



Bilkent University

Department of Computer Engineering

Senior Design Project

Project short-name: YouTalkWeSign

High-Level Design Report

Abdurrezak Efe, Yasin Erdoğdu, Enes Kavak, Cihangir Mercan

Supervisor: Hamdi Dibekliolu

Jury Members: Varol Akman, Mustafa Özdal

Innovation Expert: Mustafa Sakalsız

High-Level Design Report

December 22, 2017

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the Senior Design Project course CS491/2.

Content

1. Introduction.....	3
1.1 Purpose of the system	3
1.2 Design Goals	3
1.3 Overview.....	4
2. Current software architecture.....	4
3. Proposed Software Architecture	5
3.1 Overview.....	5
3.2 Subsystem Decomposition	7
3.3 Hardware/software mapping	8
3.4 Persistent data management	8
3.5 Access control and security	8
3.6 Global software control.....	9
3.7 Boundary conditions	9
4. Subsystem Services	10
4.1 GUI Layer	10
4.2 Application Logic Layer.....	10
4.3 Storage Layer.....	11
4. References	11

1. Introduction

YouTalkWeSign will be a web application which is designed to serve people with hearing impairments or deafness. In today's world, there are many people who suffers from hearing loss and deafness. According to the last records of World Health Organization (WHO), there are 360 million people worldwide, corresponds over %5 of the world population who have disabling hearing loss. Moreover, they also claim that 1.1 billion young people, who have age between 12 and 35 years, are at the risk of hearing loss because of the exposure to the noise in recreational settings. The general reasons of the hearing loss are genetic causes, complications at birth, chronic ear infections, the use of certain drugs and growing old. People who have a significant hearing loss which also called deaf frequently use sign language to communicate with other people.

Sign language is a language which primarily uses manual communication to have a meaning, as opposed to spoken language. It is basically combination of simultaneous hand shapes, orientation, facial expressions and movement of the hands, arms or body to express ideas and communicate with other people. Sign Language helps to build bridge between people who can hear and who cannot. In addition to this, Sign language is not only used by deaf people, but it also can be used at some circumstances among people who have suffering from hearing and talking to each other such as very crowded or very quiet areas. Despite the significant role and 2 widespread usage of the sign language in human lives, it is not cared sufficiently and there are no enough resources.

With YouTalkWeSign, we aim to provide easily accessible source to create sign language of a video from YouTube. Thanks to the YouTalkWeSign, we are hoping that people who have disabling hearing loss can watch videos they like with understanding what kind of conversations and sounds are in the videos. It would also be a good resource for other people who do not have a hearing loss, but want to learn sign language.

1.1 Purpose of the system

YouTalkWeSign will be a web application which is designed to serve people with hearing impairments or deafness. Through an on-screen avatar, the app will translate the spoken words in YouTube videos into sign language. The app can be run by replacing the "youtube" word in the video address with "youtalkwesign". Therefore, our website is <http://www.youtalkwesign.com/>.

1.2 Design Goals

- YouTalkWeSign will have a good-looking and attractive user interface. It will be user friendly and easy to use.
- YouTalkWeSign will be suitable for extensions. For example, we can increase the number of languages supported. Also, the other platforms like Twitch and Dailymotion can also be integrated.
- A user focused system may be implemented to enhance the quality of the usage- via recommendations etc.

- We are aiming to make the conversion to the sign language as much as fast because we do not want our users to wait a lot. For live videos, they may wait longer than the normal videos. However, we will try to decrease the delay as much as possible.
- The system will be reliable such that we will retrieve the video from the YouTube without making any changes on the original version.
- YouTalkWeSign website will have a responsive design. That is, its front-end will be appropriate for a variety of devices and window or screen sizes.

1.3 Overview

For this project, we first need a sign language database that includes lots of words and corresponding videos. Then, we will do reverse engineering. For each word, we will track hand movements with the hand tracker API. From the hand tracker, we will get tracking data. Next, we will send these tracking data to the avatar modelling API. At the end, we will have an avatar that translates words at our cloud. There will be lots of words with the corresponding avatar video URL at the database table.

