SHENKAR ——— שנקר
ENGINEERING. DESIGN. ART . הנדסה. עיצוב. אמנות
The Pernick Faculty of Engineering . הפקולטה להנדסה ע"ש פרניק
המחלקה להנדסת תוכנה
Software Engineering Dept.

SyncAlong

Software Engineering B.Sc. Final Project

Software Design Document

Authors:

Name	ID	Email	Phone
Shani Levi	302853619	shanilevi01011@gmail.com	0543247884
Ayman Wahbani	209138155	aymanw199816@hotmail.com	0508241000

Supervisor:

Dr. Marcelo Sihman & Mr. Eyal Nussbaum

13/12/2021

Approved by: Eyal Nussbaum Date: 05.02.2022

Approved by: Marcelo Sihman Date: 06/02/2022

Content

1. Introduction

a. System Overview

SyncAlong is a system designed to evaluate synchronized activities performed simultaneously via peer-to-peer video connection.
The system will use movement detection and image analysis and will be displayed using a web platform with peer-to-peer secured video-communication.

The users will receive live and static data:
Live stream and poses data - such as collecting the evaluation point of the users for a time of X frames in a peer-to-peer video connection .
Static data - such as centralized information on synchronization results (levels, exceptions, etc), sync-control, elderly members control, profiles and more.

Trainer user will set a meeting time with the trainee user who is under his responsibility, the meeting will be set by filling out a relevant section information about the meeting (time, date, setting of activity time and type of activity). The trainee will see all the appointments scheduled for him in the order of the action times.

b. Purpose

- A platform that enables synchronized activity with family and friends remotely.
- A system for the elderly user to increase motivation to do more sports activities that can not physically participate in joint sessions.
- A multi-participant training system designed to support interpersonal training in a synchronized manner between the participants.
- Positive response to participants when syncing in motion for the purpose of raising motivation.

c. Scope

The scope of the SyncAlong system is about synchronized joint physical activity.
Addresses a few key points:

- Real-time Communication
- Client-server computing or networking
- 3D human pose estimation involves estimating the articulated 3D joint locations of a human body.

d. Definitions and Acronyms

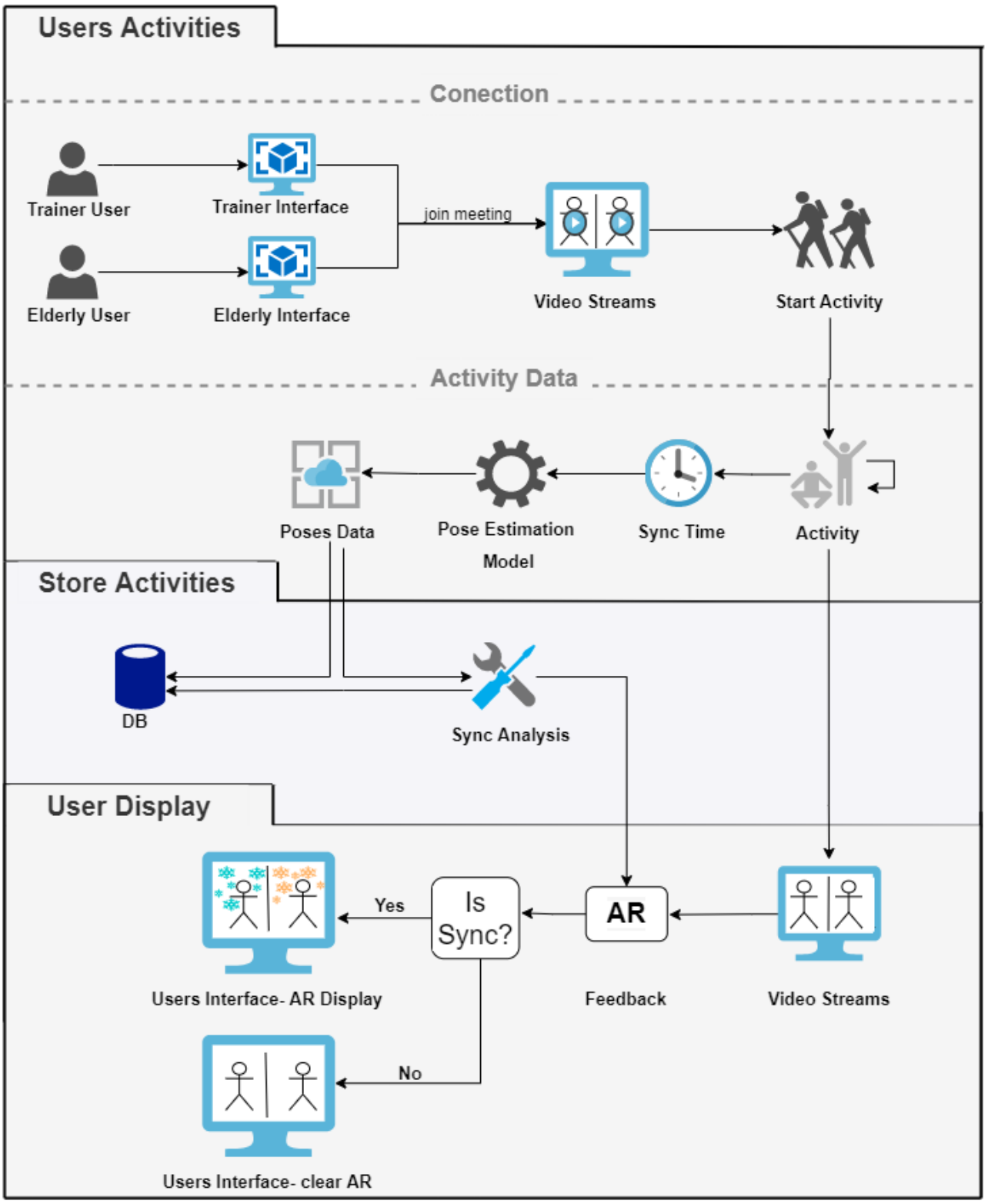
Real-time system	Send and receive data instantly online across multiple clients. Two-way flow.
Signaling	WebRTC allows real-time, peer-to-peer, media exchange between two devices. A connection is established through a discovery and negotiation process called signaling.
Outcome	Numerical result of the synchronization calculation
Synchronization of movement	Synchronization of movement is defined as similar movements between two or more people who are temporally aligned.
Sync Rang	Defined range for synchronization indices in the range x-y and within it sub-settings for synchronization level
High sync	Very close approximation of Synchronization
Low sync	Not nearly the sync setting
Non sync	Inconsistency of pose estimation or motion time estimation in performing shared motion between users
Suggestion (Improve Result)	A drawing showing a correction on top of a video taken
Period time	A period of time that elapsed from the moment an body part is in a certain place until the moment it returns to the position
Feedback	Numeric / visual / verbal view of the quality of synchronization between users
Section	A list of varying sizes (few parts) of workouts whose total predefined time
Coach	A system function which builds an automatic training set for the user and gives feedback at the end of the section

e. Constraints

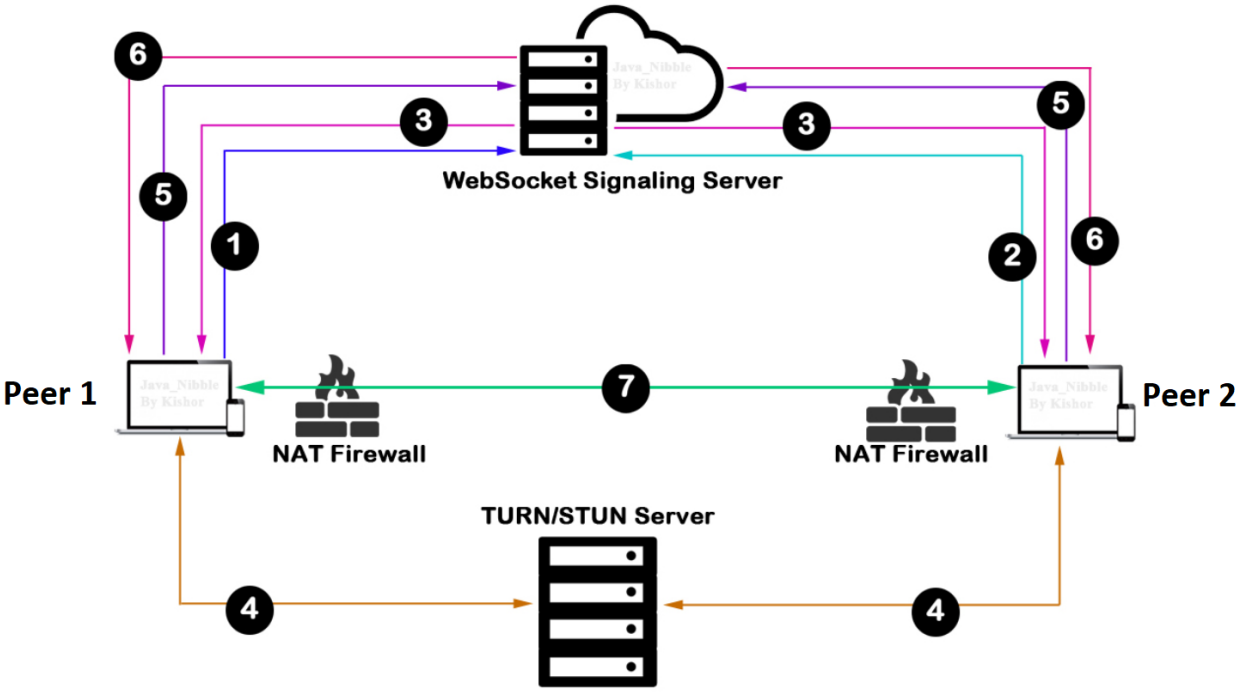
- Time delay between different computer systems.
- Environmental configurations e.g. the placement of cameras.
- Segmentation issues that are dependent on the potentially cluttered or obscuring background of a family room (e.g. its objects, decoration and lighting).
- Segmentation issues dependent on the nature of the participant - adult, pensioner, handicapped, or other structural issues. The clothes they are wearing in relation to the background and how baggy, bulky or obscuring the clothes would affect the system.
- Other movements in the room and their effects (e.g. non-participant occupants, pets, their shadows).
- Changing light conditions (e.g. window light, lights turned on / off, shadow conditions).
- The initial start-up time of the system will take a considerable amount of time for the human eye (up to a minute) for initial calculation of setting points

