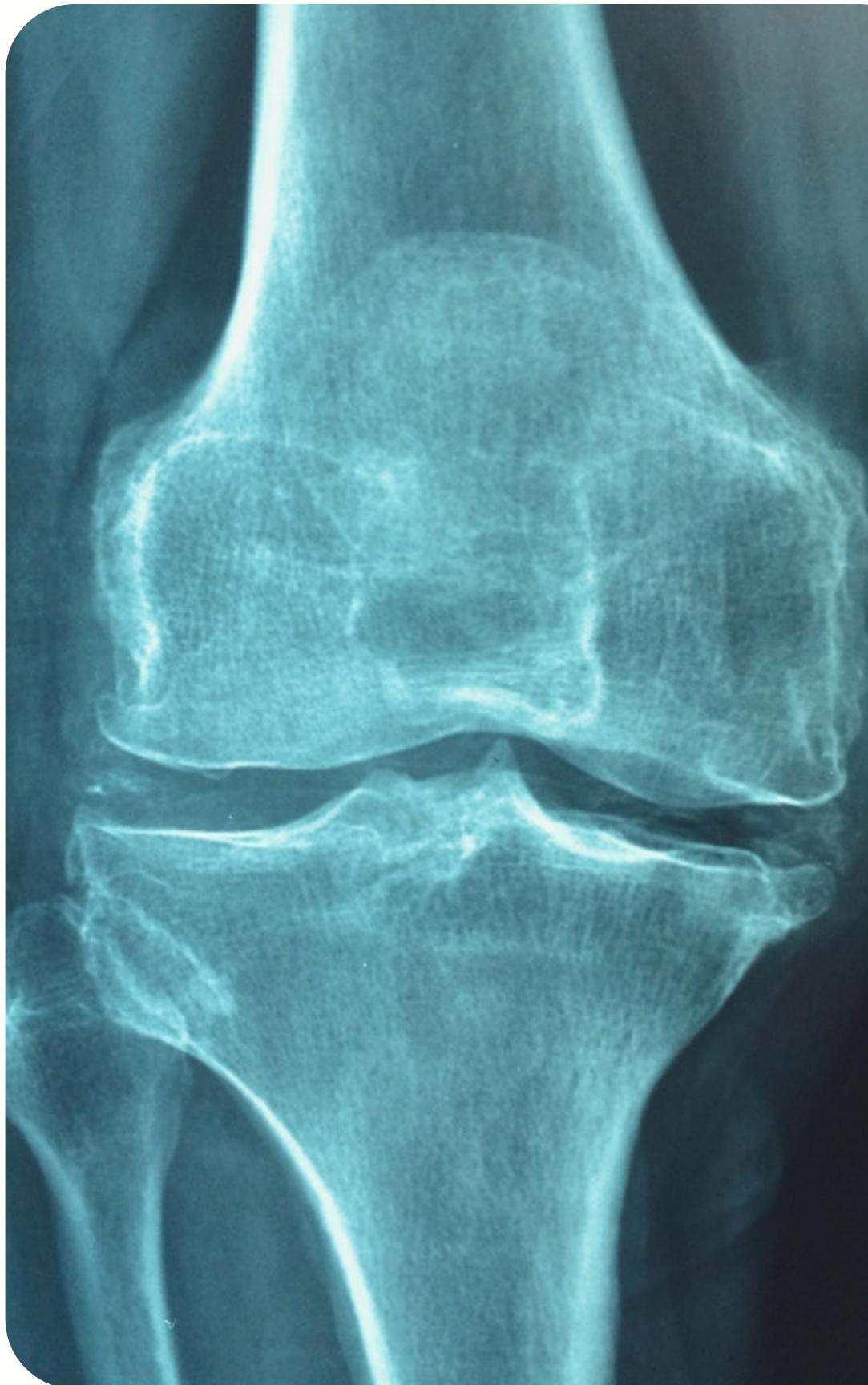
## GROUP H3

Timo Debono  
Katrín Grunert  
Andesh H. Haribhajan  
Jonathan A. Harris, MSc  
Andreea Mateescu

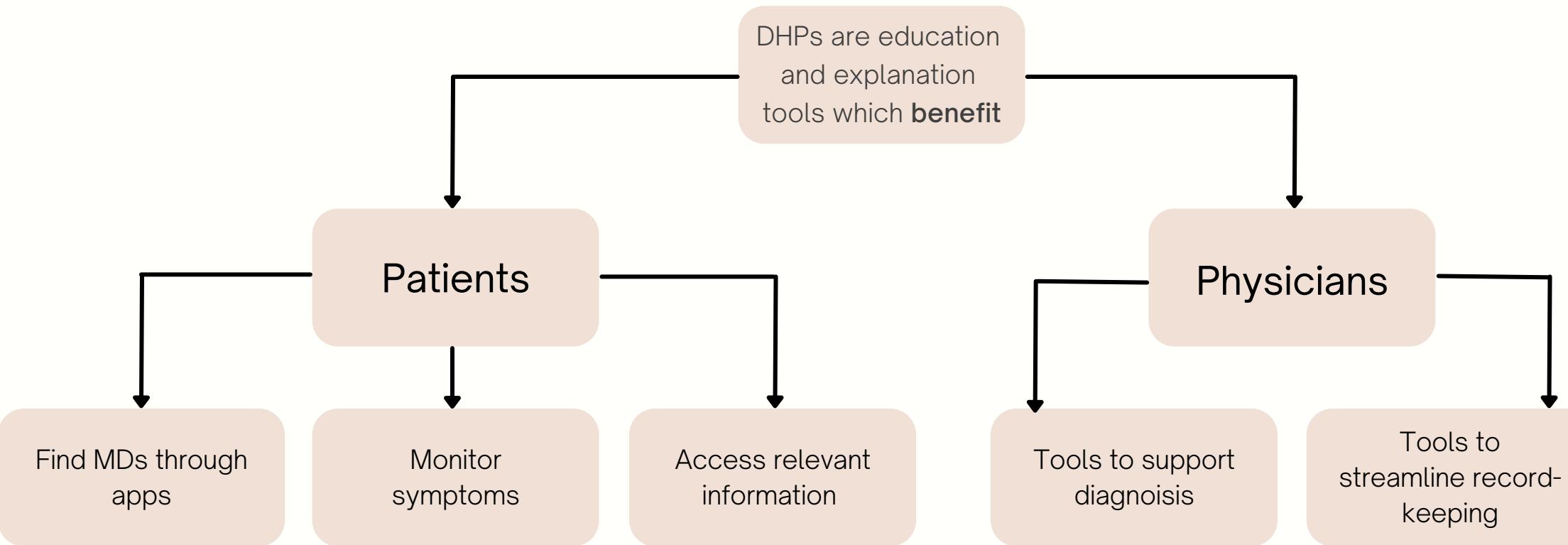
# KNEE-XAI



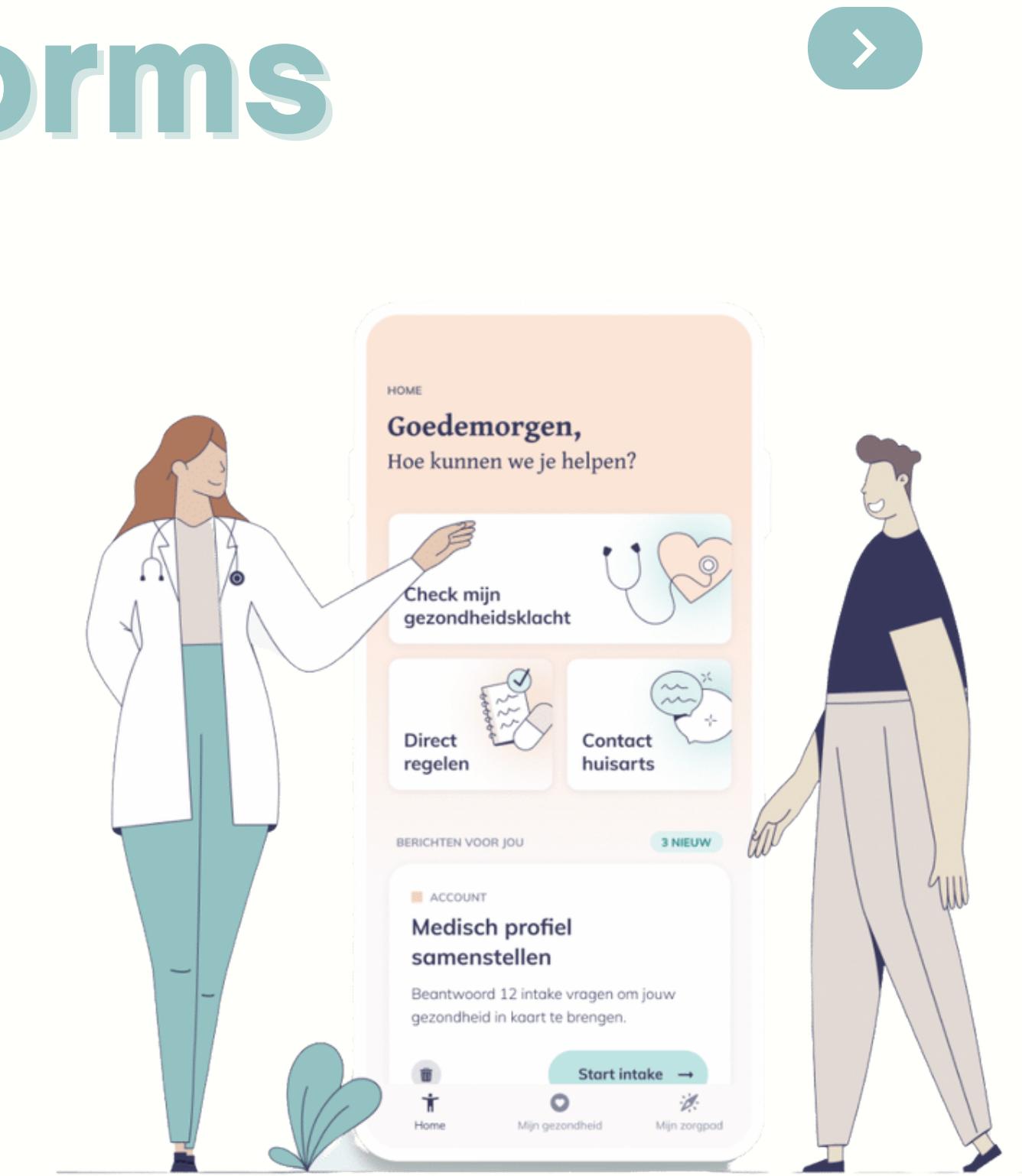
Meet your **smart assistant tool** to  
guide and supplement your  
osteoarthritis evaluation.