Next, for example, when user enters the URL:

<https://www.youtalkwesign.com/watch?v=2wseM6wWd74> we will need the transcript of the video. Like this YouTube video with the id: 2wseM6wWd74, for most of the videos, YouTube has a speech-to-text subtitles functionality. For these videos, we will simply get the speech text from the YouTube. For example: <http://video.google.com/timedtext?lang=en&v=2wseM6wWd74>. This link has a transcript for the video with the id: 2wseM6wWd74 in an XML format. We will take this XML and convert it to a Java object with the help of JAXB (Java Architecture for XML Binding).

After getting the transcript of the video, for each word we need an equivalent avatar movement. We will find it from the database. Then, our website will make word-to-word translation for the video. Since speaking is a lot faster than the sign language, we will adjust the avatar speed according to the speaking speed at the video. At the end, the video and the sign language avatar will be played simultaneously.

2. Current software architecture

There are many different applications about sign language at the internet. However, the idea of converting sound or text to the sign language is not that popular. We have found three important applications for this purpose in which two of them are Brazilian applications. Unfortunately, we could not have an information about their software architecture since they are not an open source projects.

ProDeaf

ProDeaf is a set of software that translates text and voice to Portuguese Libras - Brazilian Sign Language - to allow communication between the deaf and hearing. Their solutions are designed to allow businesses to promote accessibility and social inclusion to their customers and employees (<http://prodeaf.net/>) [1].

Hand Talk

They offer digital translation for Brazilian Sign Language. They use Hugo, avatar, their virtual interpreter, to make deaf communication more accessible (<https://handtalk.me/>) [2].

Mimix3D

Mimix3D Sign Language Translator is a mobile app that interprets spoken and written English into sign language using a 3D friendly avatar. New signs are added every week to their iOS application (<https://itunes.apple.com/tr/app/mimix3d-sign-languagetranslator/id1156035569?l=tr&mt=8>) [3].

After searching very deep through the internet, we have found these three applications. However, these three applications are not really related to our idea. They do not have anything about video to sign language translation. Therefore, since all three of them are similar in terms of functionality, there does not really exist a project about our idea. Also, the YouTube integration makes our application unique. On the other hand, projects that convert sign language to the text is available, but our idea is the exact opposite of this idea.

3. Proposed Software Architecture

- For front-end of the web application, we will use:
 - Thymeleaf (Java template engine)
 - Bootstrap v4 beta (for a responsive and a good-looking design)
 - jQuery v3.2.1 (for being a one-page app with ajax calls)
- For back-end of the web application, we will use Java technologies:
 - Spring Boot (auto configured version of Spring MVC)
 - Spring Security (JDBC authentication and authorization)
 - Spring Data JPA (database operations with MySQL [login, register, history, hearted and trending videos])
- For the implementation of the web application, we will use:
 - Eclipse IDE Oxygen Version
 - Apache Maven (for a dependency management).

Our server will be in a Digital Ocean droplet. Apache Tomcat 8 will be our web application manager where we will deploy our application.

3.1 Overview

Our application logic follows the idea of Spring MVC (Figure 1). Controllers catch the requests. Then, controllers use service classes to work for them. Service classes do their works on the database repository. For example, when user wants to heart a video, by clicking heart icon, he sends a POST request to HeartedController. HeartedController catches this and calls HeartedService to make it insert necessary data to the database. After successful insertion to the database, Controller warns view to do the necessary changes. For example, heart icon's color changes.