2. System Architecture – System Context Diagram

a. Architectural Description and Design: Roles, Activities and Data



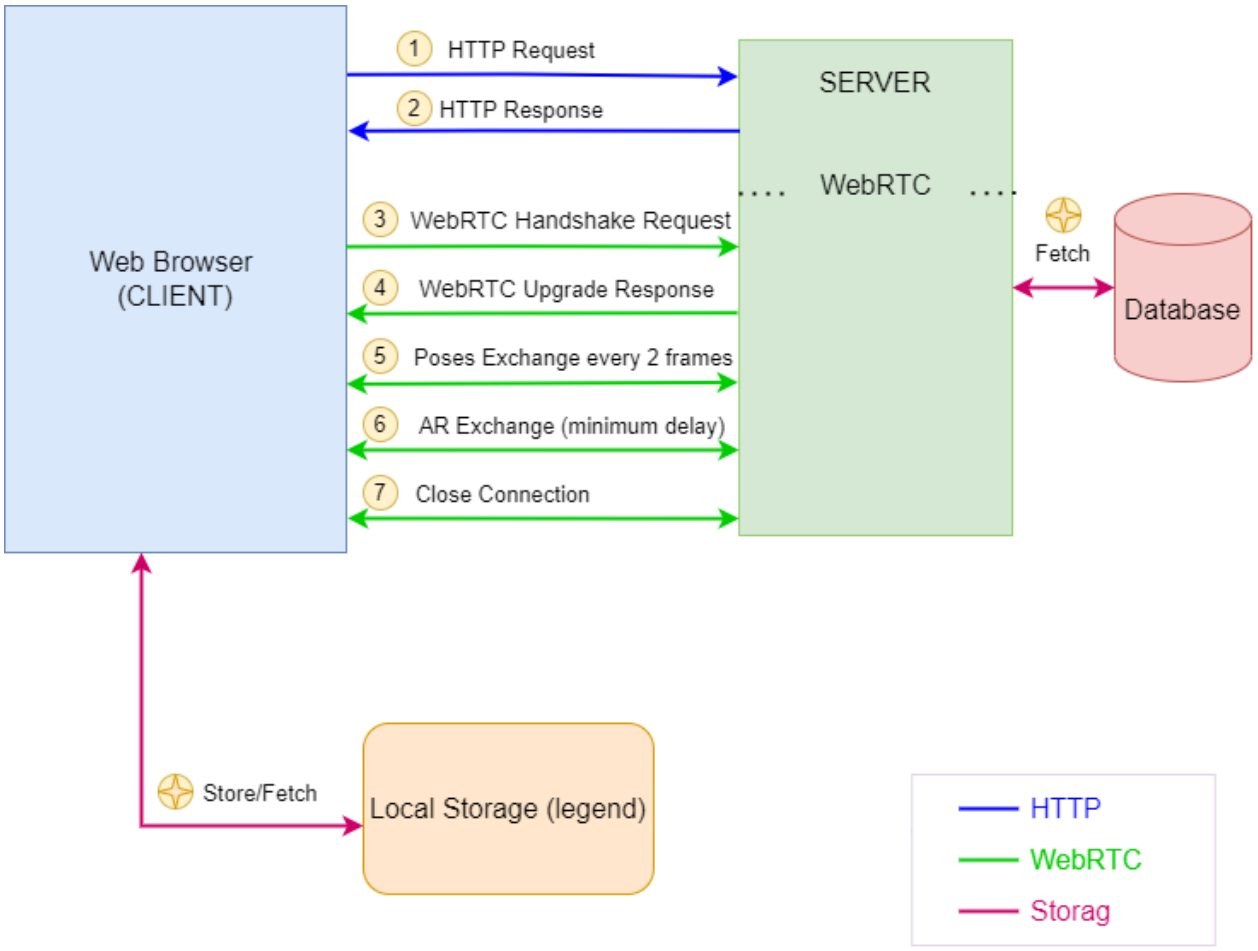
a.1. Communication on WebRTC ,Handshake req-res



The process for enabling communication on WebRTC can be seen as a 7 step flow:

1. Peer 1 registers with a signaling server.
2. Peer 2 registers with a signaling server.
3. Once both peers are registered with the signaling server, the server will send notification to both peers, indicating the other peer has arrived.
4. Both peers will connect to the TURN server and acquire the ICE candidates.
5. After successful acquisition of ICE candidates, both peers will send their ICE candidates to the Signaling server.
6. The Signaling server will send the ICE candidates to each peer for another peer (i.e, It will send Peer1's ICE candidates to Peer2 and vice versa).
7. By using other peers ICE candidates each peer will establish a successful Peer-to-Peer connection and exchange the Video/audio streams.

a.2. Real-Time Connection Architecture: High-Level Diagram using WebRTC



Steps 1 and 2: HTTP request-response for Login/Open/Entering meeting

Steps 3 and 4: Handshake request-response to switch protocols

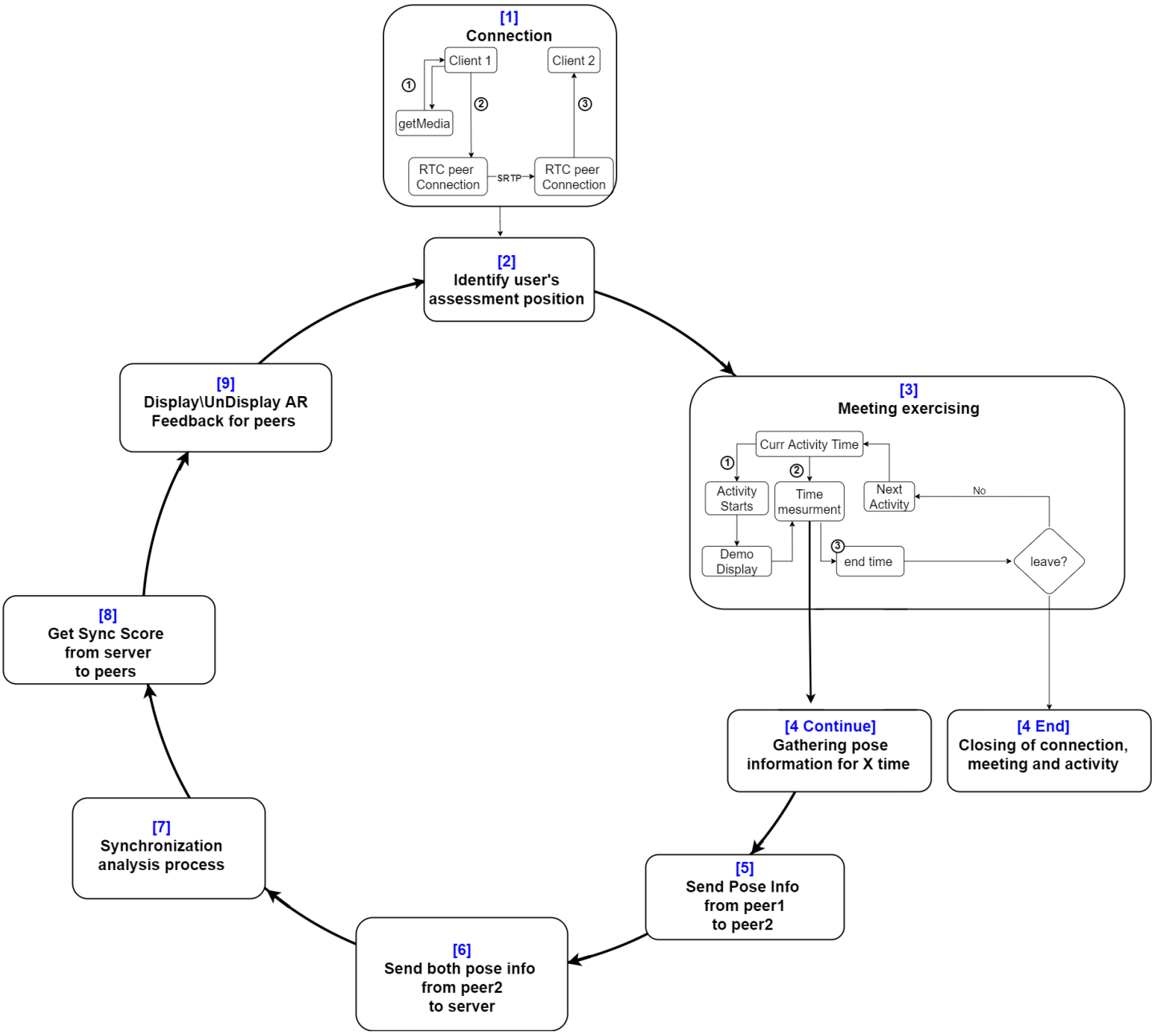
Step 5: Bi-directional Pose Estimation collected exchange.

Step 6: Bi-directional feedback exchange.

Step 7: Web Socket connection closes from either client or server

Step ✦: Verifying User from DB and set token authorization to LocalStorage

b. The Life Cycle of the System



:

3. Literature Survey

3.1)Literature Survey as a system

a. Problem Survey

The need of a system that will encourage joint sports meetings and will know how to recognize a synchronized movement and provide feedback to participants in real time about their movement coming from the customer - Kedar Center. The customer defined both the need and The functionality is made from the needs of research on the subject of synchronization.

b. Solution Survey

Two Riders IMS project[1]

The project was motivated by the creators' interest in design for interpersonal synchronization, and by a commission for an exhibition about the history, science, and design of bicycles. The creators conceived a novel experience using a physical object designed to enable, motivate, and reflect interpersonal synchronization. The project was installed for a total of two years in three science museums in three countries: Israel, Italy and Germany. In the Two Riders installation, two people sit face-to-face on a pair of bicycles that have a shared front wheel. Each participant can pedal independently. Their pedaling generates music that is dependent on their synchronization: the better they synchronize their cycling, the richer the music becomes.

"Flying with the Snark"[1]section:6.2

The Snark adventure game is a mixed reality game, where children hunt an elusive, virtual creature called the Snark, in a large interactive environment. It was designed by a team of researchers and designers headed by HCI researcher Yvonne Rogers. In this experience pairs of children were given jackets embedded with sensors to wear, and interacted with an abstract representation of the Snark, presented on a screen. The relationship between the children's actions and the Snark's "behavior" is described as such: If both jackets are banking the same way, then the Snark comes close and laughs. If both are flapping together then the Snark comes close and is very happy, soaring and swinging and gliding

c. Discussion and Conclusion

Interpersonal motor synchronization is a growing research area that presents opportunities. approaches to the design and implementation case of remote are bounded by the defining aspect of synchronization – temporal coordination. The need for precise temporal alignment, coupled with limitations of the online communication channels , presents unique challenges and opportunities; major aspects are: physical movement and the exchange of information with a special emphasis on remote interaction. Here we summarize the main need for SyncAlong system

d. Literature Survey Sources:

[1] Rinott, M., & Tractinsky, N. (2021). Designing for interpersonal motor synchronization. *Human–Computer Interaction*, 1-48..c

3.2)Literature Survey for algorithms for identifying synchronization motion between people.

a. Problem Survey

Synchronized joint activities have two key elements of synchronized movement: The similarities between the pose estimation of humans and the temporary alignment of movements. Research of OpenCv library-based interpersonal synchronization systems offers many methods for identifying human posture assessment in different techniques and obtaining different data reviewed in this section. Studies show that the way information is collected for the purpose of calculating synchronization between people can be examined in various indices.

b. Solution Survey

Research on Pose estimation, tracking and comparison[1]section:IV

In terms of the similarity of human poses: The article addresses given images, and images can be different sizes, a person can appear in a different part of the image, and by manipulating the image - by changing the size and scale of a bounding box of the person's image for consistent size. Additional steps involving normalization of the coordinate result as normalized L2 Vector array, And comparison by calculating the similarity of cosines

- The problem with this approach to the solution is the approach of image manipulation, the SyncAlong system produces a connection via stream and does not produce continuous images, continuous images can be produced but processing time will be required for this step

In terms of aligning sequence of pose: As the poses can be viewed as multi-variate time series, a method called Dynamic Time Warping (DTW)

Match Pose - A System for Comparing Poses[2]:

instead of mapping the skeletons length, the need is to compare the angles between their joints.

c. Discussion and Conclusion

During the market review, we came to the conclusion that the process of similarity in motion is not sufficiently defined to create an efficient algorithmic process. We chose to search for a solution that uses existing mathematical formulas to compare points. We chose to examine a mathematical approach, that there are reports in research literature on its successes in the field of motion comparison detection: this approach will be combined with reference to angle comparison as done in similar systems

d. Literature Survey Sources:

[1]https://www.inf.uniri.hr/images/studiji/poslijediplomski/kvalifikacijski/Sajina_kvalifikacijski.pdf

[2]https://www.ijert.org/research/match-pose-a-system-for-comparing-poses-IJERTV8IS100253.pdf

4. Technological Survey

a. Problem Survey

Real-time peer-to-peer signaling and media exchange:

Both WebRTC and WebSocket are technologies for communication capabilities.