# Digital Health Platforms



*The integration of Artificial Intelligence (AI) into Digital Health Platforms (DHP) is a natural progression that may improve the physicians' decision process.*



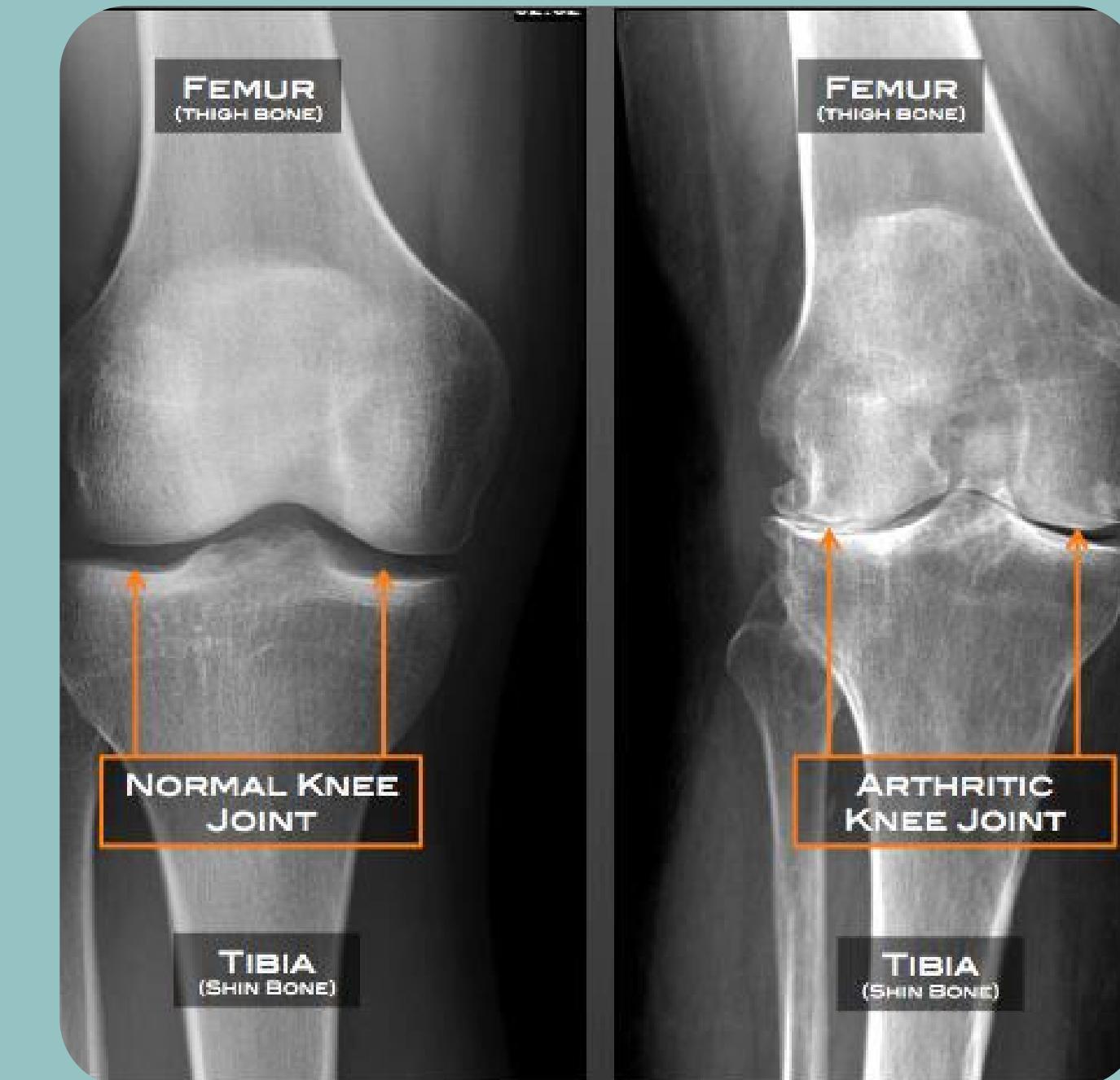
# Opportunities for DHP-based solutions

**Knee osteoarthritis is prevalent, yet evaluation is subjective**

- The degenerative condition affects **approximately 654.1M worldwide** [1]
- Defined by narrowing, osteophytes, sclerosis, deformation
- Single Kellgren-Lawrence (KL) score [Grade 0-4] attempts to describe multiple characteristics

AI prediction may support the evaluation of knee OA to predict **KL score** and identify the presence of defining characteristics of OA.

**BIGGEST LIMITATION:** the ability to explain the classification result.



KL score based on the severity of several structural characteristics [2]



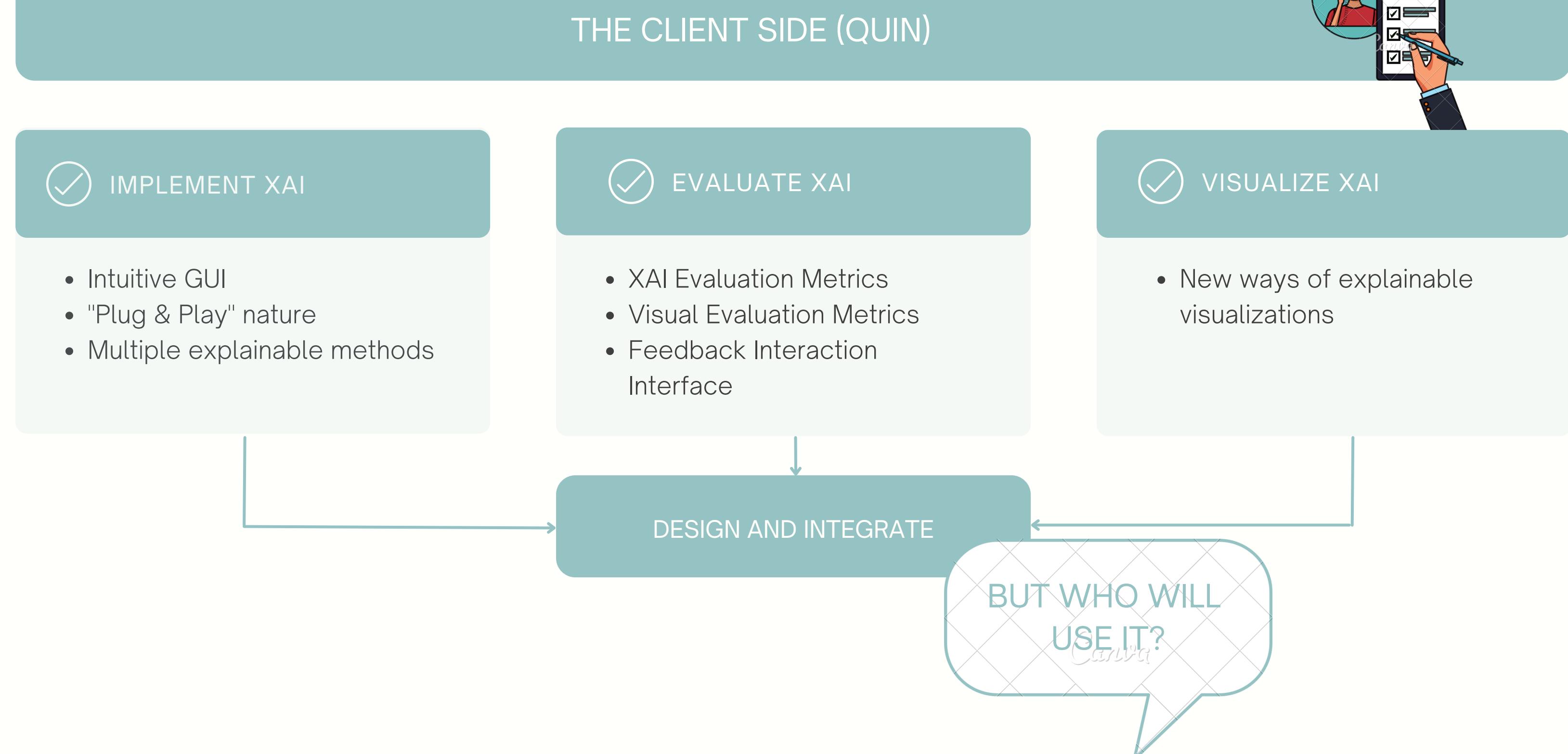
is a DHP looking to expand their radiography analysis capabilities.

" What methods are there to **implement, evaluate, and visualize XAI** for a digital health platform (DHP) that takes a **human-in-the-loop** into account for the detection of knee osteoarthritis? "

# Ideation

KNEE--XAI

# Stakeholder requirements



# User requirements

## THE USER SIDE (CLINICIANS)



### CLINICIAN PROFILE | **Informaat**



- Motivated by maximizing the patient care
- Wants an easy-to-use, intuitive app
- Wants an app that simulates the diagnostic process
- Wants the possibility to perform subjective evaluations



### INTERVIEWS

Chiropractie Watergraafsmeer

van der Schouw M.  
Chiropractor, DC,  
MChiro, BPhysio



Drs. Jaggoe - Haribhajan G.  
Physiotherapist

**ActiveHealth**  
Rehab & Physio Clinic  
*'Moving Towards A Healthy Life'*

Bersaoui M. MSc.  
Physiotherapist



Drs. Ooft. J.R.  
Orthopedic Surgeon

Drs. Lierop Van M.G.  
Rehabilitation Specialist

## KNEE-XAI

- Inspect and interact with the X-ray
- Aid in discovering what clinician 'missed'
- Few 'options' - simple to use

- Ability to second guess
- Summary of the detailed findings
- The process of the app should simulate the clinician's process

# Inspiration sources

## FROM THE MARKET

Surgimap is a spine-based DHP that aims to improve physician workflow. It facilitates the importation, annotation, evaluation, and exportation of spine and hips X-rays.