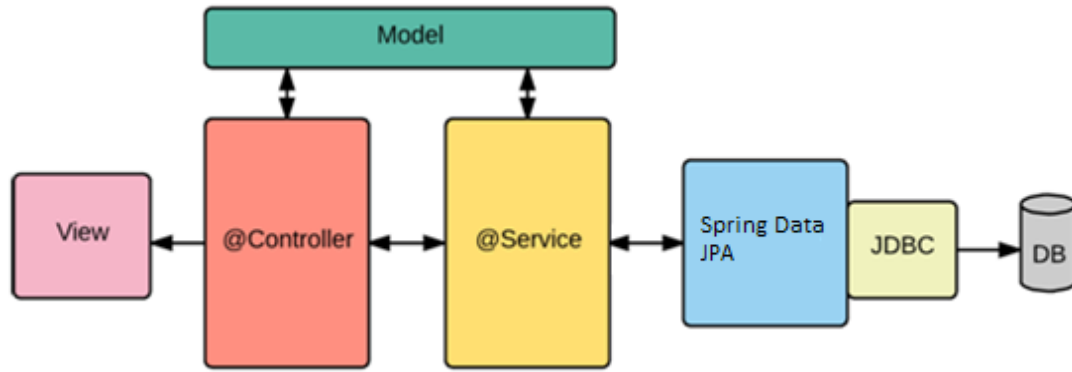


Figure 1 – Spring MVC

Our model classes are User and Video.

Our controller classes are MainController, LoginController, RegisterController, SearchController, HistoryController, TrendingController, HeartedController and AboutUsController where each of them has corresponding service classes.