Client side (Different platforms for Web applications):

- The use of our system will be made by Kedar Center for sync-data tracking
- easily and accessible by using the Web app.
- Switching between protocols, http←→webSocket.
- Supports opencv based libraries for body, head and hand recognition.
- Supports AR.

Algorithms requirements:

Motion synchronization is defined by two key elements:

- Match in human evaluation pose : a major requirement of our system is the ability to make an adjustment An engine that will adjust different posture values between users (participants in an activity).
- Match of speed/rate of movement: it is necessary to know the speed of movement of each participant, For the purpose of a synchronization index in the joint movement.

Server side:

- Permanent availability is required for end users based on Web RESTful services.
- Receiving requests from several customers at the same time.
- Switching between protocols, http←→webSocket.
- Create and update information that is in the database.
- Large videos are required to be saved

Data base:

- Entities with a fixed structure with predetermined parameters.
- Queries which consist of cutting and joining several different entities in the system.

b. Solution Survey

Real-time peer-to-peer signaling and media exchange : source-link sec-link

	WebRTC	WebSocket
Pros	• In a selective forwarding topology, each participant in a session connects to a server that acts as a selective forwarding unit (SFU). Each participant uploads their encrypted video stream one time to the server. The server then forwards those streams to each of the other participants. This reduces latency and also permits things like transcoding, recording, and other server-side integrations such as SIP which would be much more difficult in a peer-to-peer connection. • WebRTC is more secure • No software installation required • High-quality data transmission using modern audio and video • Video or audio quality adjusts to the digital environment during the connection, suppressing echoes and noise • Reliability and safety of data - the server protects connections by encrypting them via TLS and SRTP protocols • An implementation of an unlimited control interface based on HTML5 and JavaScript is available	• WebSockets is compatible with almost all existing browsers • The need for a video conferencing server (which mixes video and sound) for audio and video conferences in a group, since the browser is unable to synchronize two or more incoming streams • Cross-communication (cross origin communication) - it creates certain security risks but is needed for extensive functionality; • Cross-platform compatibility (web, computer, mobile devices); continuous connection from the backend to frontend in a web application working with a server
Cons	• The need for a video conferencing server (which mixes video and sound) for audio and video conferences in a group, since the browser is unable to synchronize two or more incoming streams	• **Slow** interaction response. The result of sending data to the WebSocket becomes known after 75 seconds of the timeout. • Security issues in WebSockets: processing denial; possible private data leaks; client-server encryption, etc.

Client side : source-link

	Angular	React
Pros	• Provides a complete application development application in the MVC architecture. • Allows testing to be performed using one tool. • Allows for a uniform version update for the entire infrastructure	• UI software development infrastructure that uses additional libraries as needed for development Client side, which allows flexibility both in terms of UI and in terms of functionality. • Easy to understand and use libraries that allow faster development. • Uses virtual DOM trees that enable enhanced performance and Improved user experience.
Cons	• Large library that requires deep understanding and a lot of learning time for new developers. • Includes many features embedded in the heart of the infrastructure. On the one hand it is an advantage because it does not require the use of additional directories but on the other hand limits the development. • Uses real DOM trees which are updated from beginning to end in each Update which degrades the performance of the system.	• Requires several tools to perform tests. • Requires the use of external libraries for consuming various functions. • Updating the version is not a simple procedure because it is necessary to verify versions of Third-party libraries used in development

Algorithms :A Libraries for Machine Learning in JavaScript, For Pose,Hand and Head Estimation :

*More ML libraries	TensorFlow.js	MidiaPipe / MP
Pros	• It has better computational graph visualizations. • It has the advantage of seamless performance, quick updates and frequent new releases with new features. • It can be deployed on a gamut of hardware machines, starting from cellular devices to computers with complex setups. • Tensorflow is highly parallel and designed to use various backends software (GPU, ASIC), etc	• MediaPipe offers cross-platform, customizable ML solutions for live and streaming media. • Working on a Graph-based framework, MP is tremendously faster than TensorFlow, especially so on the browser. • The visualizer runs these models in browser as wasm, WASM is a compilation product that technology allows to run heavy software at a speed that is almost as fast as they run on a computer.
Cons	• No support for Windows • Missing Symbolic Loops	• It has limited support for hardware acceleration. • It does not have default access to the file system in the browser host environment
Runtime (FPS)	**Desktop:** 42 | 35 | 29 **MacBook Pro:**52 | 40 | 24 **iPhone :** 43 | 32 | 22 **Slower**	**Desktop:** 150 | 130 | 97 **MacBook Pro:**75 | 67 | 34 **iPhone :** 9 | 6 | N/A **Faster**

Server side: source-link

	Node.js	PHP
Pros	• Faster development, requiring fewer lines of code. • Same language to server and client side helps synchronize between them. • There are external modules that can be used according to the development needs. • Known as a fast engine in the market compared to the other options. • Open source with broad community support.	• Open source programming language is common, especially for server-side web application programming. • Can be run on a wide range of operating systems and servers. • Enables efficient work with tabular databases using the ORM system. • Enables development cost savings if development is done by non-software infrastructures cost.
Cons	• Creates a bottleneck with heavy processes. • Known as problematic in performing asynchronous operations.	• Is known to have multiple security vulnerabilities. • Some of the software infrastructures for working with this language have little support.

Data base: source-link sec-link

	SQL	NoSQL
Pros	• The information is stored in a readable and simple way in the tables. • Cuts and combinations of tables can be made. • Avoidance of duplications. • Information Integrity (ACID)	• Very flexible and unlimited data structure. • Has enormous storage capabilities • NoSQL offers speed at the expense of information integrity (ACID).
Cons	• Limit a size in a particular field so that data may be lost. • Has a limited structure.	• Less readable and easy to track. • Limited ability in comparisons, analyzes and combinations of database tables

c. Discussion and Conclusion

- Real-time peer-to-peer signaling and media exchange:
The SyncAlong system is required to provide a solution of continuous and fast connection between users, the use of WebRTC designed for high-performance media delivery meets the system requirements in the best way and therefore we will use WebRTC for connection.

- Client side :
Development through the React open source library will save the team time learning and developing in a new language. We will also have support from external libraries for accelerated component development user interface. Since this language is based on the JavaScript language we can create uniformity in the whole system on the client-server side

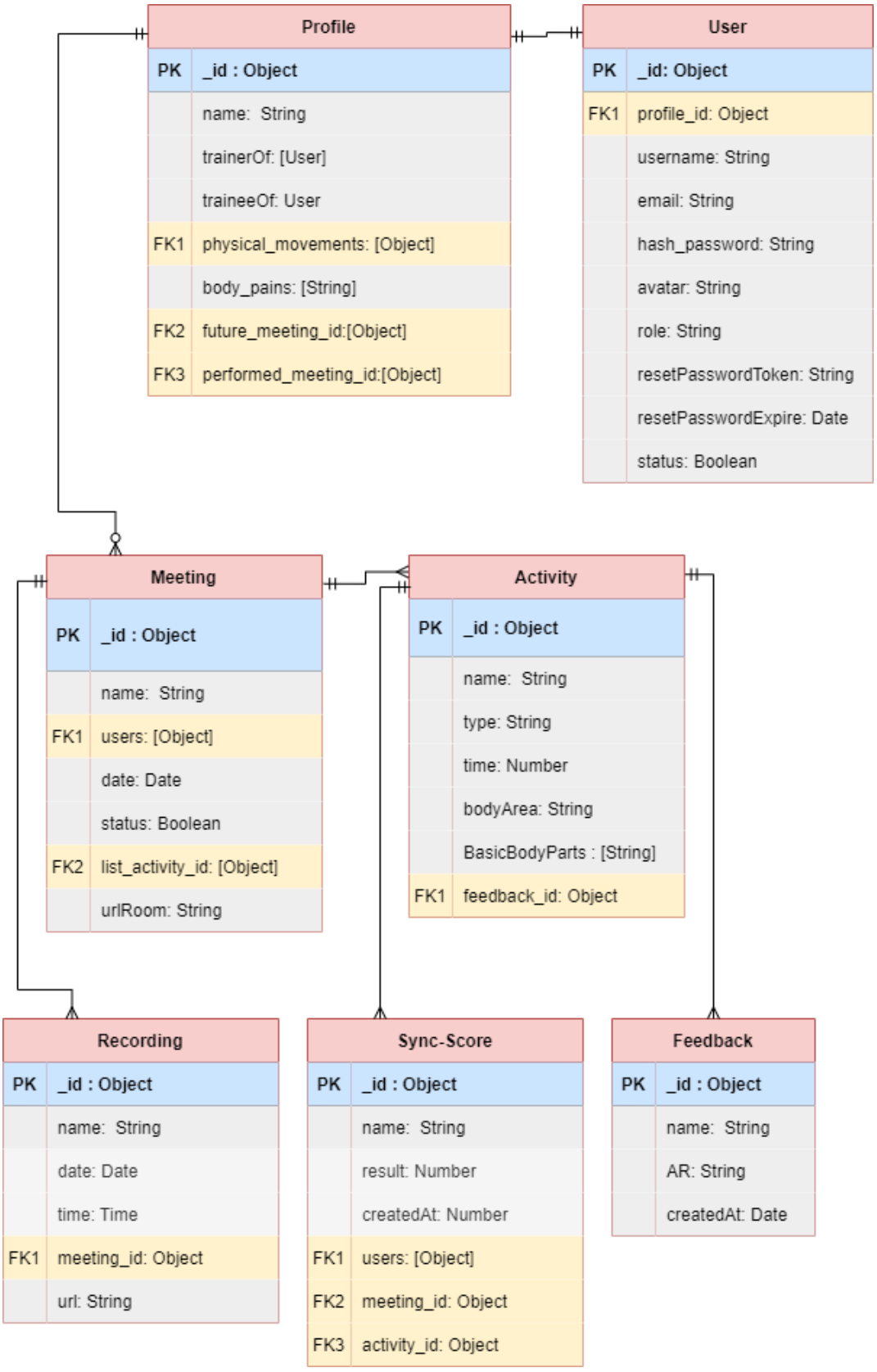
- Libraries for Machine Learning in JavaScript:
MediaPipe offers us a fast framework that although it has limited use at the moment it addresses the needs of the system. We can expect to identify human poses assessments with the help of mediaPipe library, and for the identification of hand marks, we'll use the `fingerpose package` to identify hand marks.

- Server side:
Because the client-side development language will be executed in JavaScript, maintaining consistency and usage in one language the development of the whole system (Full-Stack) will lead to a great advantage , which is why NodeJS is chosen for server development language.