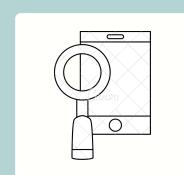


Image analysis:  
angles & distances



Case planning  
tool

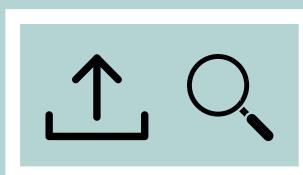


Tailored packages  
for different needs

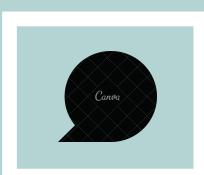


Tool for clinicians'  
collaboration

### Relevant considerations:



file upload &  
inspection



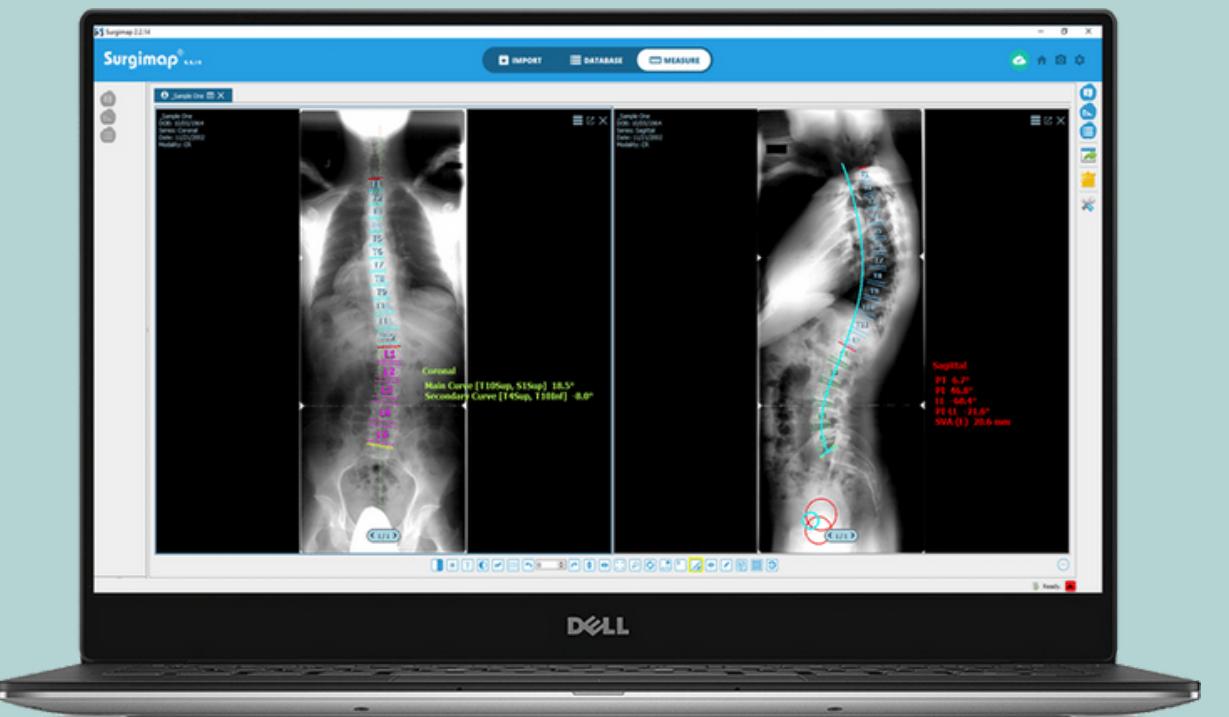
enable  
annotations



generate  
report



UI  
format



## FROM ACADEMIA

### Human-in-the-loop concepts

**A SURVEY OF HUMAN-IN-THE-LOOP FOR MACHINE LEARNING**

Xingjiao Wu<sup>1,2</sup>, Luwei Xiao<sup>2</sup>, Yixuan Sun<sup>3</sup>, Junhang Zhang<sup>2</sup>, Tianlong Ma<sup>1,2,\*</sup>, Liang He<sup>1,2,\*</sup>

<sup>1</sup> Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai, China  
<sup>2</sup> School of Computer Science and Technology, East China Normal University, Shanghai, China  
<sup>3</sup> Fudan University, Shanghai, China  
(wxingjiao2885, louishxw008, madmaxkgb, junhangzhang68)@gmail.com, {lma, lbe}@cs.ecnu.edu.cn

**ABSTRACT**  
Human-in-the-loop aims to train an accurate prediction model with minimum cost by integrating human knowledge and experience. Humans provide training data for machine learning applications and directly accomplish tasks that are hard for computers in the pipeline with the help of machine-based approaches. In this paper, we survey existing works on human-in-the-loop from a data perspective and classify them into three categories with a progressive relationship: (1) the work of improving model performance from data processing, (2) the work of improving model performance through interacting with model training, and (3) the design of the system independent human-in-the-loop. Using the above categorization, we summarize major approaches in the field; along with their technical strengths/ weaknesses, we have six classification and discussion in natural language process, computer vision, and others. Besides, we provide some challenges and opportunities. This survey intends to provide high-level summarization for human-in-the-loop and motivates interested readers to consider approaches for designing effective human-in-the-loop solutions.

Fig. 1: The development cycle of model.

941v2 [cs.LG] 29 Nov 2021

entropy MDPI



### Review Explainable AI: A Review of Machine Learning Interpretability Methods

Pantelis Linardatos , Vasilis Papastefanopoulos and Sotiris Kotsiantis

Department of Mathematics, University of Patras, 26110 Patras, Greece;  
vasilis.papastefanopoulos@upatras.gr (V.P.); sotiris@math.upatras.gr (S.K.)

\* Correspondence: p.linardatos@upatras.gr

**Abstract** Explainable Artificial Intelligence (XAI) has led to its widespread industrial adoption, with machine learning systems demonstrating superhuman performance in a significant number of tasks. However, this surge in performance, has often been achieved through increased model complexity, turning such systems into "black box" approaches and causing uncertainty regarding the way they operate and, ultimately, the way that they come to decisions. This ambiguity has made it problematic for machine learning systems to be adopted in sensitive yet critical domains, where their value could be immense, such as healthcare. As a result, scientific interest in the field of Explainable Artificial Intelligence (XAI), a field that is concerned with the development of new methods that explain and interpret machine learning models, has been tremendously reignited over recent years. This study focuses on machine learning interpretability methods; more specifically, a literature review and taxonomy of these methods are presented, as well as links to their programming implementations, in the hope that this survey would serve as a reference point for both theorists and practitioners.

**Keywords:** xai; machine learning; explainability; interpretability; fairness; sensitivity; black-box



1. Introduction  
Citation: Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S.

Artificial intelligence (AI) had for many years mostly been a field focused heavily on theory, without many applications of real-world impact. This has radically changed

### Computer Vision / XAI methodologies



# Design process

## KNEE--XAI

# Deriving product requirements



## MODULARITY

### FRONTEND

- Simplistic and linear process with clear choices
- Zoom, pan, and manually crop the image
- Note-taking
- User-calculated KL-score (**Human-in-the-loop**)
- Report Generation

### BACKEND

- Present & explain the calculated KL-score
- Choice of XAI methods: simple, sophisticated, revealing (a heatmap with an opacity slider, bounding box, arrows)
- Evaluation of different XAI methods

# Deriving product requirements



## MODULARITY

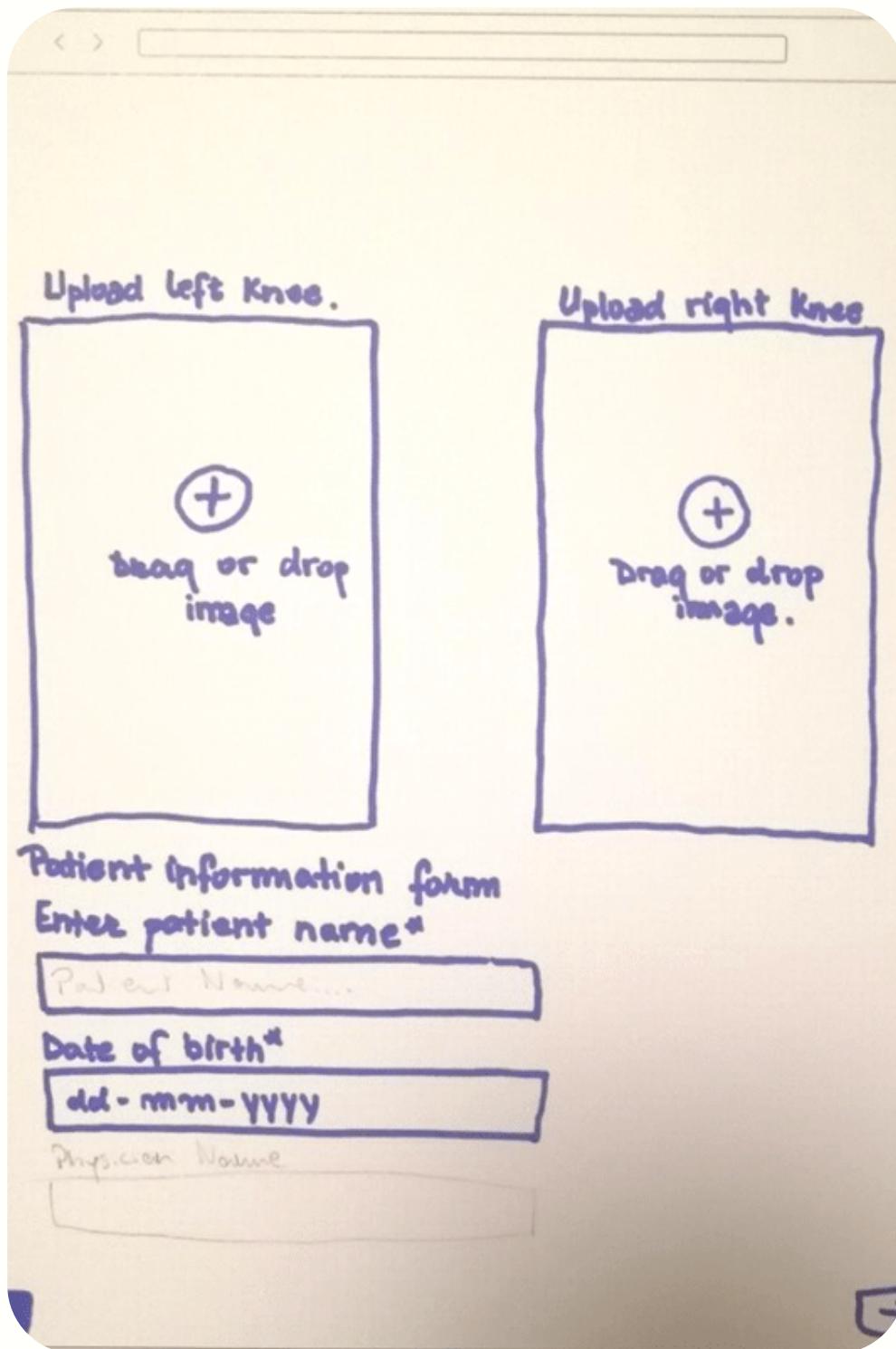
### FRONTEND

- Simplistic and linear process with clear choices
- Zoom, pan, and manually crop the image
- Note-taking
- User-calculated KL-score (**Human-in-the-loop**)
- Report Generation

