

# WeTube Projektdokumentation

## Team:

Helena Aldaloul  
Elif Akpinar  
David Fliesser  
Manuel Seubert

## Inhalt

1. Interpretation der Aufgabe .....	3
2. Anforderungen/ Umsetzung.....	4
2.1. Gesture Mode .....	4
2.2. Cursor Mode.....	4
3. Gestaltungskonzept und Mock-up-Erstellung .....	5
3.1. Homepage .....	5
3.2. Cursor Mode.....	6
3.3. Gesture Mode .....	8
4. Full Screen .....	9
5. Finale Projekt.....	10
5.1. Homepage .....	10
5.1.1. Gesture Mode .....	10
5.1.2. Cursor Mode .....	10
5.1.3. Deaktivierte Kamera Mode.....	11
5.2. Videoseite.....	11
5.2.1. Gesture Mode .....	11
5.2.2. Cursor Mode .....	12
5.2.3. Deaktivierte Kamera Mode.....	12
5.2.4. Full Screen.....	13
5.3. Search-bar .....	13
6. Unterschiede zur Mock-Ups .....	14
6.1. Allgemeine Style.....	14
6.2. Verfügbare Gesten .....	14
6.3. Videoplayer Controls und Full Screen .....	16
6.4. Feedback .....	16
6.5. Aktivierung / Deaktivierung von Kamera und Gestures.....	17
6.6. Videoliste & Search Bar.....	17
7. Verwendetes API .....	18
8. Fazit .....	19

## 1. Interpretation der Aufgabe

In der heutigen Zeit nutzen viele Menschen während des Essens, Lehrens oder anderen Tätigkeiten soziale Medien, beispielsweise über Fernseher, Smartphones oder Tablets. Um die Nutzungsszenarien zu unterstützen, wurde ein Konzept entwickelt, das sich an bestehenden Videoplattformen orientiert und diese gezielt erweitert. Die entwickelte Plattform ist in ihrer Grundfunktion mit YouTube vergleichbar, bietet jedoch einen zusätzlichen Mehrwert.

Der zentrale Vorteil der Plattform liegt in der gestenbasierten Steuerung von Videos. Nutzerinnen und Nutzer können Inhalte ansehen und vollständig ohne physischen Kontakt bedienen, ausschließlich durch Handgesten. Dadurch wird eine komfortablere, hygienischere und intuitivere Interaktion ermöglicht, insbesondere in Situationen, in denen der Einsatz von Maus, Tastatur oder Touchscreen unpraktisch ist.

Die entwickelte Plattform trägt den Namen „WeTube“.

## 2. Anforderungen/ Umsetzung

Im Rahmen der Entwicklung der Anwendung wurden verschiedene Anforderungen analysiert und definiert, um festzulegen, welche Funktionen umgesetzt werden sollen. Dabei wurde bewusst auf die wesentlichen Kernfunktionen fokussiert und diese prototypisch realisiert.

Zu den zentralen Funktionen der Anwendung gehören unter anderem:

- Play / Pause
- Vor- und Zurückspulen
- Wechsel zum vorherigen bzw. nächsten Video
- Anpassen der Lautstärke
- Vollbildmodus
- Neuladen des Videos
- Wechsel zwischen den Bedienmodi

Die Anwendung verfügt über zwei unterschiedliche Modi mit jeweils spezifischen Interaktionskonzepten: den „Gesture Mode“ und den „Cursor Mode“.

### 2.1. Gesture Mode

Im „Gesture Mode“ kann das Video direkt über definierte Handgesten gesteuert werden. Damit eine Geste erkannt und korrekt ausgeführt wird, muss sie für eine bestimmte Haltezeit vor der Kamera gehalten werden. Dieses bewusste Halten der Geste erhöht die Erkennungsgenauigkeit und reduziert unbeabsichtigte Auslösungen.