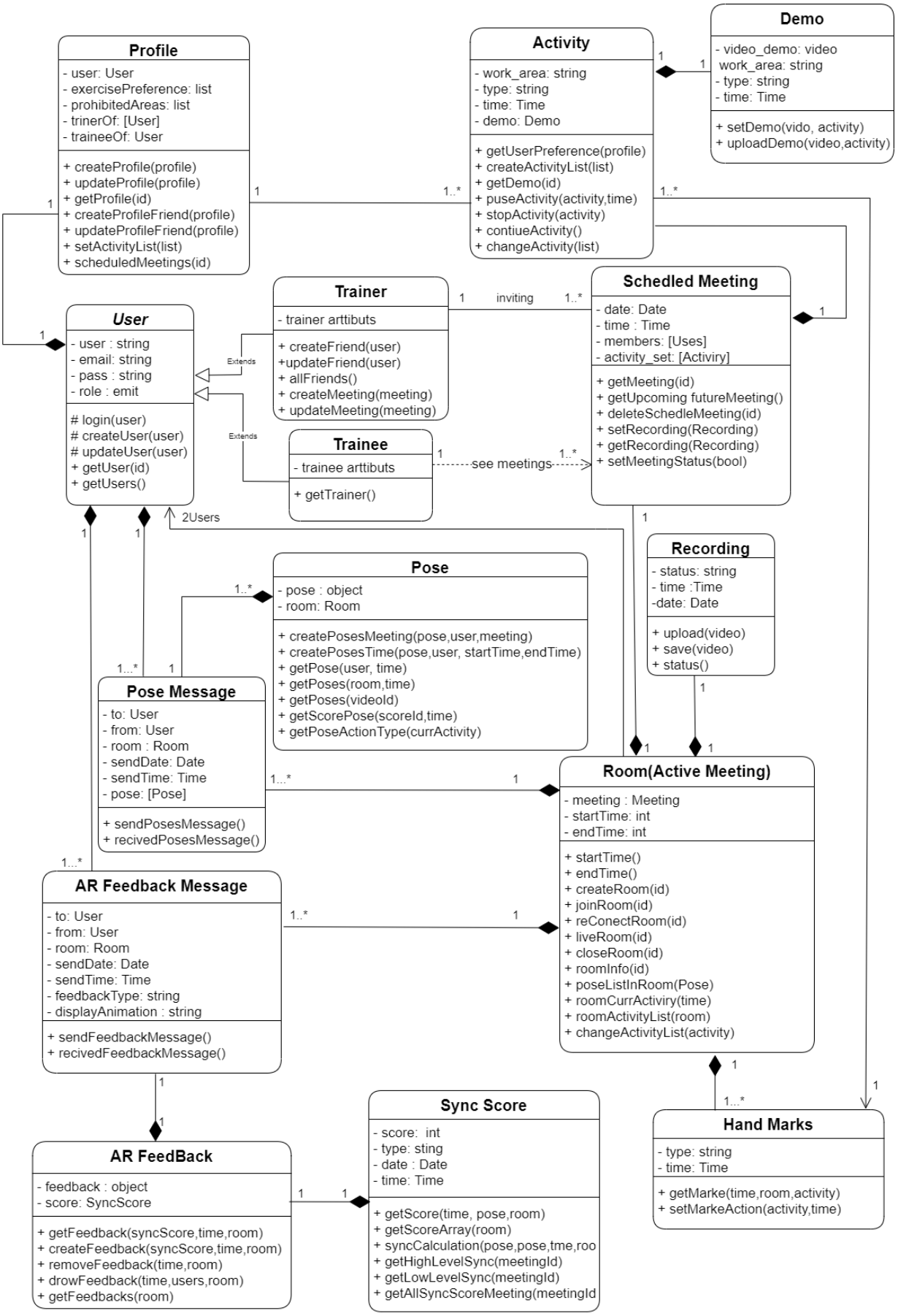
- Data base:
Because the SyncAlong system needs to store a lot of passing data such as data about sync results and about the feedback, new data enters every 2 seconds (=initial setting that can change). NoSQL is a database model that will give us a solution for a large volume of information organized by a particular method . Data integrity is not critical (to some extent) but speed is very important, NoSQL databases offer speed at the expense of information integrity (ACID). This is why a non-relational database such as MongoDb is the preferred option.