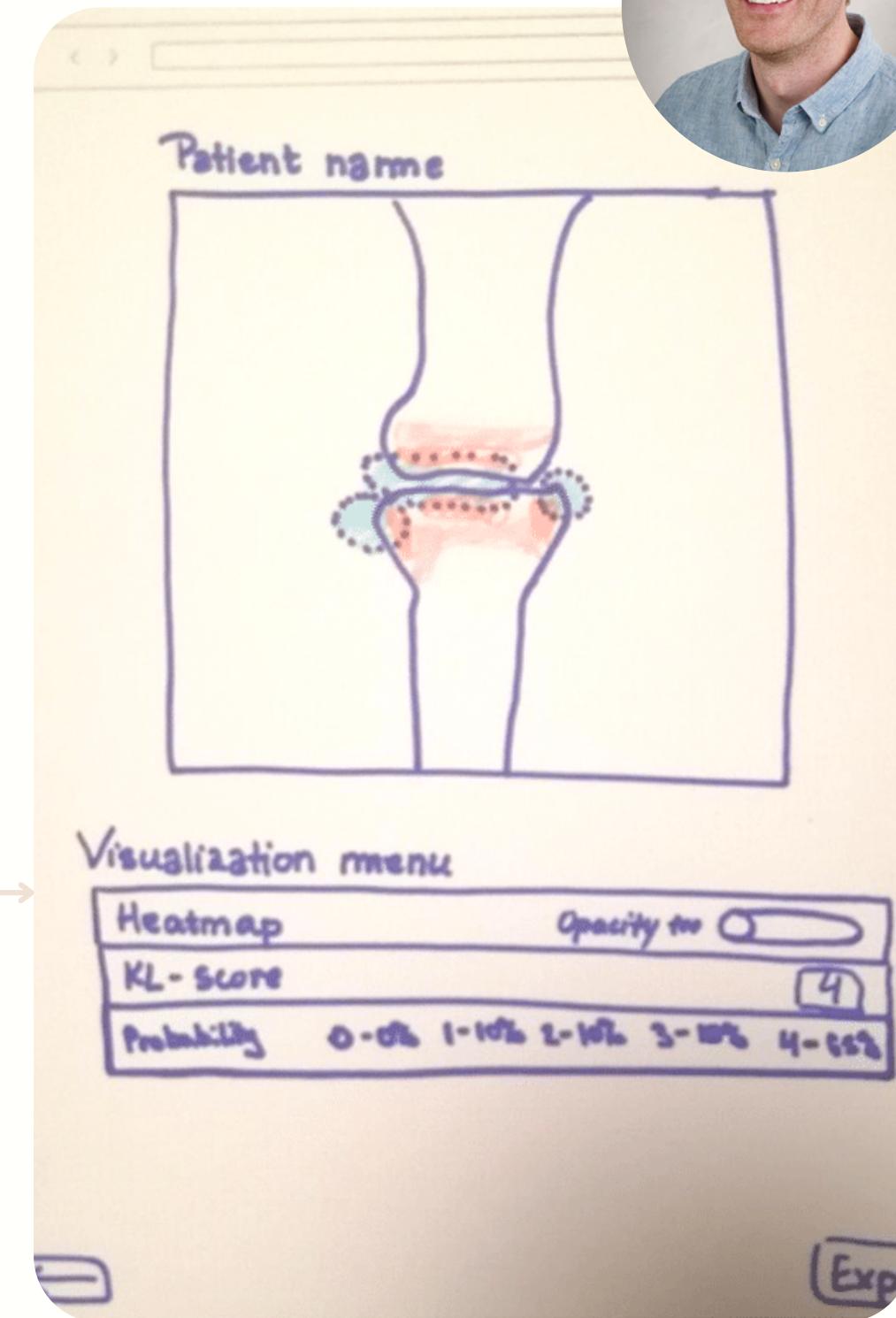
### BACKEND

- Present & explain the calculated KL-score
- Choice of XAI methods: simple, sophisticated, revealing (a heatmap with an opacity slider, bounding box, arrows)
- Evaluation of different XAI methods

# Paper prototyping



Upload X-ray page

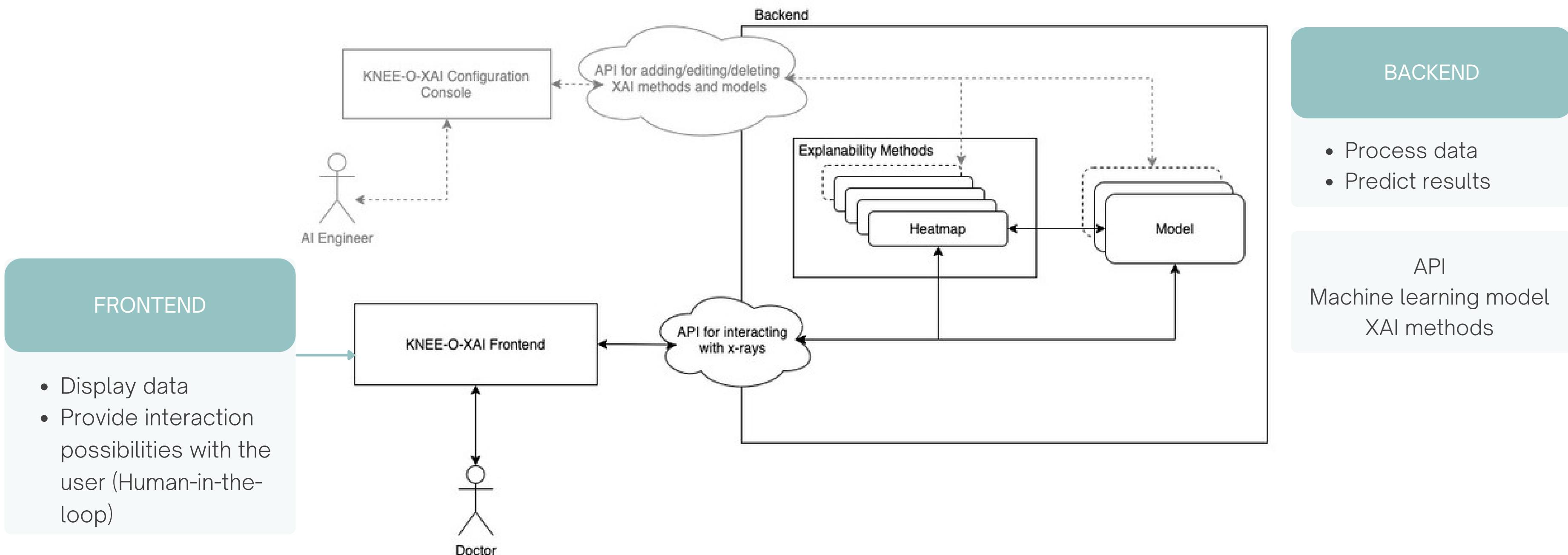


Analysis page

Jon: And now we are done, right?