3.2 Subsystem Decomposition

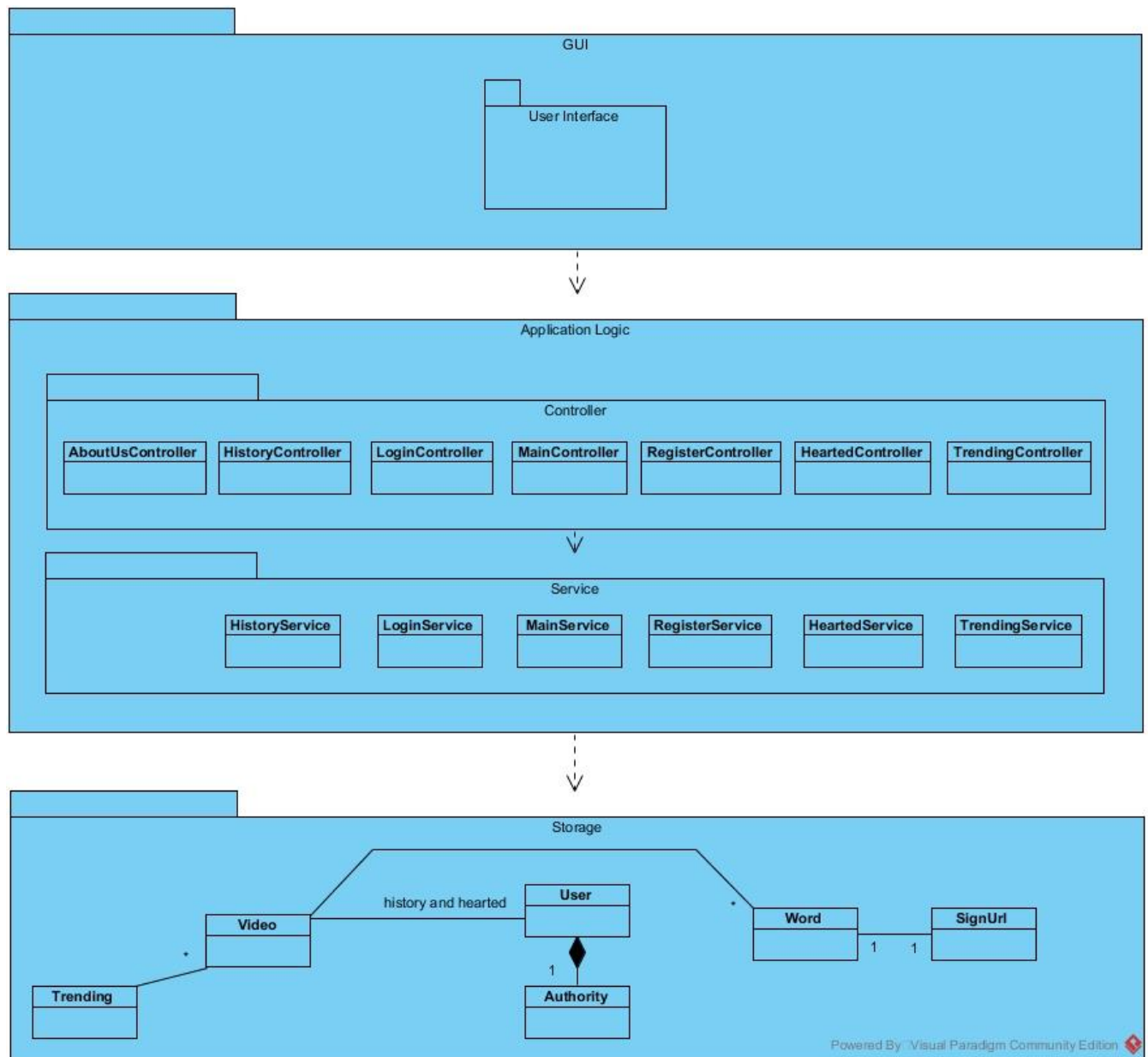


Figure 2 - Subsystem Decomposition

For YouTalkWeSign, we choose the three-tier architectural style which will organize our subsystems into three layers: GUI, Application Logic and Storage (Figure 2).

The **GUI** layer will have classes that are responsible for providing an interface between the user and the system: HTML/CSS pages that are powered by JavaScript and Thymeleaf.

The **Application Logic** layer will include all controllers and their services.

The **Storage** layer will take care of the storage of information in the database.

3.3 Hardware/software mapping

Our backend will be implemented in Java programming language using the latest JDK (1.9). Also, as mentioned before, we will benefit from Java EE frameworks. We will use MySQL database for storage.

3.4 Persistent data management

The information we keep into database are user credentials, trending videos, user history and hearted videos. This information updated when the user makes a request on the application side.

There will not be so many tables in database, therefore, every request can be handled. There will be some basic table such as users, trending, user history and user hearted videos tables. To keep data persistent, we will use a relational database. With relational database, we can handle relations between tables easily.

3.5 Access control and security

YouTalkWeSign's security will be provided with Spring Security framework. Spring Security will be used for password encryption, authentication and authorization. Also, it will be used for securing a URL access.

First, when the user successfully registers we do not save them to database as a plain text thanks to the spring security. We encrypt passwords before saving them to database because if the database ever gets hacked, all the plain text passwords will be exposed and that would be a great disaster. Hence, passwords must be encrypted with a good hashing algorithm which will be very hard for any hacker to crack. Spring Security supports one of the best hashing algorithms which is bcrypt [3]. Example (Figure 3):

<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>		username	email	password	enabled	
<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	cihangir	mercancihangir@gmail.com	\$2a\$10\$zosl.GMT2oq.qG3L.iy02.GJAXJOM.8.Mzd6nxdy6Xr...	1
<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	ds	ds@ds	\$2a\$10\$4CfOOy38vFERSLfjFjsVm.3XwtETULmfUfBSnbsv/l9...	1
<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	<div><div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div></div></div>	asd	asd@asd	\$2a\$10\$54aPrZV/yBanr49xC5us.RGFTcmXI6sbCj8D1u29Kv...	1

Figure 3 - Encryption

Second, as soon as the registration is done, new user is given a ROLE_USER. This will allow user to see pages which requires login. Example (Figure 4):

id	username	role
1	cihangir	ROLE_USER
2	ds	ROLE_USER
3	asd	ROLE_USER

Figure 4 - Authorization

Spring Security handles this authentication and authorization process via JDBC with the user and user_roles tables stored in MySQL database.

3.6 Global software control

Our system has an event-driven software control. When user request to login after entering his username and password, the system send request to database and check whether there is any user account in the database or not. If there is user with having these credentials, the system will display welcome page to user. When the user wants to change password, the system gets new password from user and saves it to database.

User can look trending videos by clicking trending button. When user request to see trending videos, the system will request to database to retrieve current trend videos list. This situation is same as for looking history and heart videos. However, in these situations, when the system gets request from user for seeing history and hearted videos, the system firstly check the user whether s/he has already logged in. If the user has already logged in, the system send request to database to retrieve user's own history or hearted videos.