5. Design

a. **Data Design** - Database Description

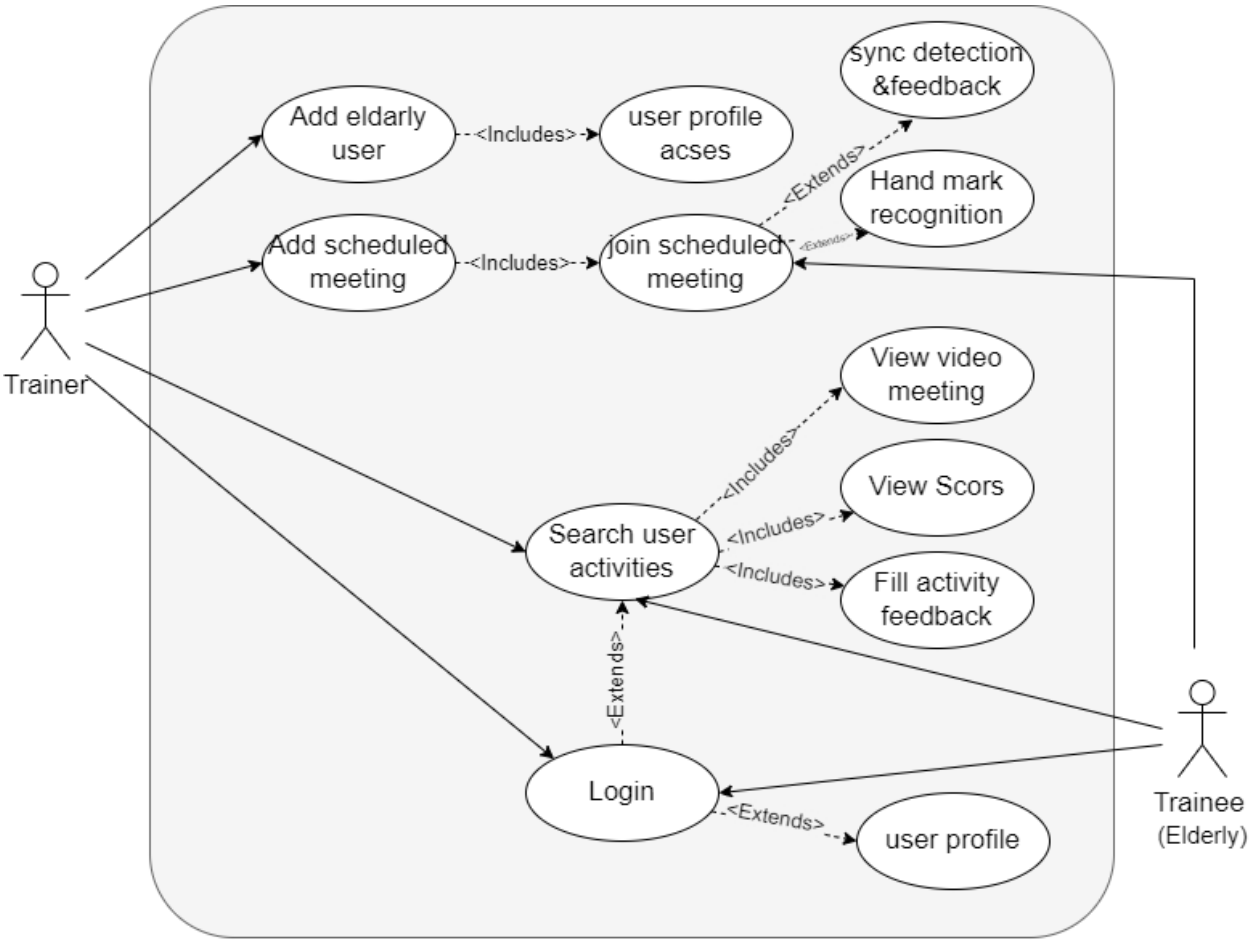


b. **Structural Design** - Class Diagram

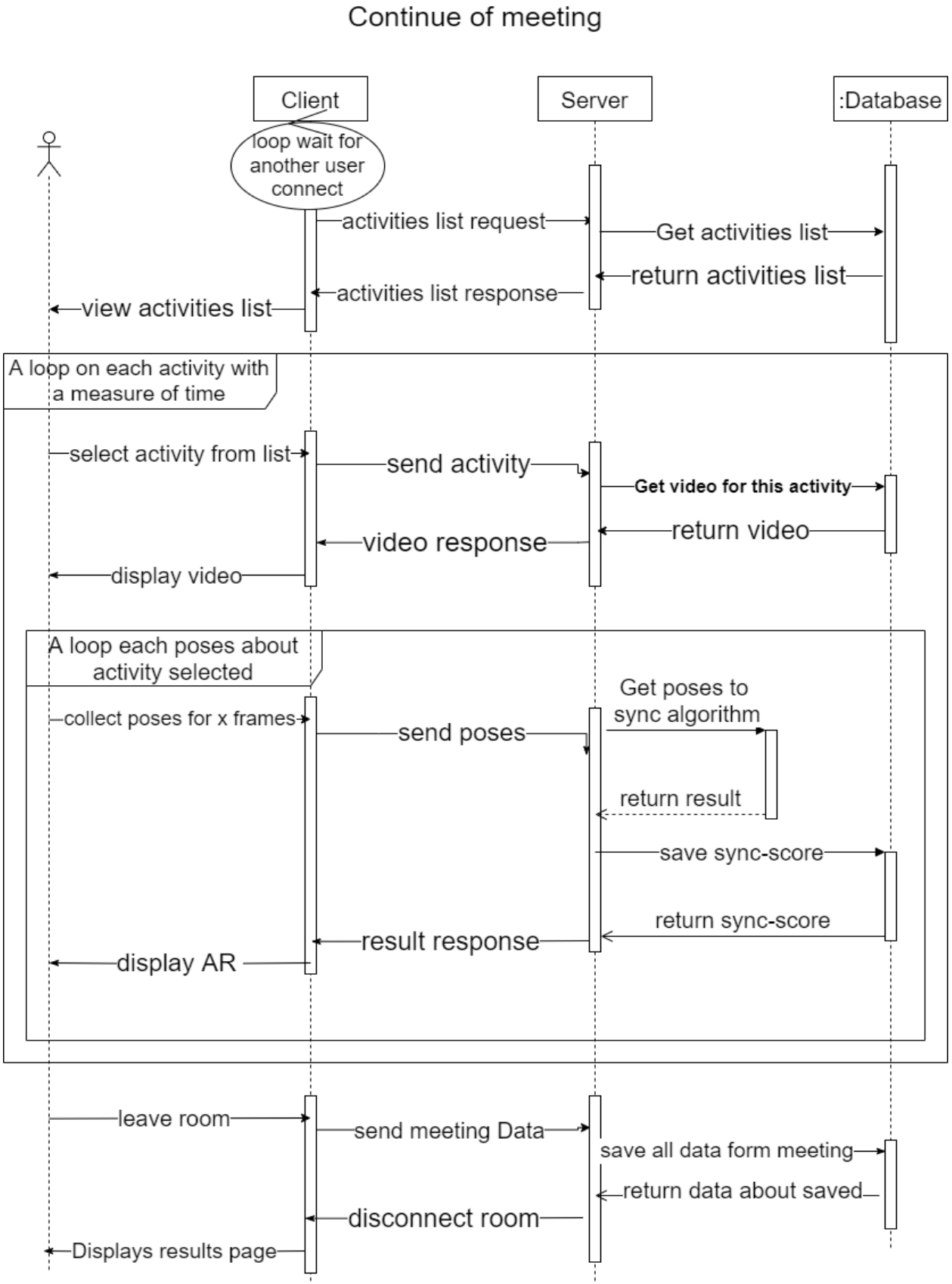


c. Interactions Design

i. Use Cases

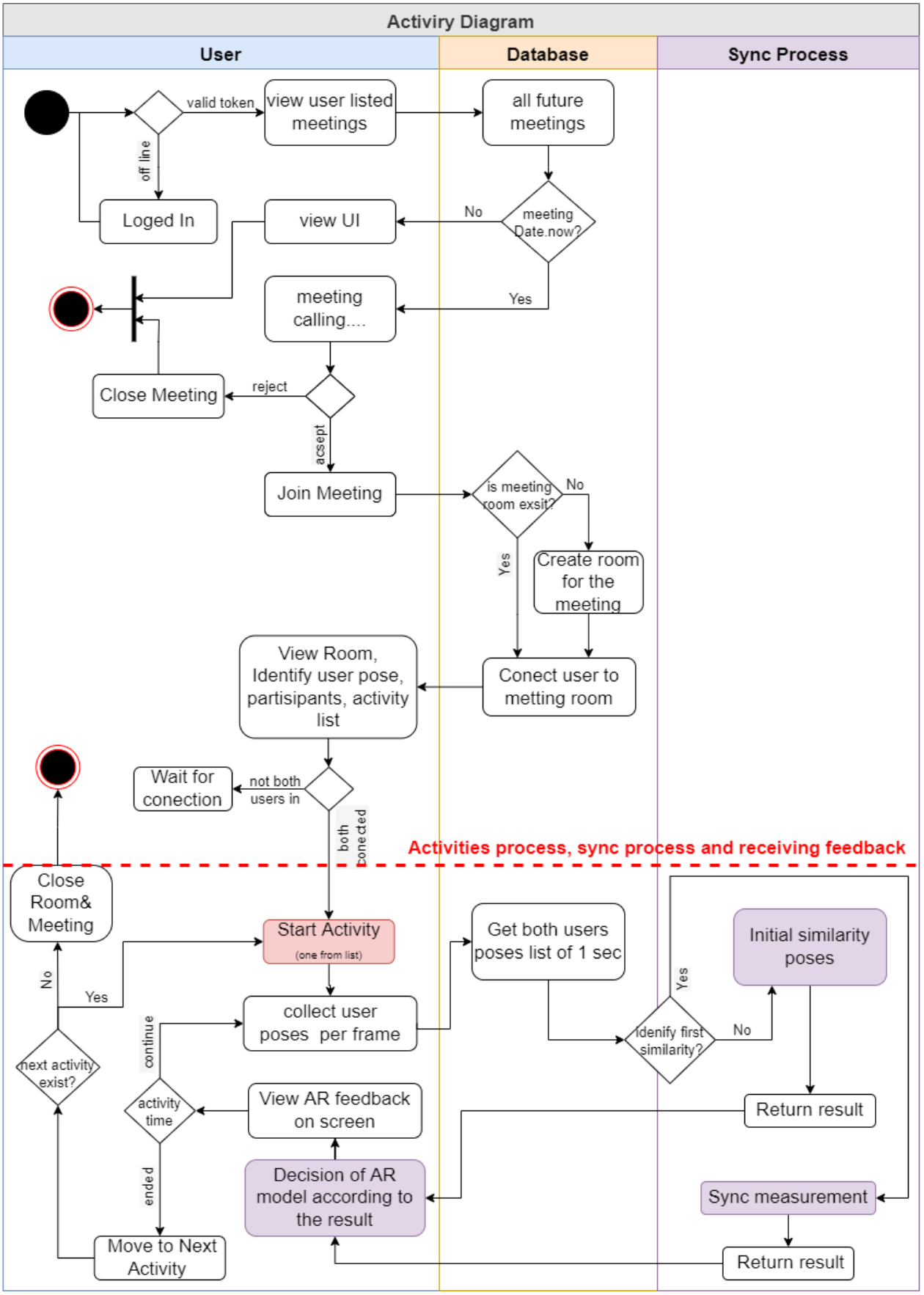


ii. Sequence Diagram

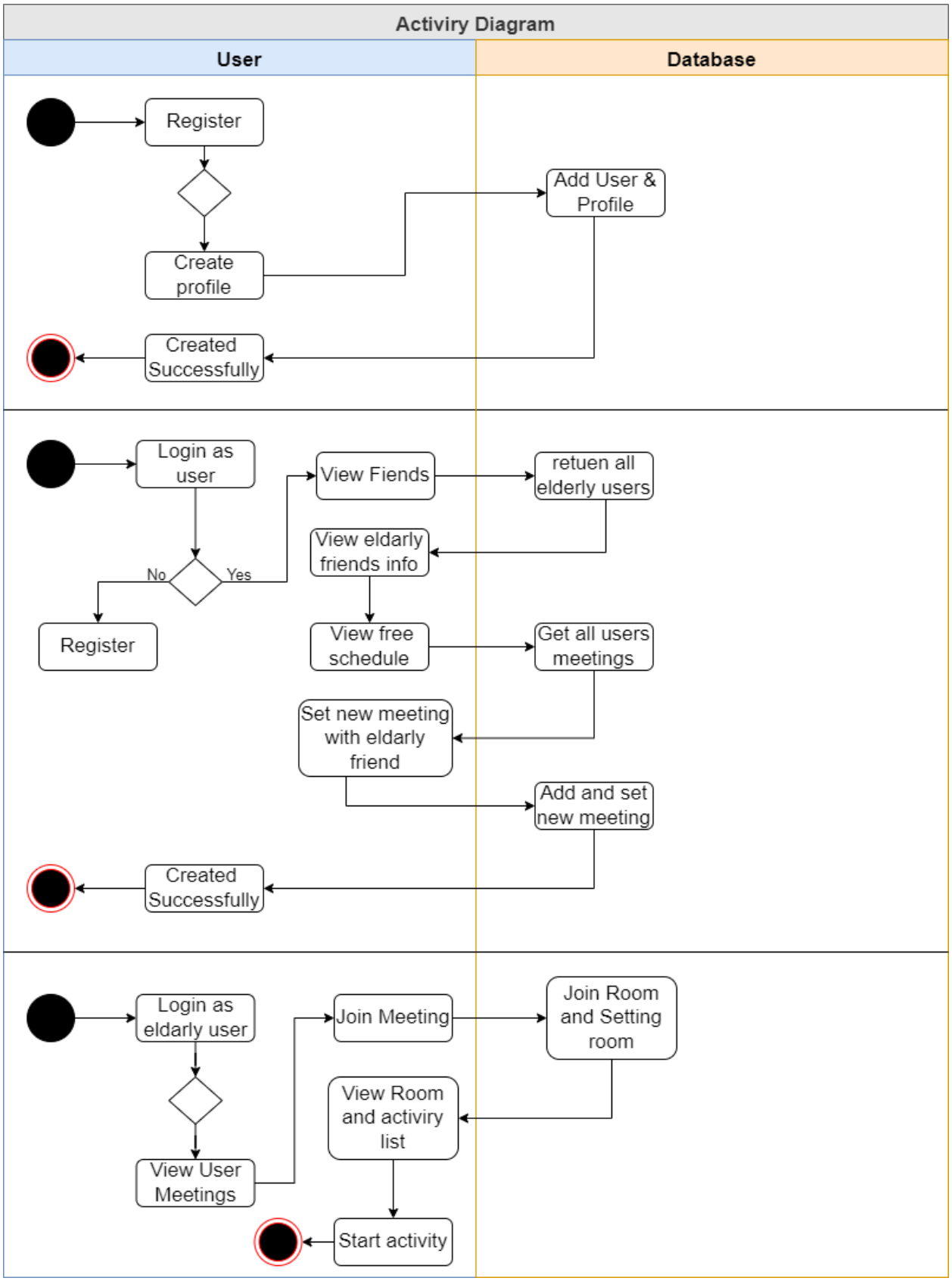


(1)

iii. Activity Diagram / State / Processes



[SyncAlong], [Shani, Ayman]



d. Description of Algorithmic Components

We are going to get two sets of this type of data: one for the live webcam feed of peer_1 that we need to match to, and one for live webcam feed of peer_2.

d.1. Sync algorithm overview

Important clarification: The process of calculating the similarity of humans will be in the aspect of reference to pose assessment only!!! (not with hands and face)
ref from MidiaPipe library:

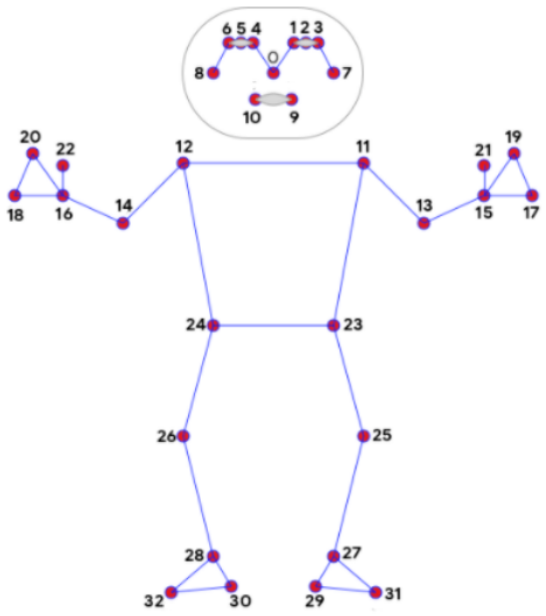
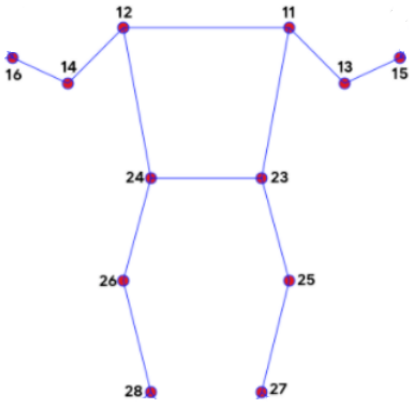
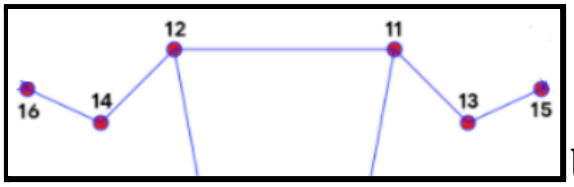


Fig 4. 33 pose landmarks.