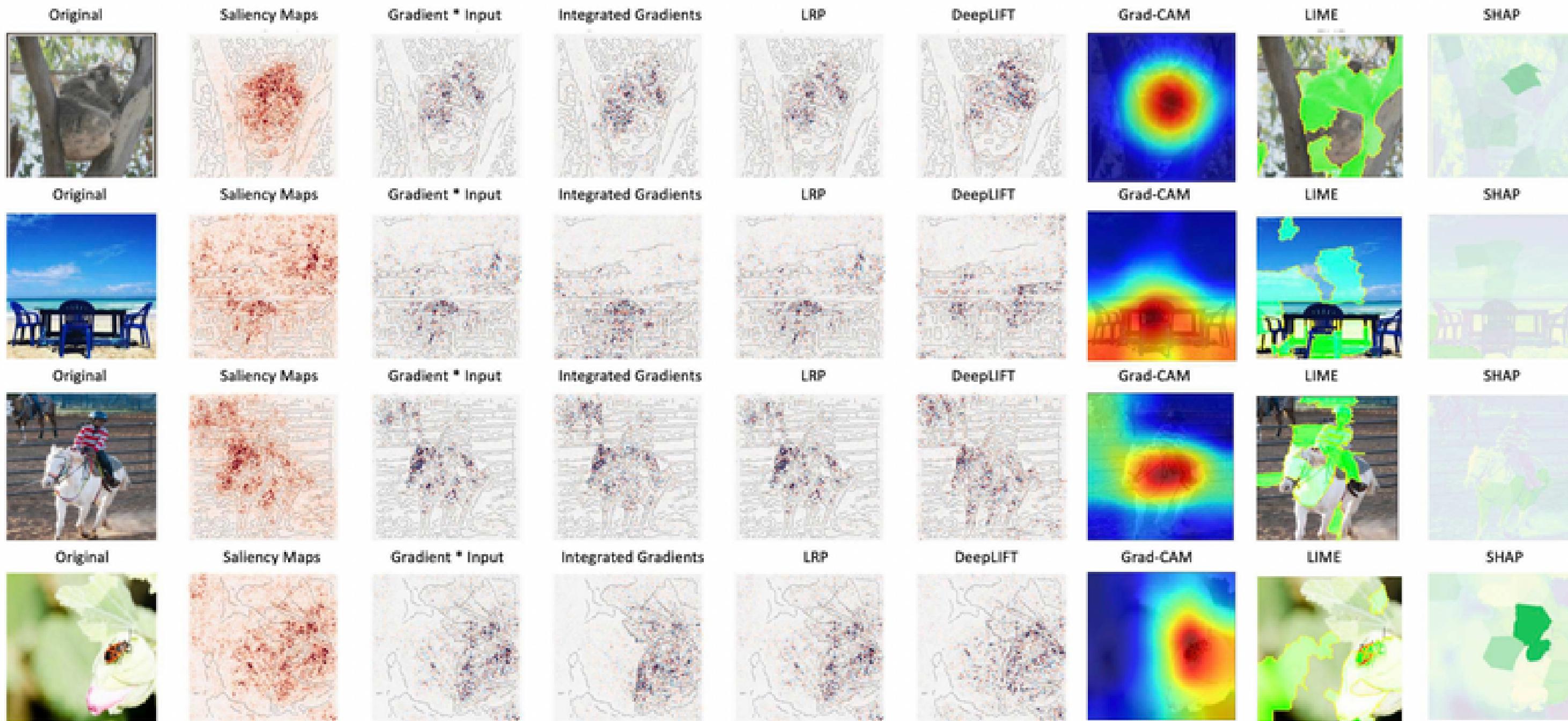


# Prototype architecture



# A pool of XAI choices

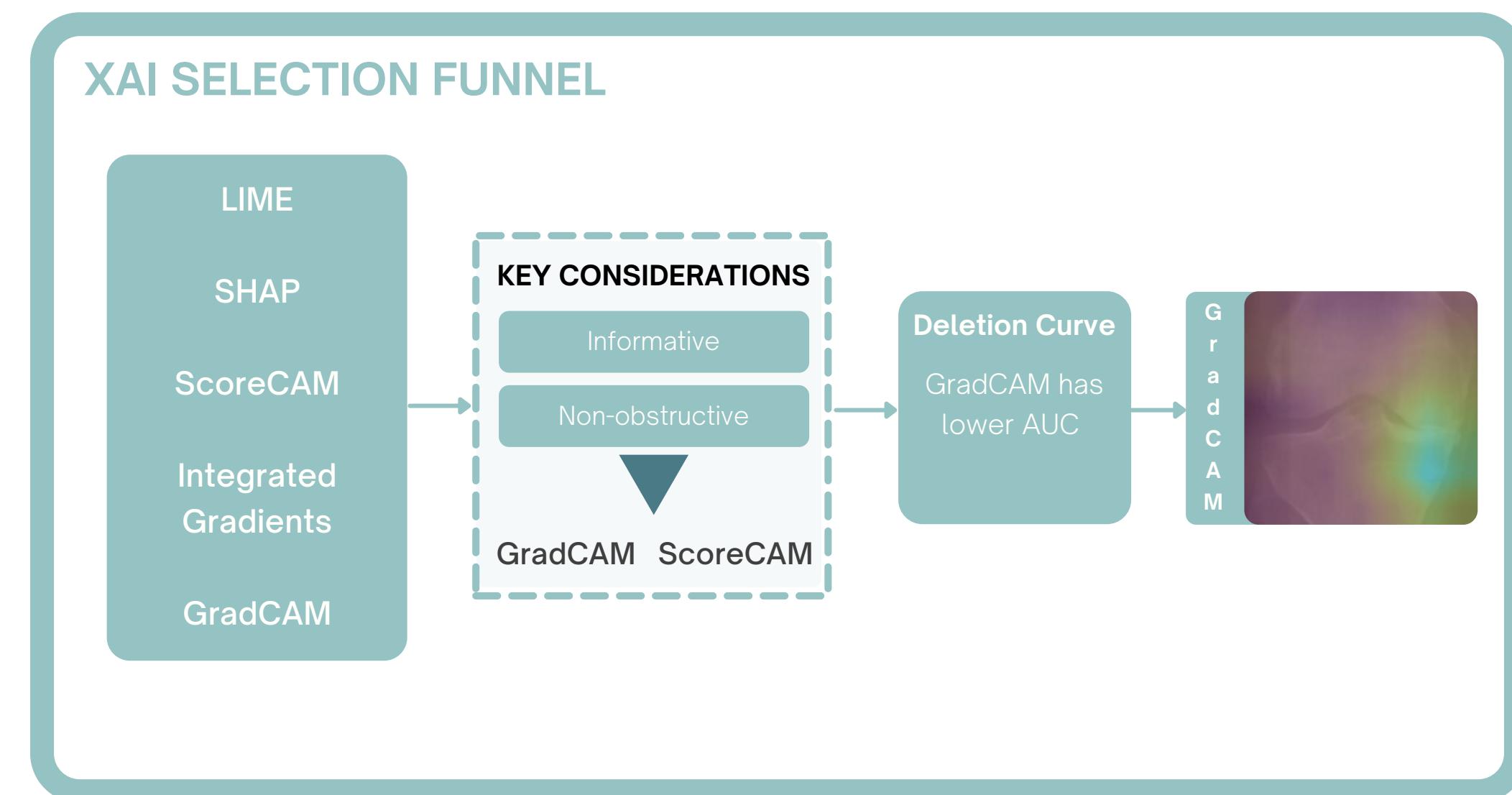
WHAT FITS OUR PROBLEM BEST?



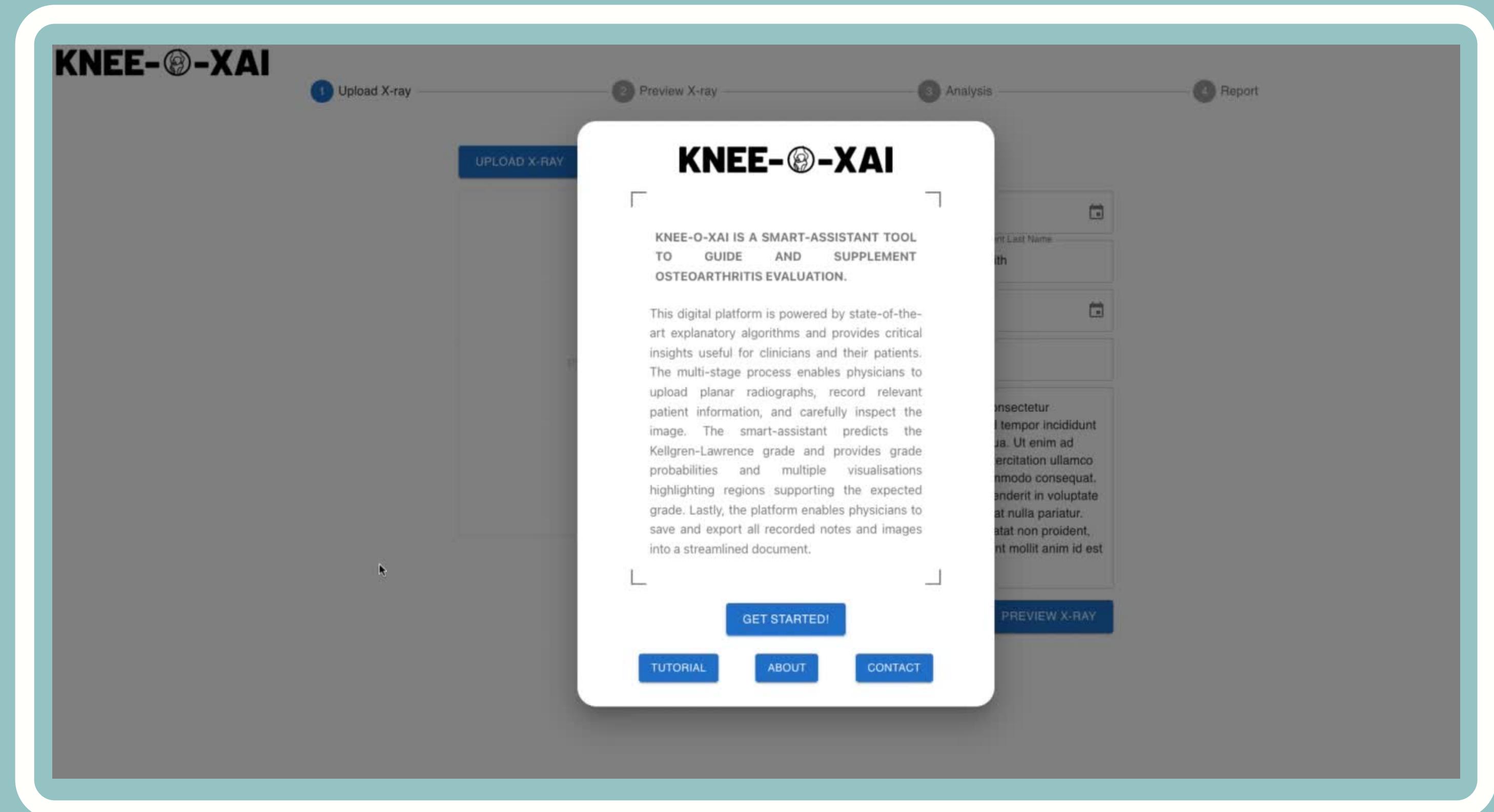
# Explored models and XAI methods

**MODEL ARCHITECTURES**

Model	Test accuracy
ViT (ImageNet)	0.7150
ViT	0.6794
DenseNet-169	0.6374
ResNet-152	0.5361
VGG-19	0.5289