### 2.2. Cursor Mode

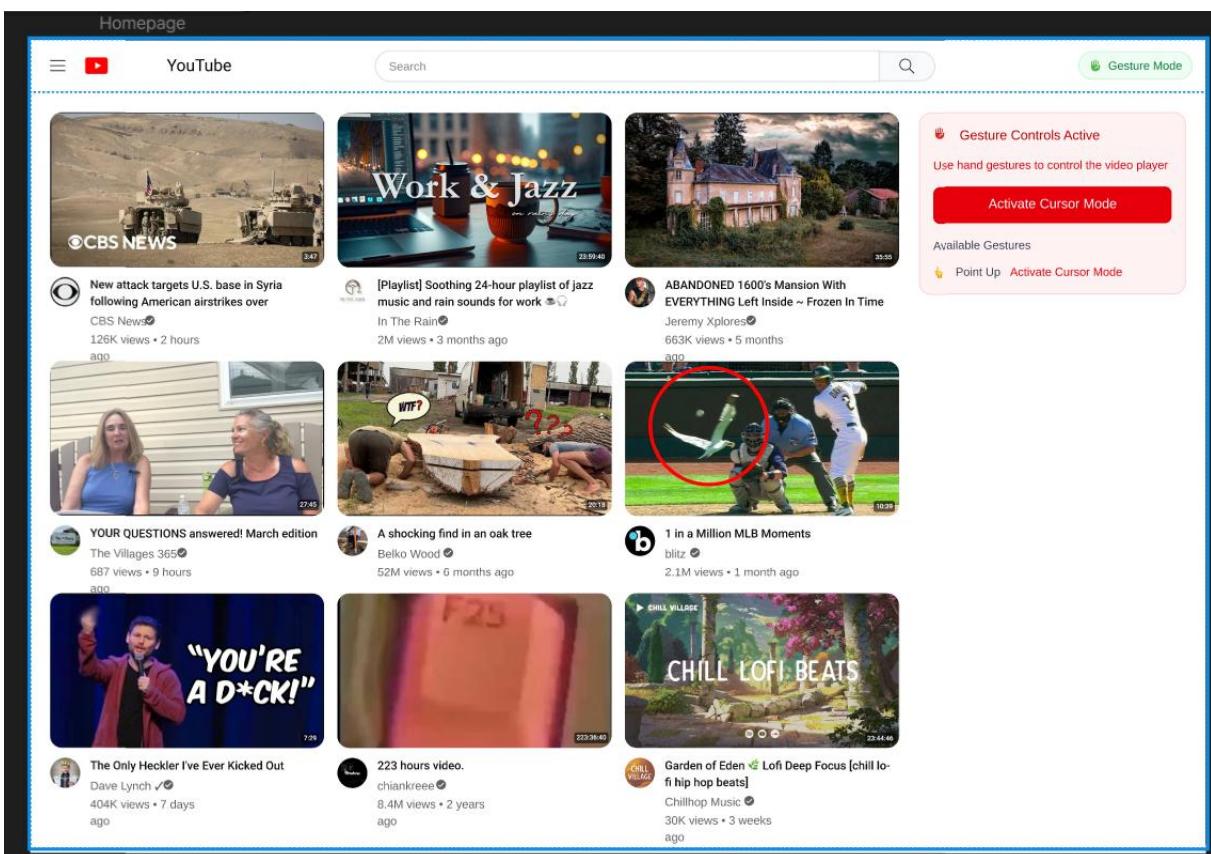
Der „Cursor Mode“ erweitert die Interaktion um einen virtuellen Cursor, der als roter Punkt auf dem Bildschirm dargestellt wird und über die Bewegung des Zeigefingers gesteuert wird. Mithilfe einer Zwei-Finger-Geste ist es möglich, durch die Videoübersicht zu scrollen, während eine zusätzliche Daumenbewegung (Pinch) eine Klick-Aktion auslöst. Dadurch wird eine präzise und zielgerichtete Navigation innerhalb der Benutzeroberfläche ermöglicht.

### 3. Gestaltungskonzept und Mock-up-Erstellung

Als Vorlage diente die Videoplattform YouTube. Diese Entscheidung bietet mehrere Vorteile für Nutzerinnen und Nutzer, darunter eine vertraute Bedienstruktur, ein bereits hochwertig gestaltetes Interface sowie eine geringe Einarbeitungszeit. Insbesondere im Hinblick auf Barrierefreiheit ist die Orientierung an einem etablierten und bekannten Design eine sinnvolle und unterstützende Maßnahme.