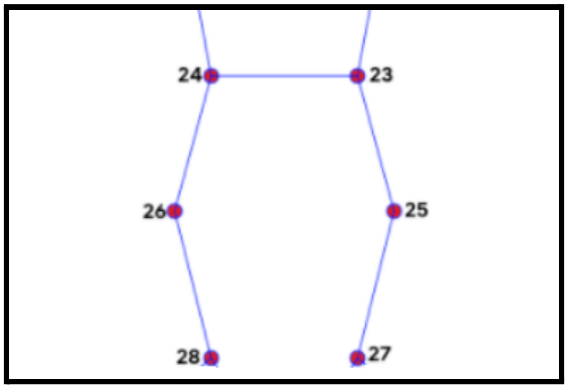
we will be forced on 12 points :



Divided into central areas for the purpose of planning coordination for the type of operation



Upper part



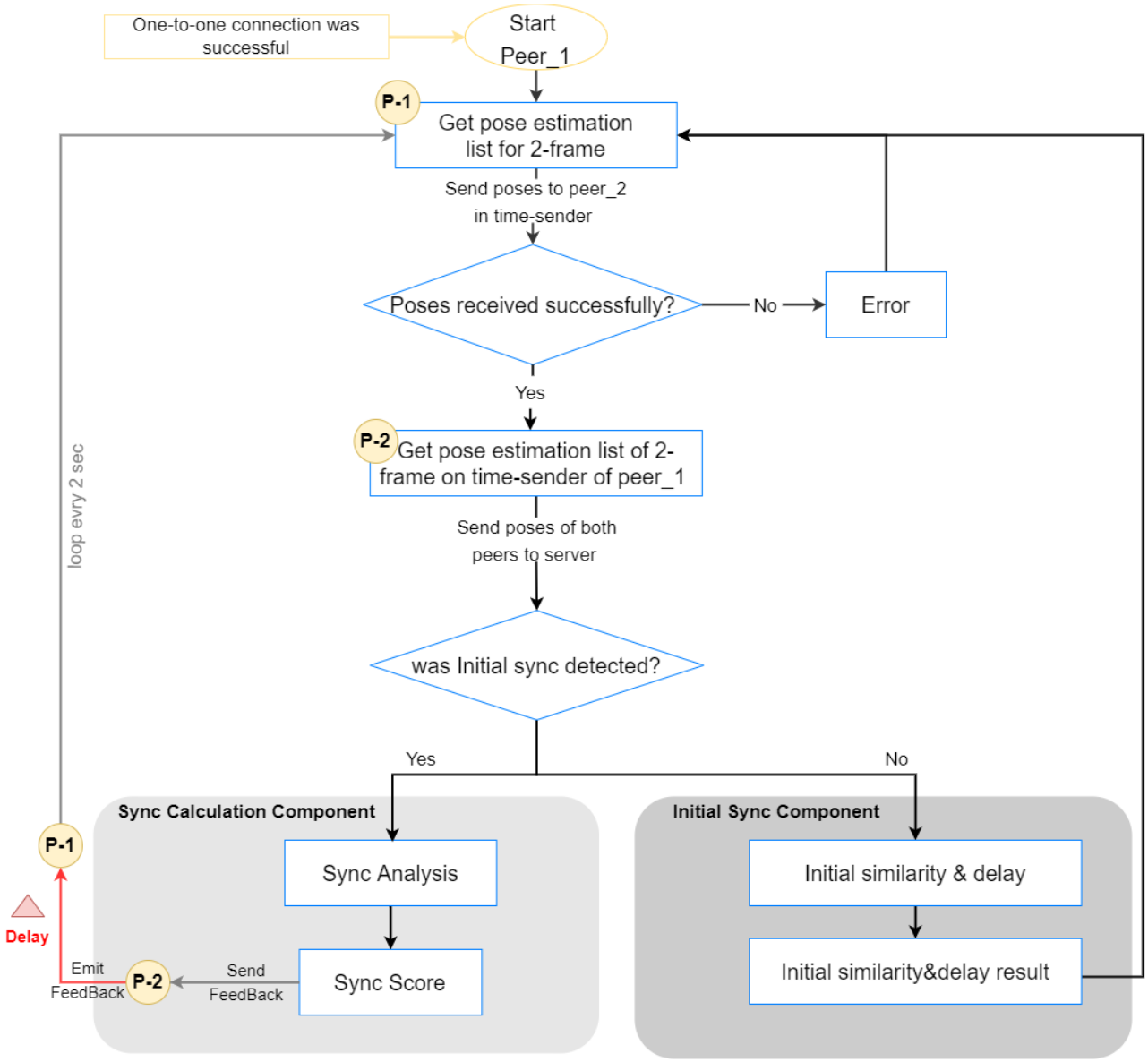
Bottom part

data receiving from pose for each frame:

```
{

  image: ImageBitmap,

  leftHandLandmarks: Array(21),

  faceLandmarks: Array(478),

  poseLandmarks: Array(33),

  ea: Array(33), …

}

poseLandmarks{

  0: {x: 0.5797398090362549, y: 0.7509459257125854, z:-2.0631303787231445, visibility:
0.9994146227836609}

  1: {x: 0.6200095415115356, y: 0.6761197447776794, z: -2.052156925201416, visibility:
0.999018669128418}

       … (and more points)...

 32: {x: 0.5109288692474365, y: 0.6718034744262695, z:-2.0453882217407227, visibility:
0.9992113709449768}

  length: 33 }
```

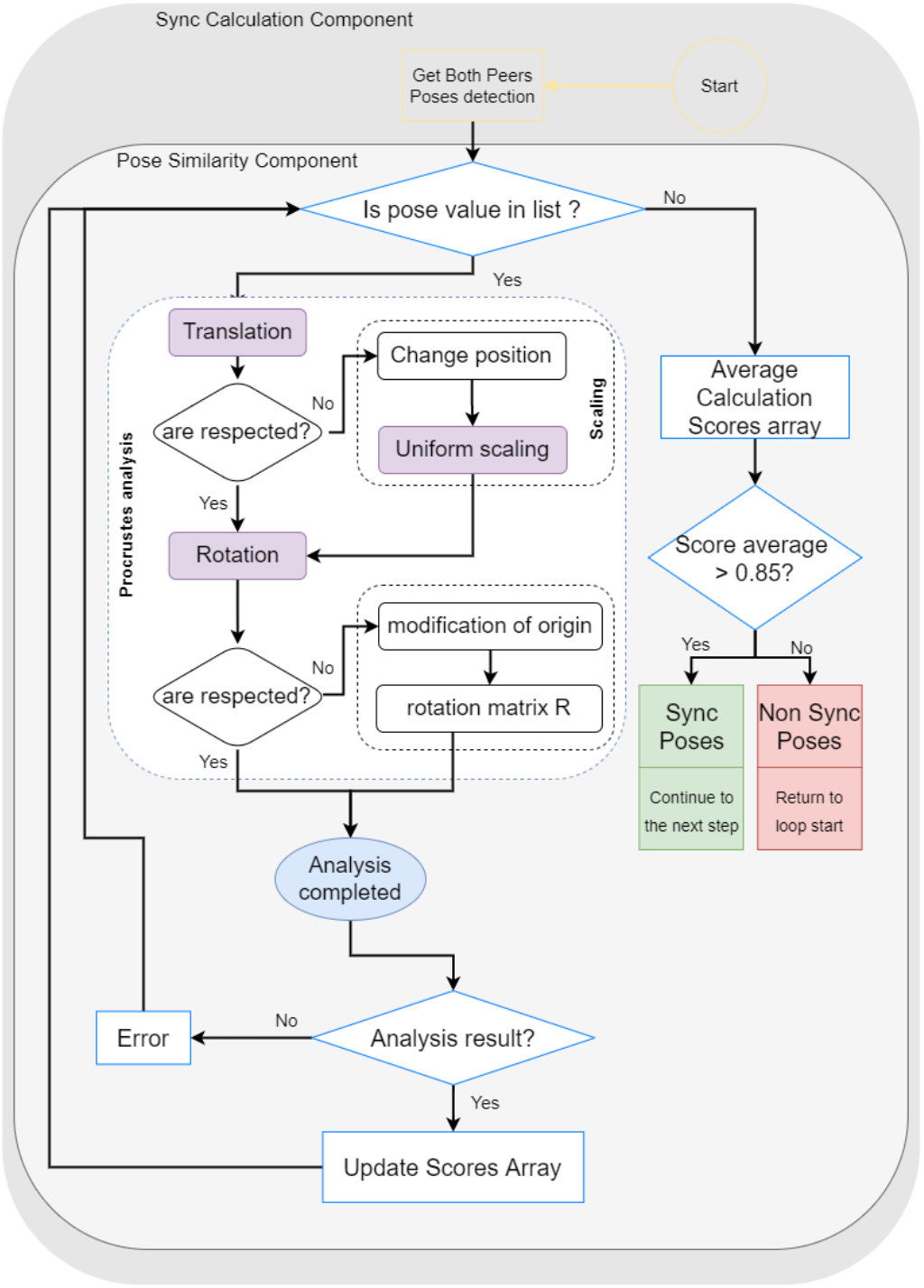
In order to shorten the running time on all points, the points will be filtered for the type of action performed (hand action will be considered an action on the upper torso) so that only points of the same body area will be sent to the sync algorithm.

d.1.1 Overview scheme:



d.2. Pose estimation similarity algorithm

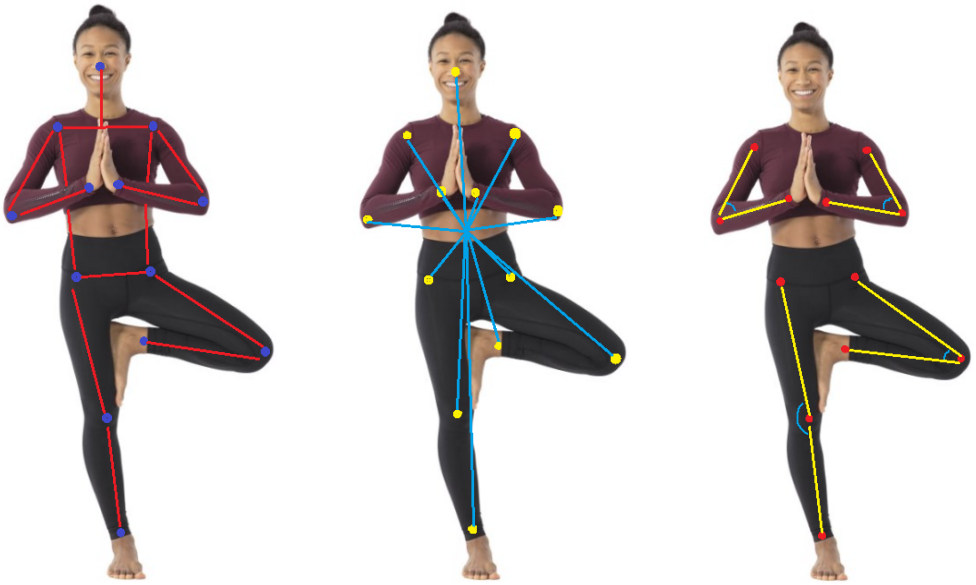
STEP 1: The first challenge is to define similarity. When we think about the problem,we see that there are many uncertainties to be addressed: humans can have different heights and body shapes, they might be in different parts in the picture: one person may have been standing close to the camera, another might have been faraway. All these problems have to be solved in order to output a correct result. we decided to take all the steps followed Procrustes analysis



The first step of Pose Similarity Component is based on mathematical steps that match an object to a model (a model is the first user in the connection - who created the room) After adjustment by translating, resizing and optimally rotating them.
The comparison is made using the square root and the calculation of distances between them.
Because the calculation of similarity matching is based on calculating distances between worn objects on top of each other we need to also address the matching of similarities in the main connection areas that produce an angle and check the similarity of the magnitude of the angles between the objects.
And that brings us to step 2.

d.3. Angles between joints similarity algorithm

STEP 2:



[1] **[2]** **[3]**

[1] Euclidean Distance, Scaling and Transformation
(Part of Stap1 of Pose Similarity component)

[2] Radial Vectors from the center of the torso

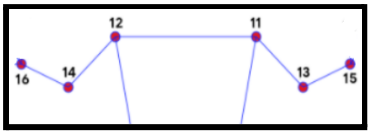
[3] Angles between joints of the body

Figures 2 and 3 above describe the steps required for examining similarities differences in angles.