# Video walkthrough





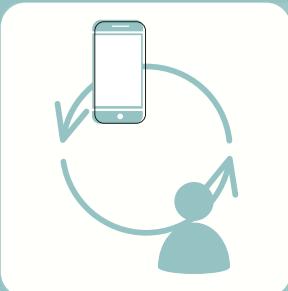
# Evaluation

KNEE--XAI

# Key design decisions



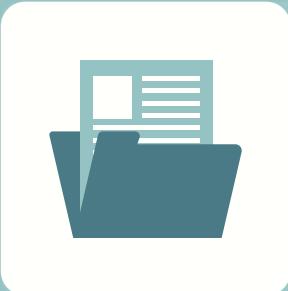
## FRONTEND



Human-in-the-loop  
annotations

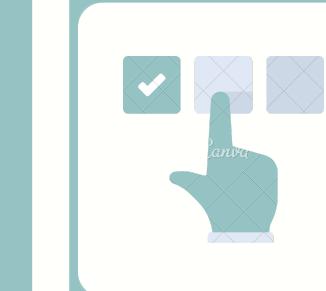


Interactive  
Visualizations



Generate Report

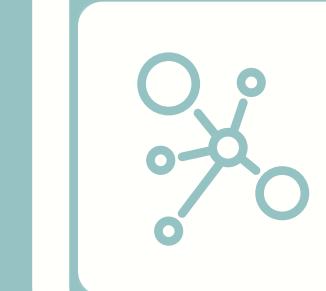
## BACKEND



Choice of XAI  
approaches

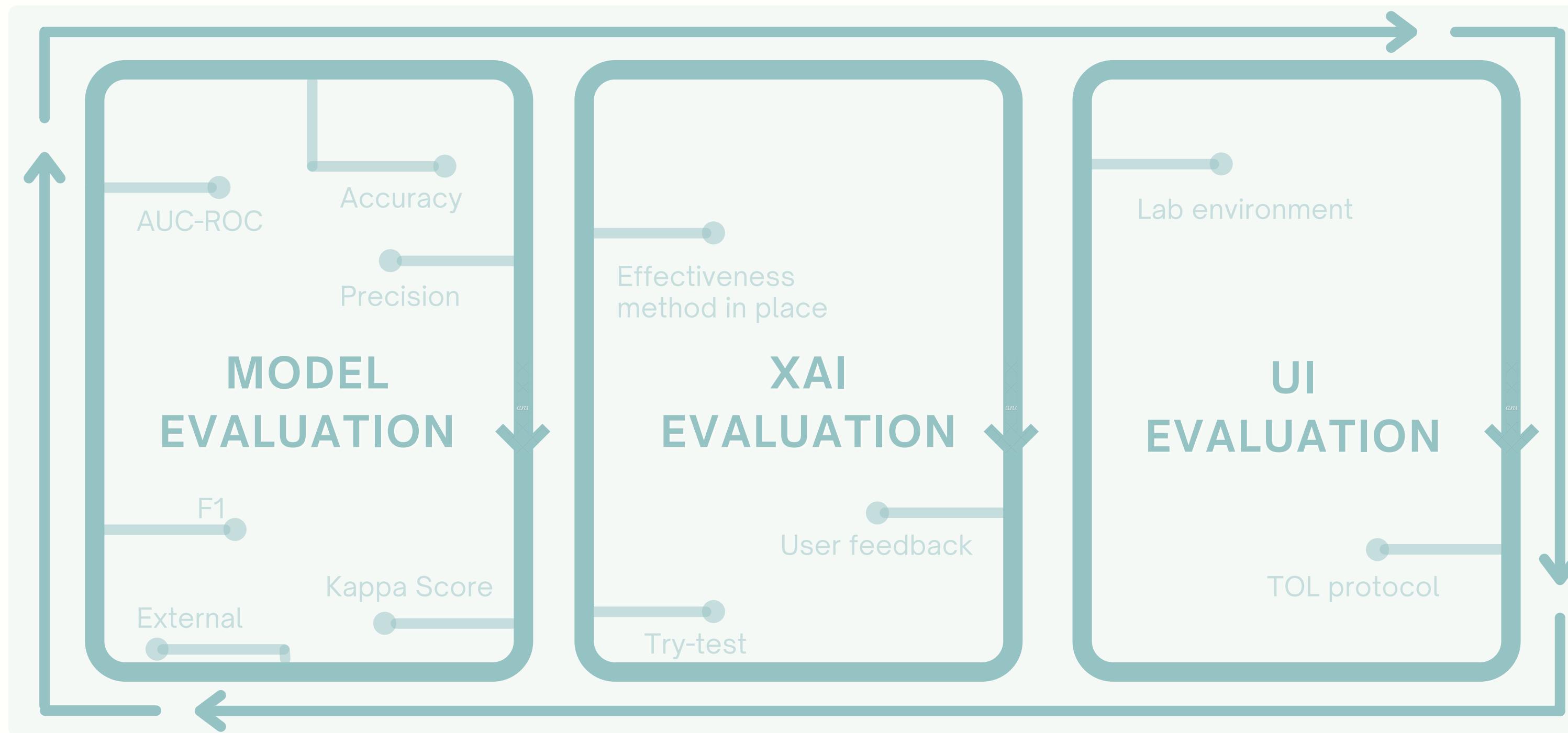


Visualization of XAI  
results

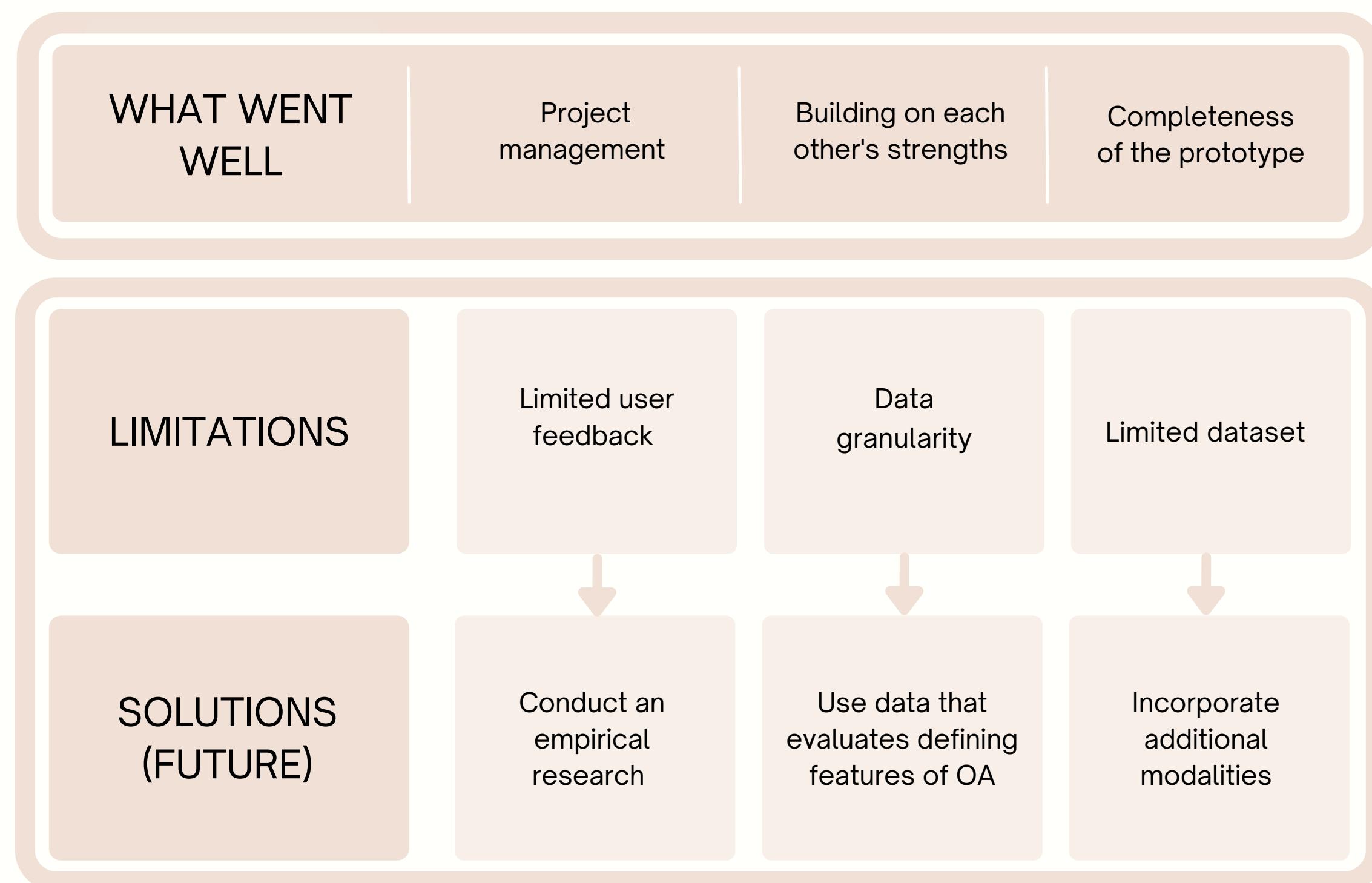


Model architectures

# Product Evaluation



# Reflection

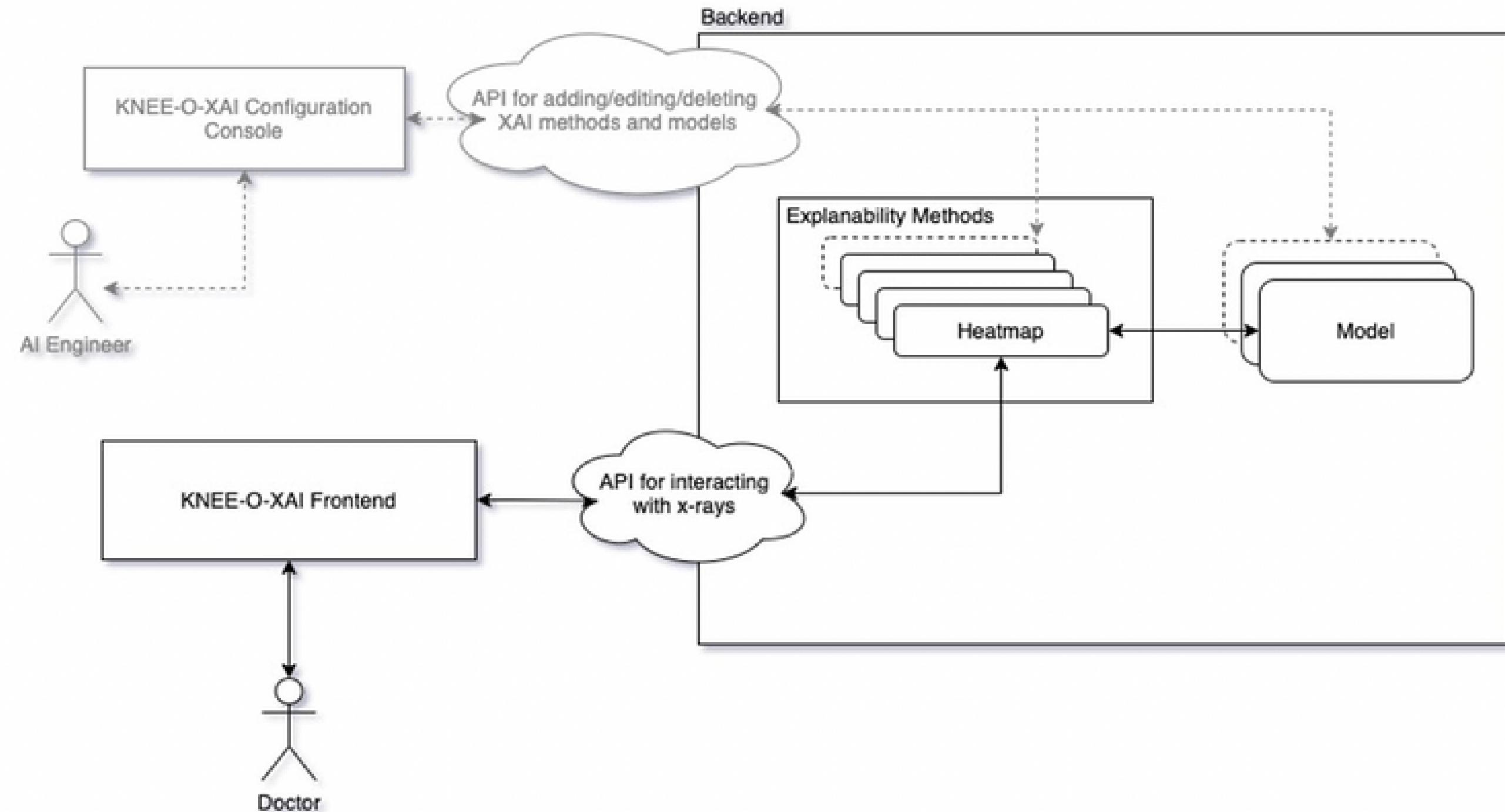




Thank you!  
Questions?