When user wants to heart video, the system gets request from user and it firstly check whether s/he is authenticated user. If the user is not authenticated user, the system will redirect to user to login page. If the user is already authenticated user, the system will save video id with user id into the database. For authenticated user, after opening new video to watch, the system send request to database to save the opened video into database.

3.7 Boundary conditions

Initialization

The user should have internet access and any web browser to reach youtalkwesign.com. Due to fact that YouTalkWeSign is a web application, there is no need other than these two. To be able to access YouTalkWeSign web application, user can enter YouTube video URL by replacing YouTube with YouTalkWeSign or the user can directly enter youtalkwesign.com. After reaching website, application displays the video with avatar in its own media player. There is no need to have account for some basic operation like watching video with avatar and looking trend videos in the web application. However, to see history and hearted videos, the user needs to log in or register.

Termination

The user can logout of the website by clicking a Logout button, if s/he is authenticated user. Before logging off, the system saves recent user data like last viewed video into database. If the user clears the cache, s/he will be automatically logged out. When the user anonymous user, s/he does not have logging out option. All user can close the application by closing web segment.

Failure

If the user has failed to login, username and password text boxes will be displayed again. In addition to these, application will display link named as "Forgot password?". The user can create new password by clicking the link.

4. Subsystem Services

Here is the detailed explanation of our subsystem layers (Page 5 - Figure 2).

4.1 GUI Layer

The GUI layer consists of HTML/CSS pages that are powered by Javascript and Thymeleaf Java template engine. For the design of the web pages, Bootstrap v4 beta is used. Therefore, our website is full responsive and has a consistent design. In addition, for the purpose of being a dynamic webpage, we have benefited from jQuery a lot. Also, the icons that are used in our website is from the Font Awesome library. Below is an example view from our website (Figure 5):

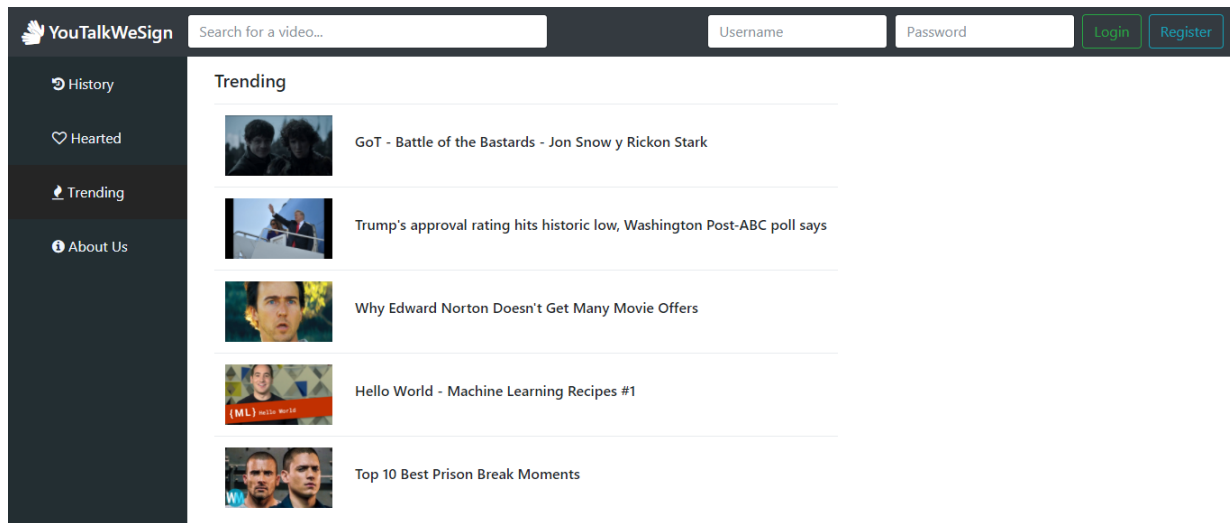


Figure 5 - GUI

4.2 Application Logic Layer

Our application logic layer uses Controller/Service procedure (Figure 6). In this procedure, first, request (GET or POST) from the user comes to the related controller and controller catches it. Then, controller calls its service class to get the job done.