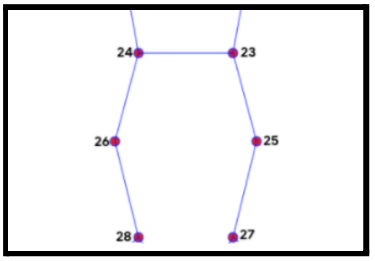
STEP2 - Algorithm for Angles between joints analysis:

1. Select each shape and base the keypoint data positions on a center, average position in the middle of the chest.
2. Transform the distances between each of the key points of :

- left hand keypoints 16,14,12
- right hand keypoints 11,13,15



- left leg keypoints 24,26,28
- right leg keypoints 23,25,27



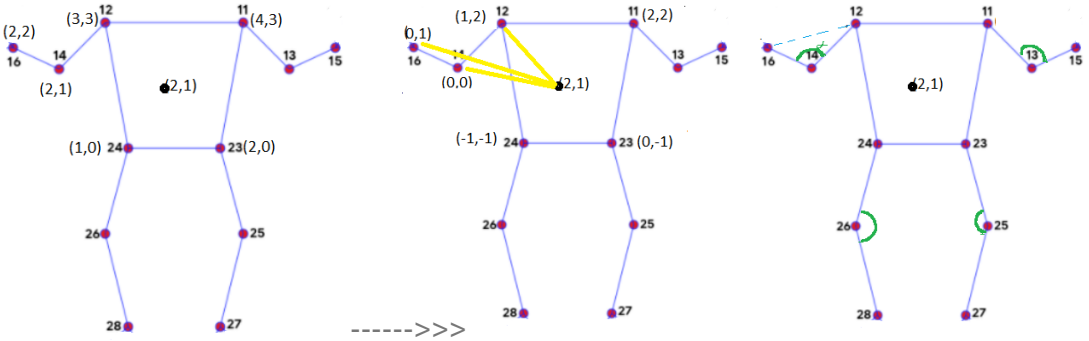
a. Calculate the centroid of each point in the landmarks array (set new landmarks points) in relation to the center of the body.
b. Define a new set of points in relation to the center of the body.

3. Calculation of angle in selected areas
4. Comparison of similarity and angle angles

—----------

---

example:



A. position landmarks of left hand : [{2,2},{2,1},{3,3}] | (collected from pose lib) | →left hand : [{2,2},{2,1},{3,3}] |

B. left hand relation to center: [{0,1},{0,0},{1,2}] | → points of 16,14,12

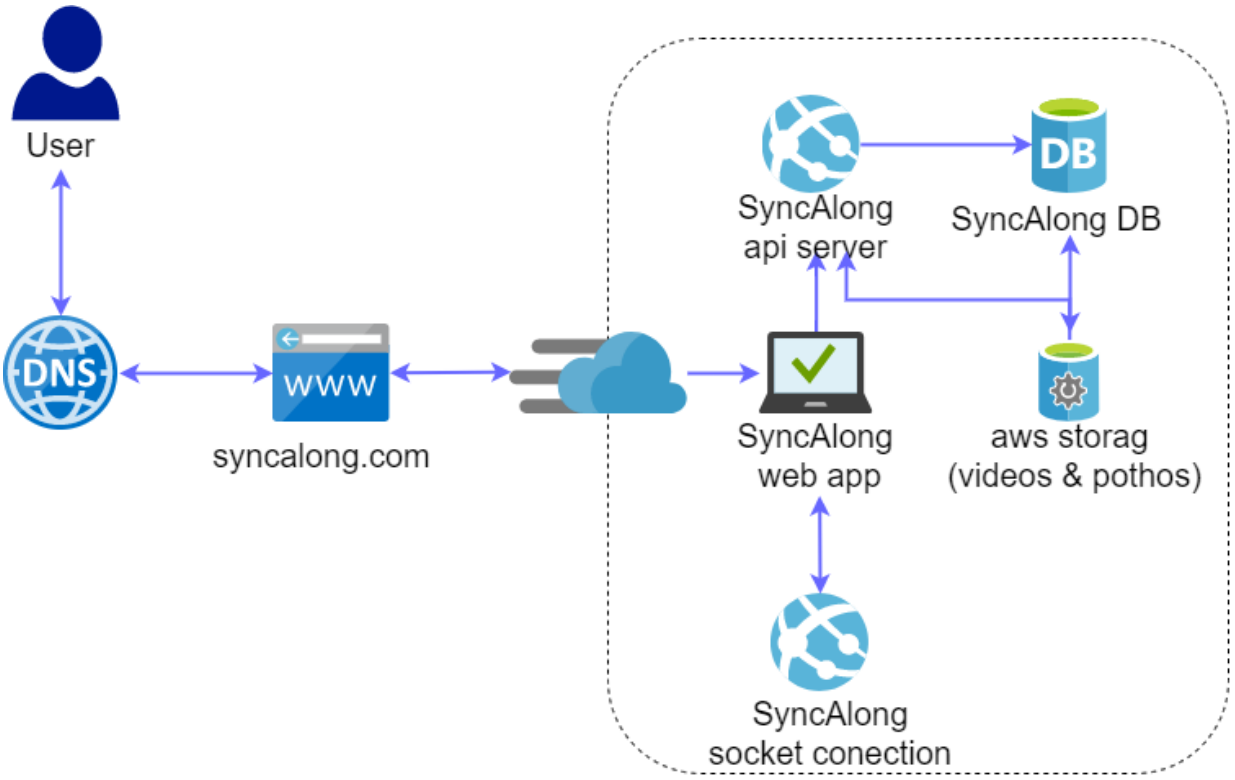
—----------

d.4. Speed similarity Element

According to the definition that the human position has 12 points that represent a coordinate value that changes over time, This feature allows us to treat this data as a time series,
Coordinates of human movement locations are collected in the POSES list.
The list defines the change in motion in a sequence (collection of a motion sequence of 2 seconds).
In other words, the calculation of the similarity of a person's assessment points (presented in step1 and step2 above) is also a test of a person's speed of movement in the same sequence being tested.

e. Software Architecture Pattern

i. N-tier: Data, Logic, Service, Presentation tiers etc.



6. Risk Management

- Sync algoritem, accuracy and reliability.
- Delay response in users connection.
- Delay of AR feedback.

7. Verification

a. Validation and Evaluation Plan

- **Delay :** Timing of server-client response.
- **Reliability :** we will have to test the reliability of the APIs requests and responses. Furthermore, we will test the ability of the API to deal with many requests simultaneously.
- **Performance monitoring** Key Performance Indicators (KPIs) are used to evaluate the performance of a website . It is critical that a webrtc web page must be lightweight to accommodate the signaling control stack javascript libs to be used for offer answer handling and communicating with the signaller on open sockets or long polling mechanism.

b. Testing Platform

Jest can be used to run tests in web applications. It is typically used with data-driven frameworks like React to ensure that components have the correct condition and accessories for a given input.

Lighthouse tab in chrome developer tools shows relevant areas of improvement on the web page from performance , Accessibility , Best Practices , Search Engine optimization and progressive Web App

8. Project Management

a. Source Control Platform for BU and Sync

For version management we will use git for local management and github for shared management

b. Alpha/Beta scopes

Alpha:

- Shani
  - server:
    - User management, model and functionality.
    - **Sync Algorithm and measuring delay response in real time connection.**
  - client:
    - Login interface.
    - User profile setting System function for adjusting activity list adapted to user data.
    - Connecting to a training room and disengagement treatment
    - Set up user posture evaluation.**Collecting pose array and sending for Sync Algorithm.**
    - Hand recognition.
- Ayman
  - server:
    - Establishment of a meeting room model and pose model.
    - Establishing user.
    - settings Connecting webSocket and handling real-time readings.
    - Authentication settings and basic protection.
  - client:
    - Setting cookies.
    - Treatment of access problems.
    - **AR settings in chat connection.**

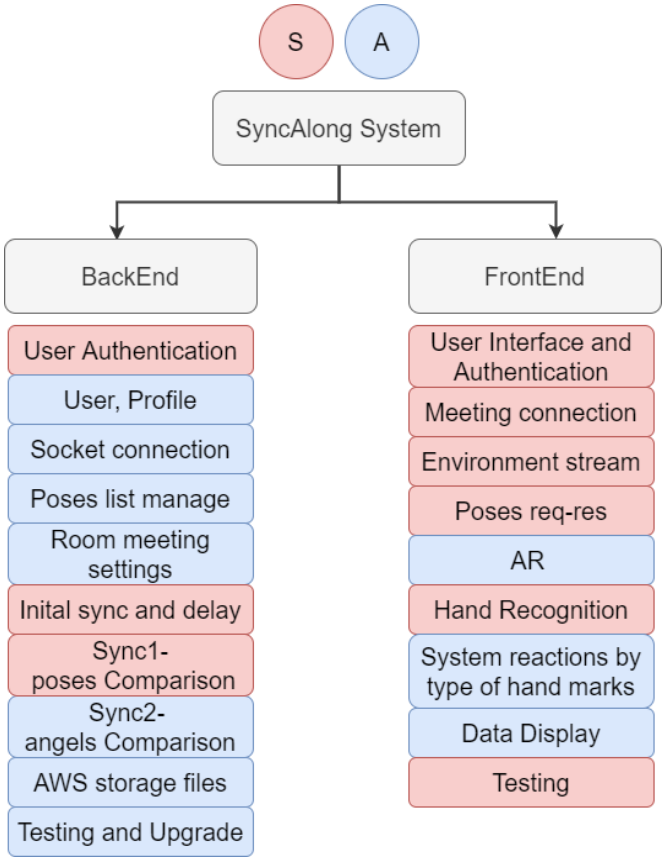
Beta:

- Shani:
  - server:
    - Pose model.
  - client:
    - Tests, improving reading times of the synchronization algorithm.
    - Main algorithm accuracy.
    - Hand recognition system function and system responses to the identification set.

        - User summary display interface and content creation for user feedback.
        - A social network for connecting users to a training room.
- Ayman:
    - server:
        - Handling data retention and reading using the AWS system
        - User Meeting Tracking Model and Model Management records meeting synchronization results.
        - Improving and focusing on clean and efficient code.
    - client:
        - Tests, improving reading times of the synchronization algorithm.
        - AR combination depending on synchronization result.