# Appendix





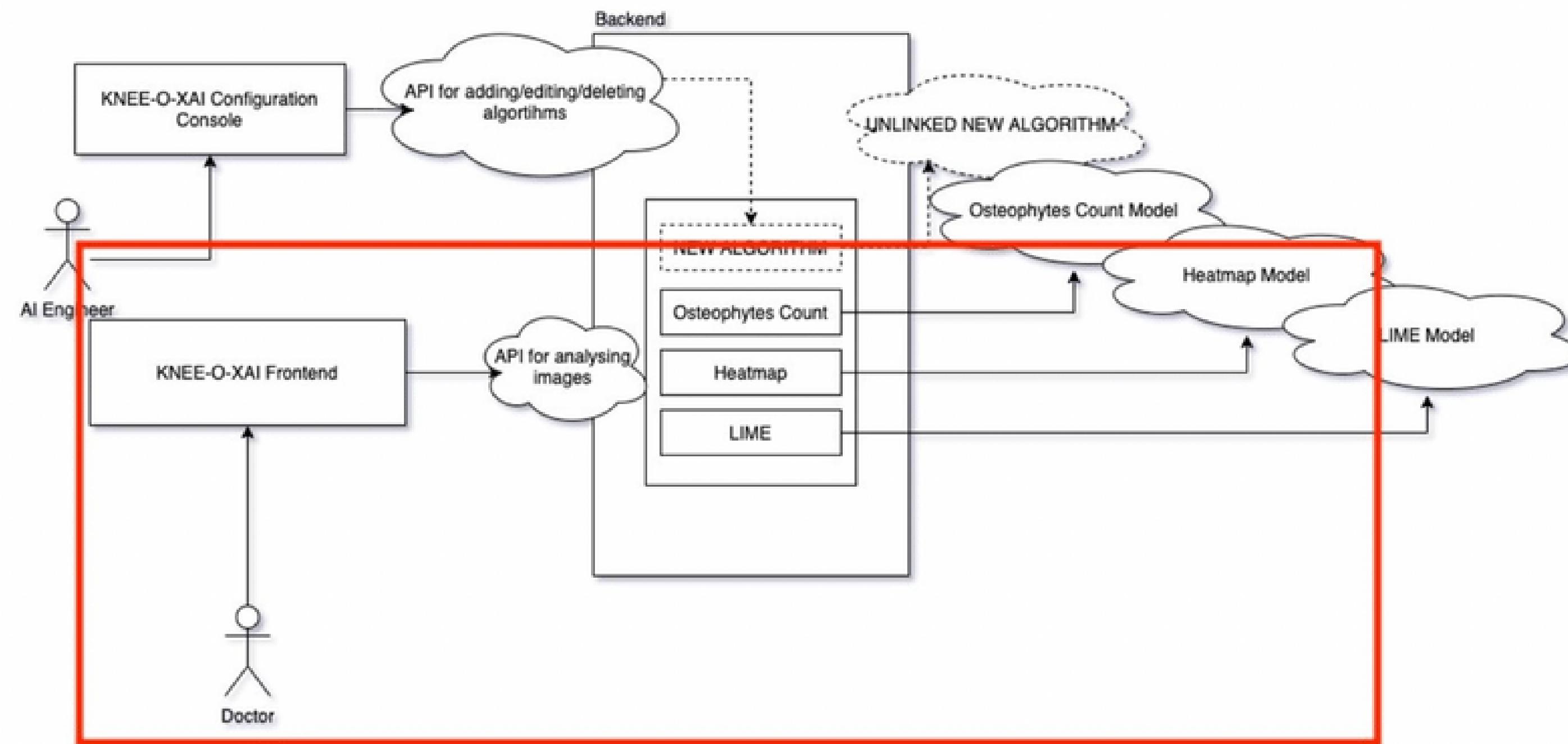
## User evaluation

**Agile Development of the Product with feedback and evaluation:**

**Alpha version evaluation by gathering UX data from select users in an HCI-lab using the TOL-protocol**

**Beta Launch and Empirical Evaluation of XAI and the GUI by real users**

**Product Launch with periodical evaluation and updates based on new requirements and understanding**





# KNEE--XAI

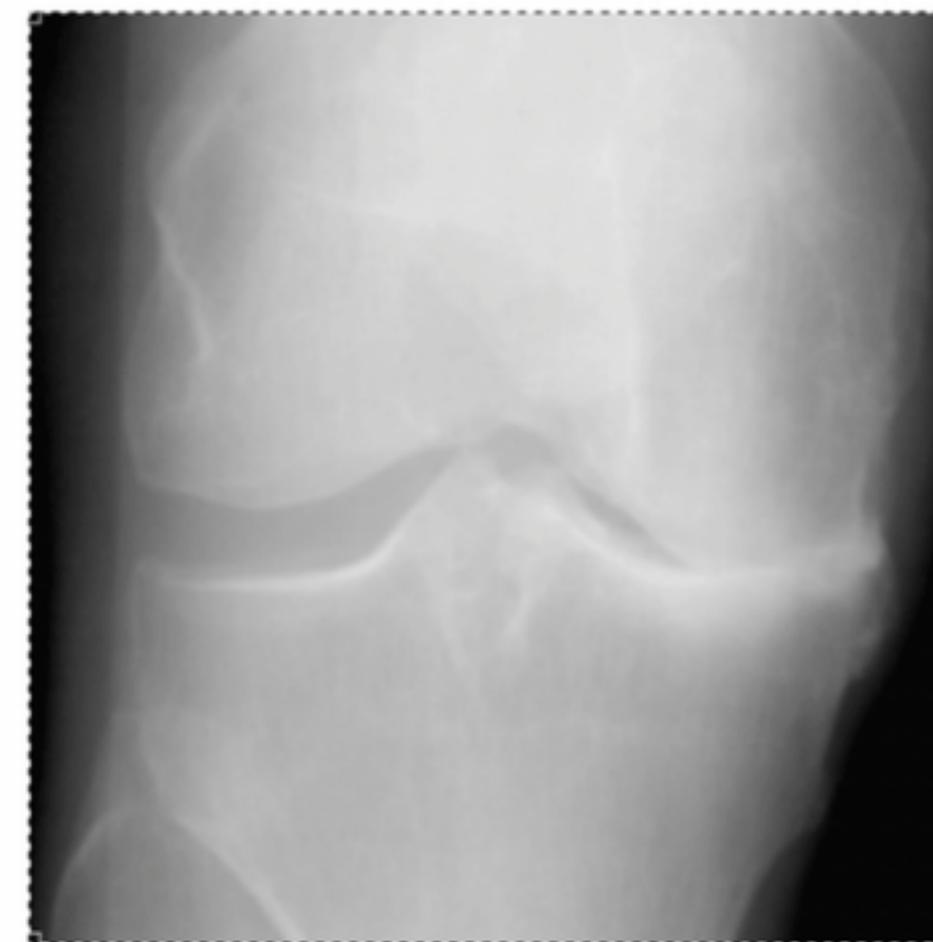
1 Upload X-ray

2 Preview X-ray

3 Analysis

4 Report

UPLOAD X-RAY



Appointment Date



Patient First Name

Patient Last Name

Date of Birth



Physician Name

Add Medical Notes

PREVIEW X-RAY



## Initial evaluation from the clinician

Inspect the x-ray by panning, zooming in the image

KNEE--XAI

1 Upload X-ray      2 Preview X-ray      3 Analysis      4 Report



### Enter Radiographic Grading of Knee OA

Select Osteophyte Formation  
2 - Large

Select Joint Space Width  
3 - Gone

Select Subchondral Sclerosis  
3 - Severe Sclerosis w/ Cyst

Select Deformation  
2 - Strong

User Calculated Kellgren-Lawrence Score

4



Add XRay Notes

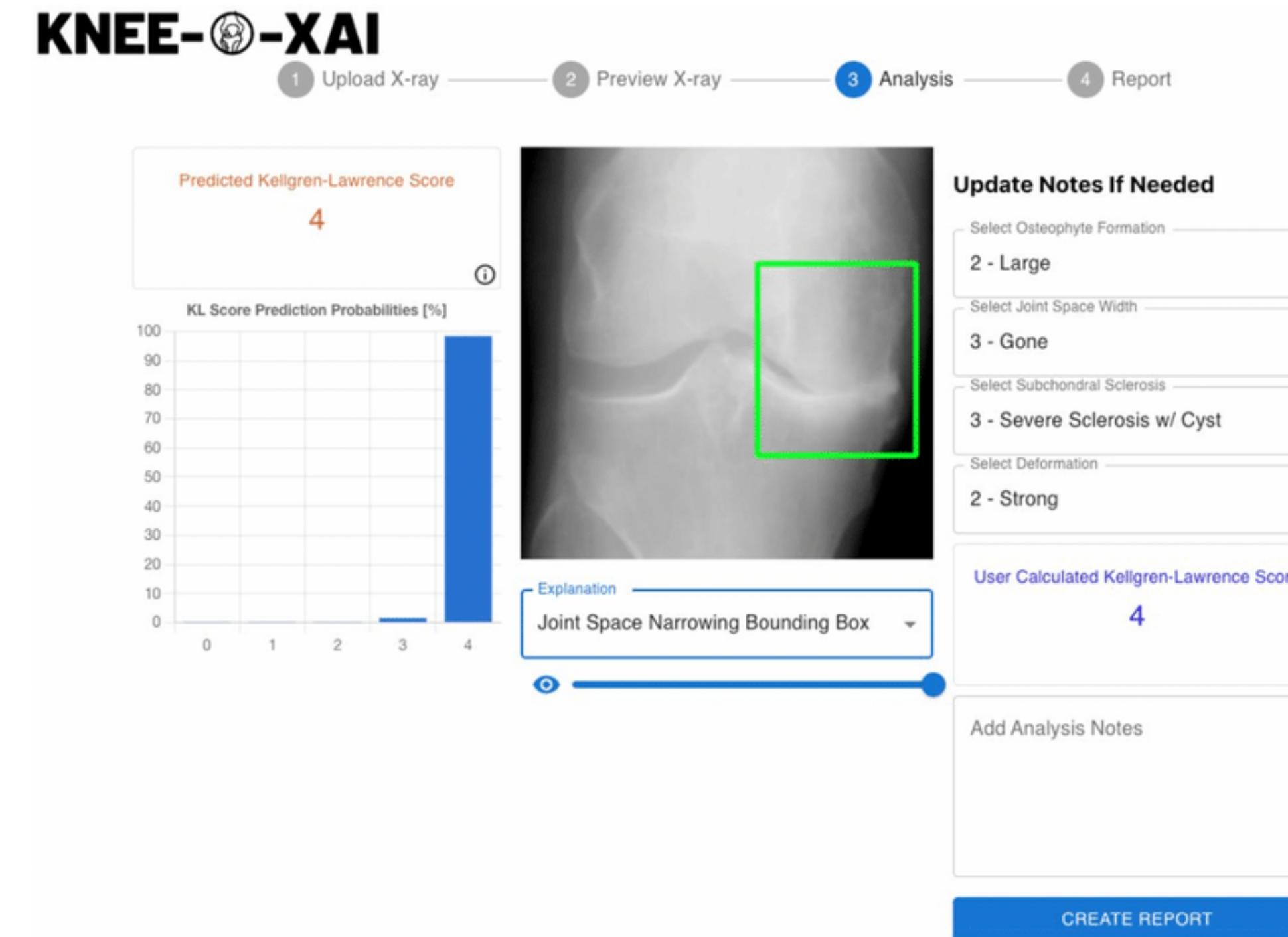
GO TO ANALYSIS



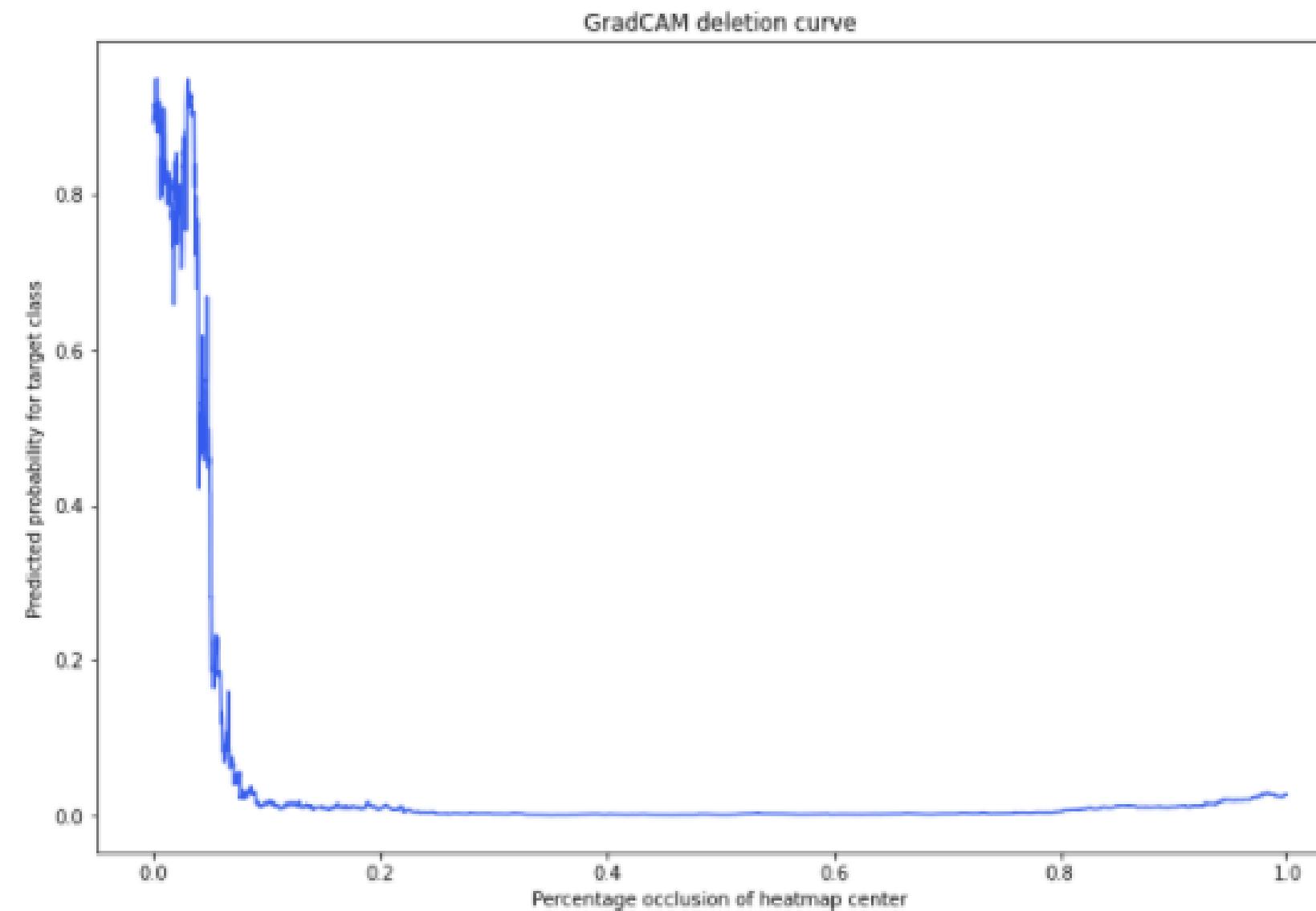
- Side by side comparison of the smart assist evaluation and the clinician's evaluation

Possibility to edit the previous clinician evaluation

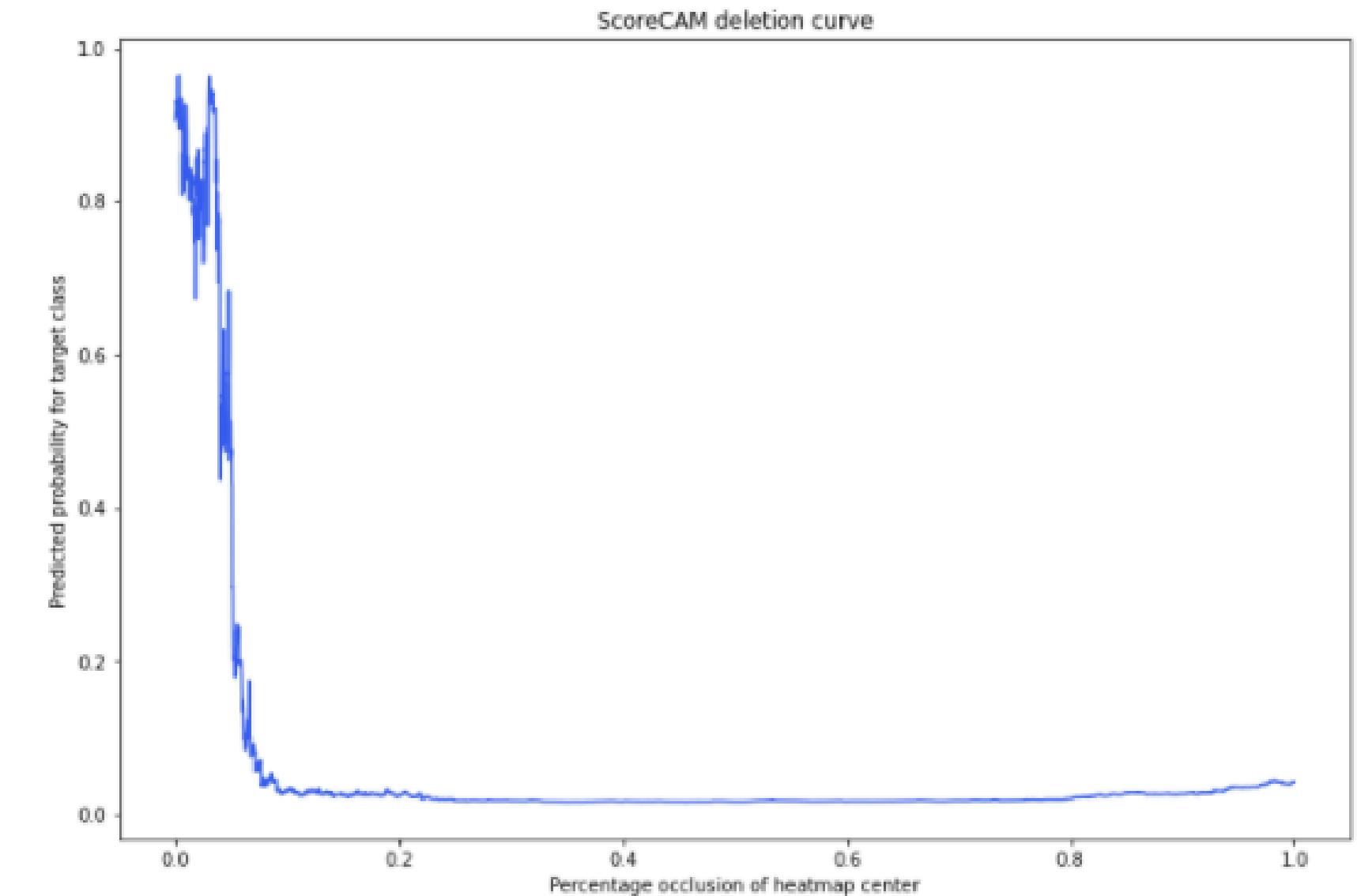
Gives confidence in the prediction by displaying the probability distribution



# Deletion curves



AUC = 0.048



AUC = 0.063

Note: Calculations averaged over a subset of the test set (n=20)

# Interview Example

1. What are sections/areas of interest when analysing an x-ray image for knee OA and what are the deciding factors when making a diagnosis?

*[Which are the sections of the knee that might indicate OA when analysing an X-ray? Which are the main stages of knee OA? Are there any standard thresholds that can be analyzed for each stage? (Example: a combination of sclerosis with other features might indicate an advanced phase of OA)]*

---