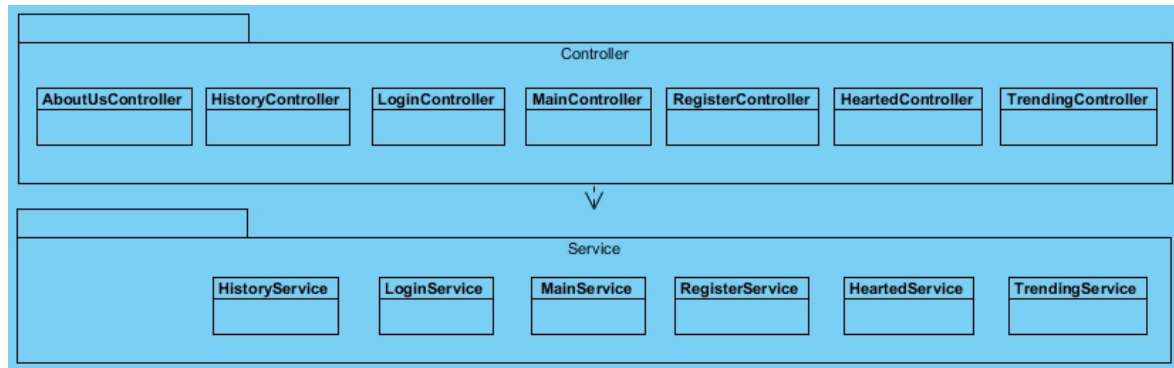


Figure 6 - Controller/Service

AboutUsController is responsible for showing our project's information to the user when clicked to 'About Us'.

HistoryController is responsible for keeping user's watched videos at the database.

LoginController is responsible for authentication.

RegisterController is responsible for authorization.

MainController is responsible for displaying the video with the corresponding sign language translation. It is the most important one because its job starts by getting the transcript of the video. Then, it prepares the words that are spoken in the video with the signs of them. Then, it administrates the GUI to make video and avatar play simultaneously.

HeartedController is responsible for keeping user's hearted videos at the database.

TrendingController decides the most watched videos at our website and keeps them at the database.

4.3 Storage Layer

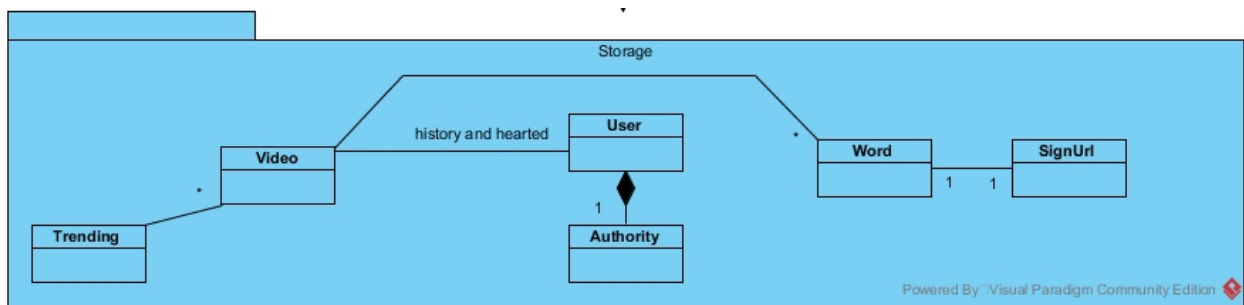


Figure 7 - Storage

First of all, we will have a word table which keeps words along with their sign URLs. Videos will have these words from their transcripts.

Second, Users' relation with the videos will be kept at database; their histories and hearteds.

Finally, we will have a trending table that keeps trend videos dynamically.

4. References

- [1] Pro Deaf. <http://prodeaf.net/> [Accessed: Oct 6, 2017].
- [2] Hand Talk. <https://handtalk.me/> [Accessed: Oct 6, 2017].
- [3] Mimix3D. <https://itunes.apple.com/tr/app/mimix3d-sign-language-translator/id1156035569?l=tr&mt=8> [Accessed: Oct 6, 2017].