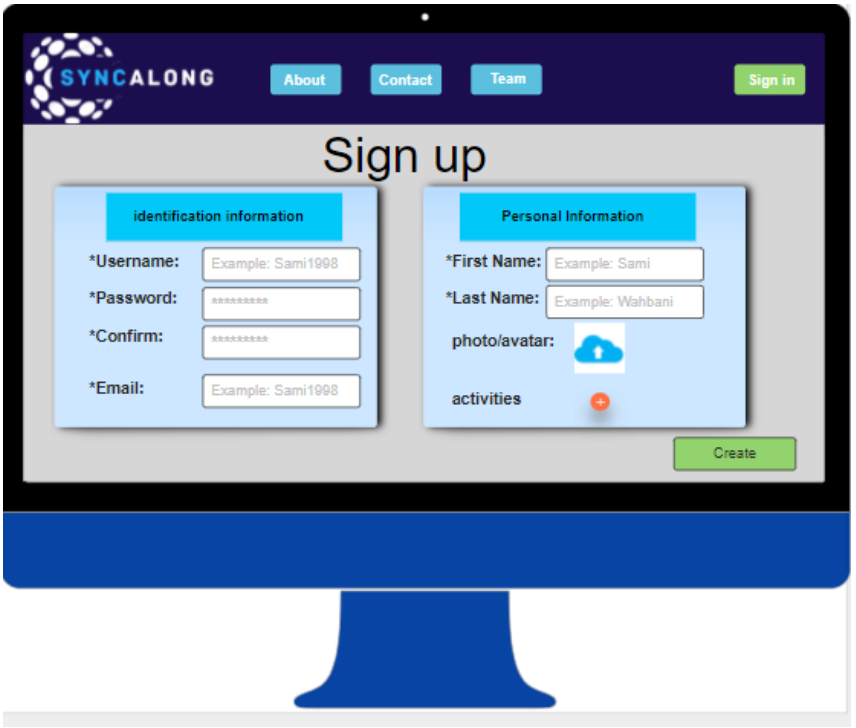
c. Schedule / Gantt
[SyncAlong system schedule link]

d. Team Roles - final

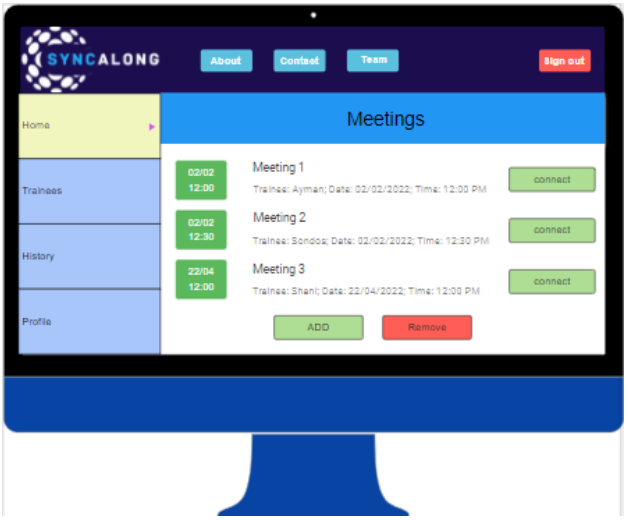


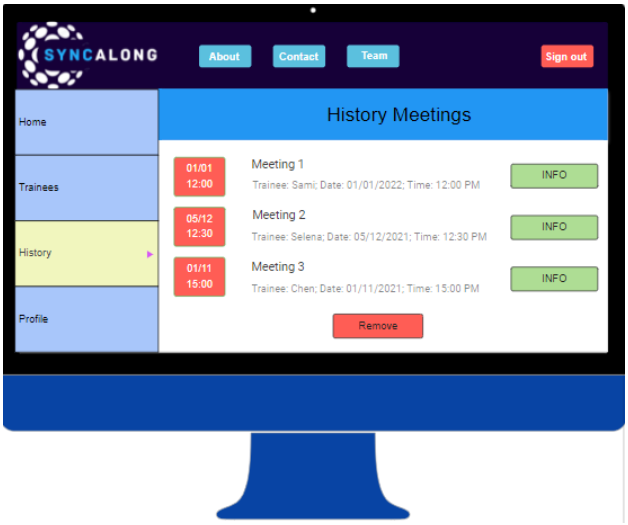
9. System Functional Screens

a. Home and Sign-in pages SyncAlong Website



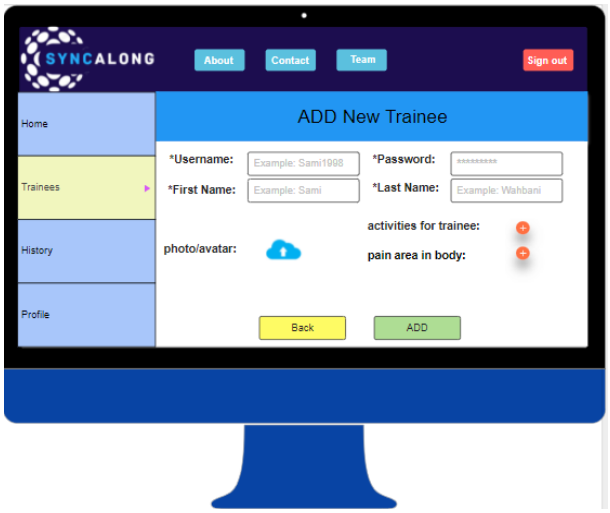
b. Trainer pages

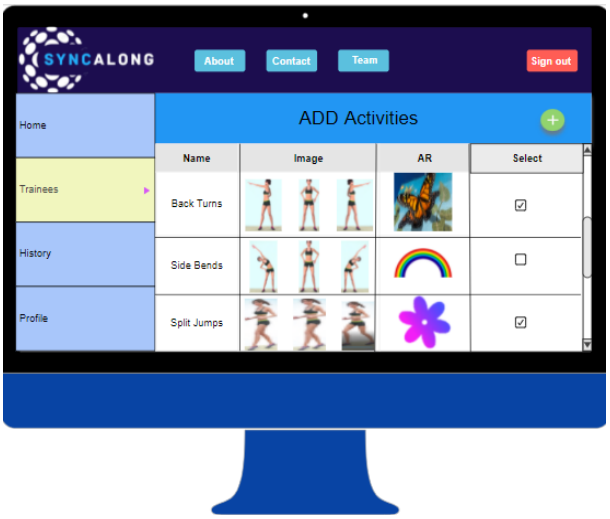




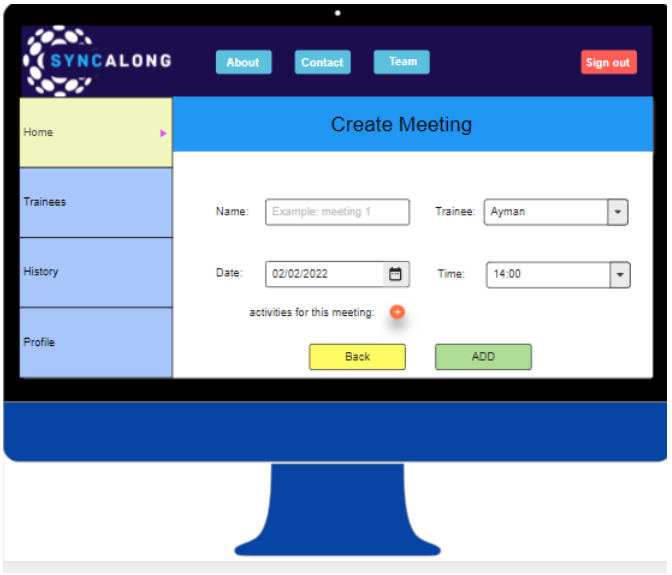
c. Show Activities for Trainer

d. ADD Trainee

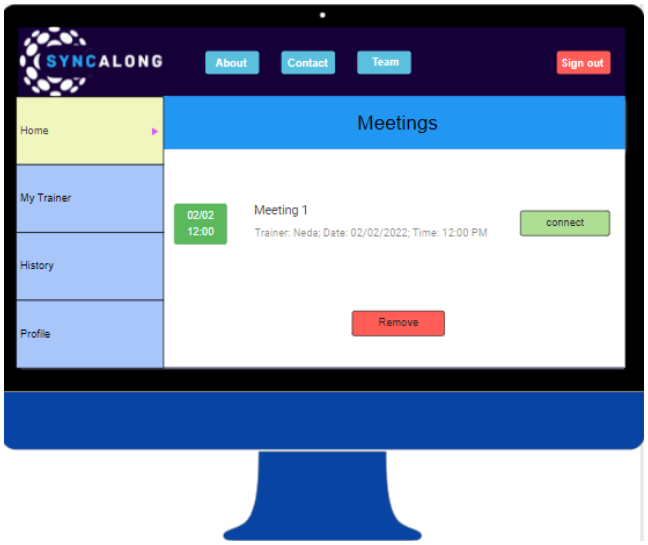


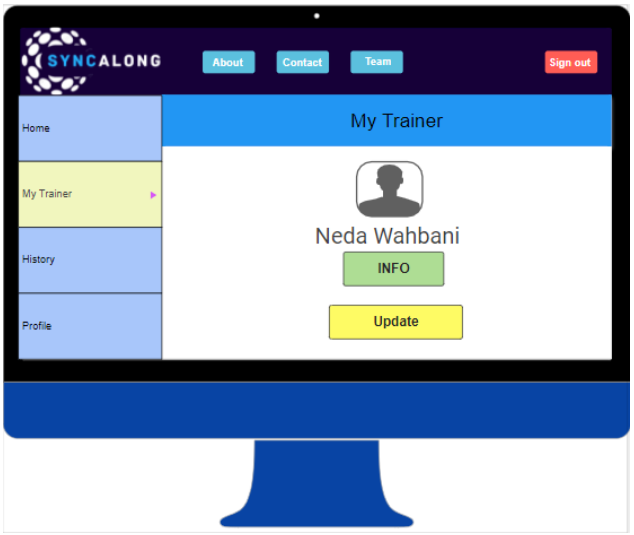


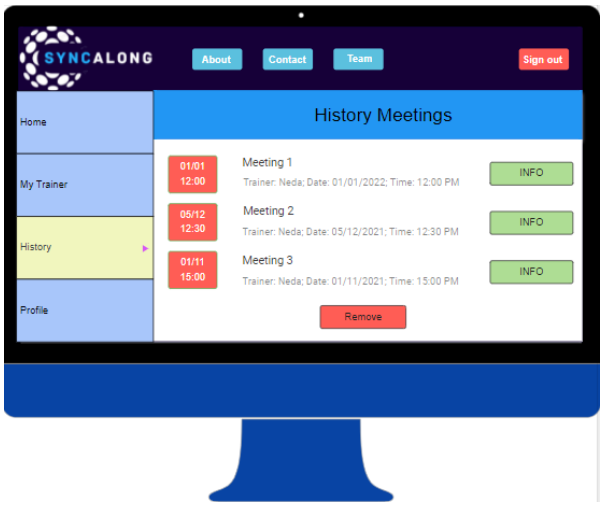
e. ADD Meeting and Activities for meeting



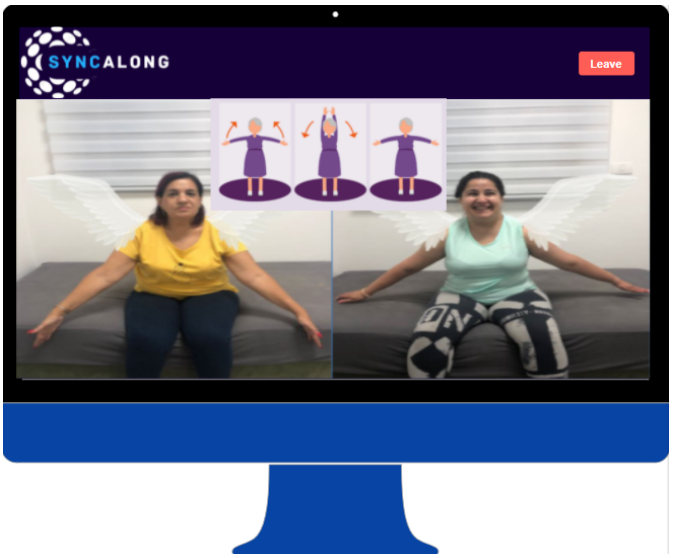
f. Trainee pages







g. Activity Screen (See activities then performed activities)



10. Appendix-A: Data Security

- User Authentication Process.
- Restriction of user privileges.
- User password encryption with Salt and Hash algorithm and writing passwords after encryption in the database.
- Secure data transfer between client side and system servers over HTTPS protocol Which encrypts the transmitted data.

11. Appendix-B: POC

a. POC discussion (if relevant)

The Challenge:

The SyncAlong system is a real-time system that requires it to respond within a pre-set time and consistently. Failure to meet the time conditions leads to incorrect results of system responses of feedback Synchronization calculation in real time.

Required:

It is required to ensure that the system responds within any time frame by receiving data, processing them and returning the results quickly enough to impact the environment at that time.