The x-ray is just a part of the diagnosis. In order to rule out Rheumatic OA, for example we do

blood tests as well. The general health, age, weight and diet are also important indicators.

For the x-ray I mostly ask for both the anteroposterior and lateral images of my patients.

In the x-ray we generally look for if the space between the joints have narrowed or not. This can be seen by comparing the other areas of the joint with the narrower area. In very severe cases the whole space will be narrowed. Another thing we look for is the occurrence of osteophytes. These are seen better in the lateral x-ray images.

We also look out for other things such as minor fractures, that might be causing the pain and help us rule out the occurrence of OA in many cases.

2. What does the process of explaining an x-ray image to a patient following a diagnosis (OA vs no-OA) look like?

---

I'm not sure what you mean by this question, but normally I would explain the patient in the same way I explained to you. In short, I will go over each 'proof' of OA in the x-ray with the patient.

Normally patients don't ask too many detailed questions. However, I like to show them the narrowing and osteophytes in their pictures. If there is no OA, I generally don't need to convince the patient. If there are still doubts I will advise them to get a second opinion ad or do an MRI for a better diagnosis. In my whole career this has happened only once. And it turned out that the patient didn't have OA.

3. How many experts are included in the decision/evaluation of an x-ray when diagnosing knee OA?

---

Normally only I will evaluate and decide about my diagnosis. I will however consult with my colleagues in case of doubts. In some cases the radiologist might add a note if they observe some

# Evaluation (Survey) Example

Grading Knee X-Ray ☆ Print Feedback

Questions Responses 1 Settings

How severe are the observed osteophytes? \*

0 1 2 3 4

None  Severe

Do you observe any joint space narrowing? \*

Yes  No

How severe is the joint space narrowing? \*

0 1 2 3 4

None  Severe

Would you advice for (clinical) action to be taken? \*