Für die Erstellung der Mock-ups wurde das Design-Tool Figma verwendet. Zunächst wurde mithilfe von Figma-Make eine Vorlage auf Basis der YouTube-Oberfläche erstellt und anschließend schrittweise weiterentwickelt. Der Schwerpunkt lag dabei auf dem gezielten Weglassen nicht notwendiger Elemente, um eine übersichtliche und reduzierte Benutzeroberfläche zu schaffen.

#### 3.1. Homepage

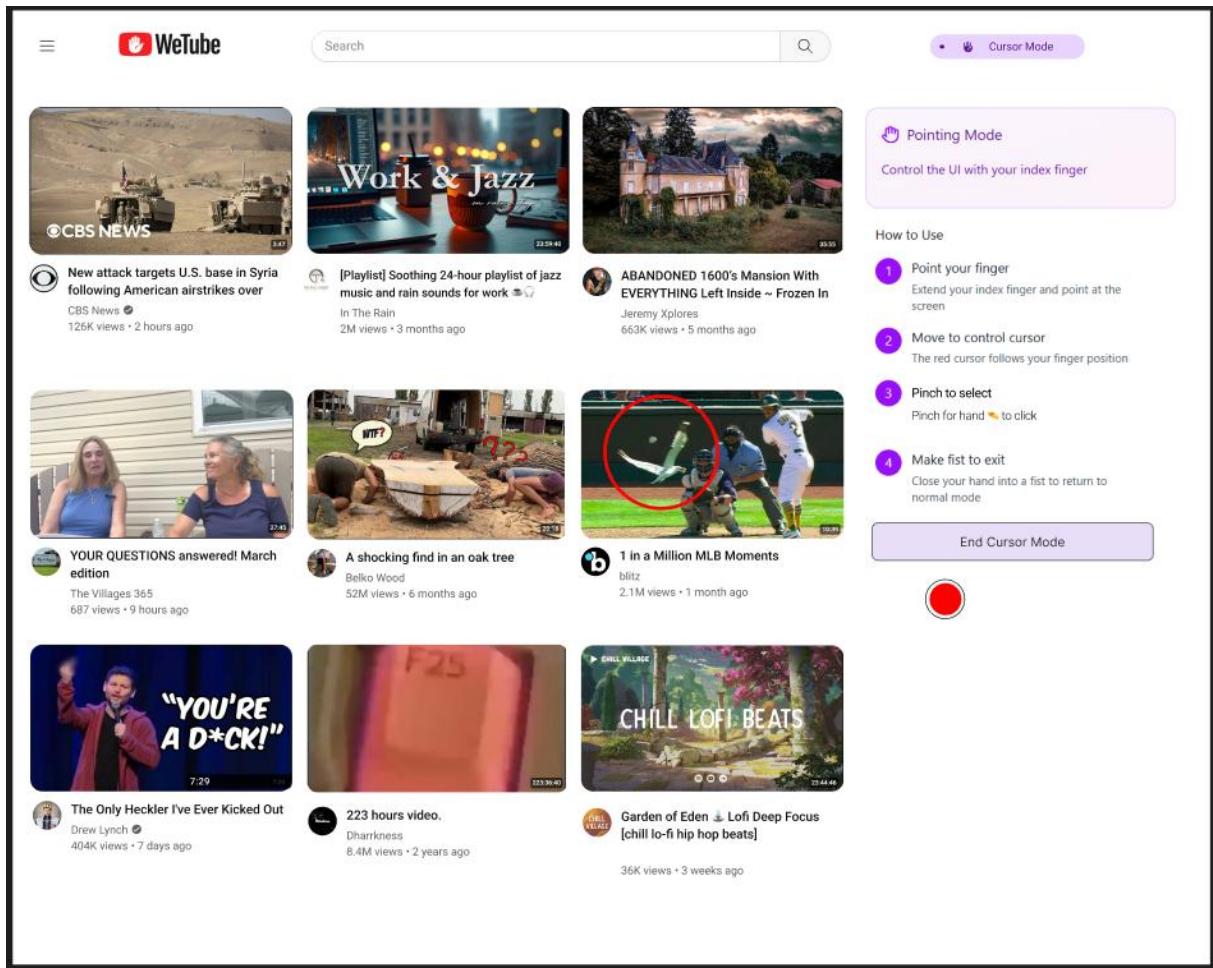


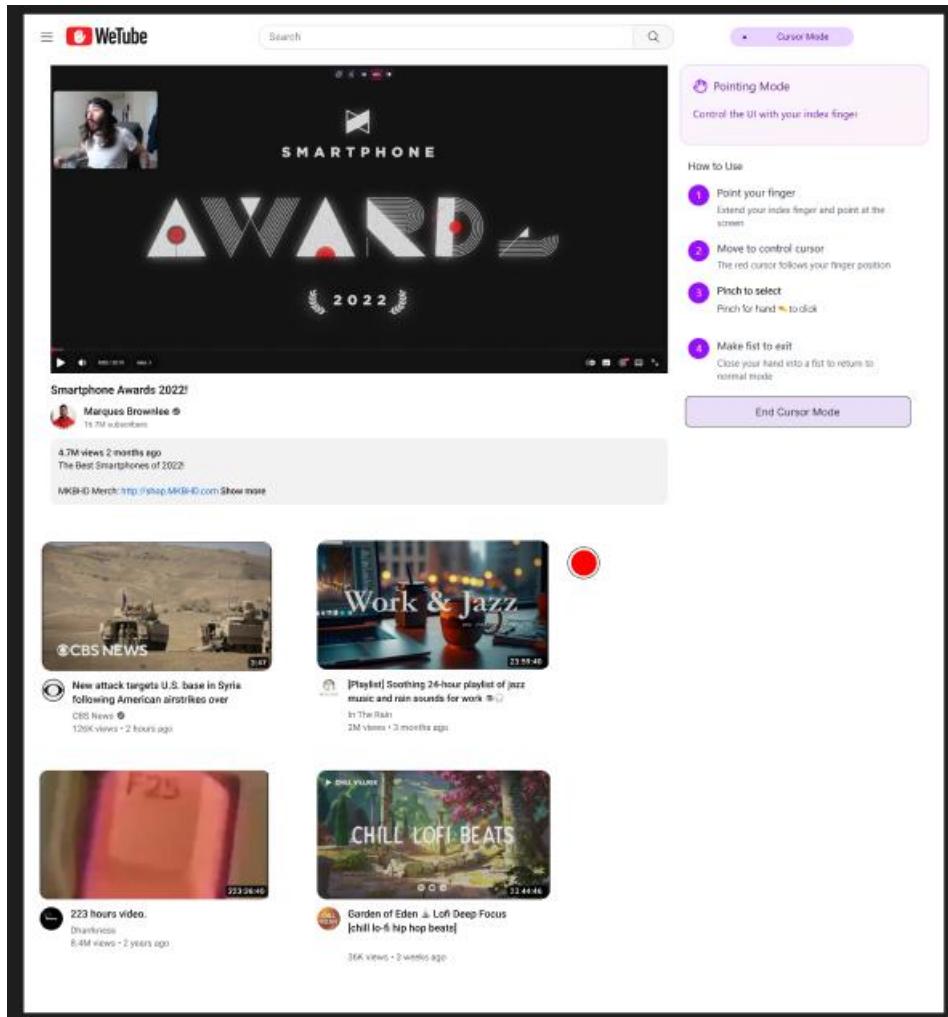
Während der Erstellung der ersten Folie (Homepage) wurde analysiert, welche Aktionen für Nutzerinnen und Nutzer vorgesehen sind und welche bewusst ausgeschlossen werden. Ziel war es, den Funktionsumfang auf wesentliche Interaktionen zu begrenzen und eine klare, übersichtliche Nutzung zu ermöglichen.

**Nicht vorgesehene Aktionen** sind unter anderem das Erstellen oder Hochladen von Videos, Spracheingabe, Abonnements, fortgeschrittene Sortierfunktionen (zumindest im Rahmen des MVP) sowie das Anlegen von Playlists.

**Vorgesehene Aktionen** umfassen das Auswählen und Anschauen von Videos, grundlegende Interaktionen mit der Benutzeroberfläche (z. B. Vollbildmodus, Vor- und Zurückspulen), die Nutzung der Suchfunktion sowie eine cursorartige Navigation zur Bedienung der Oberfläche.

### 3.2. Cursor Mode

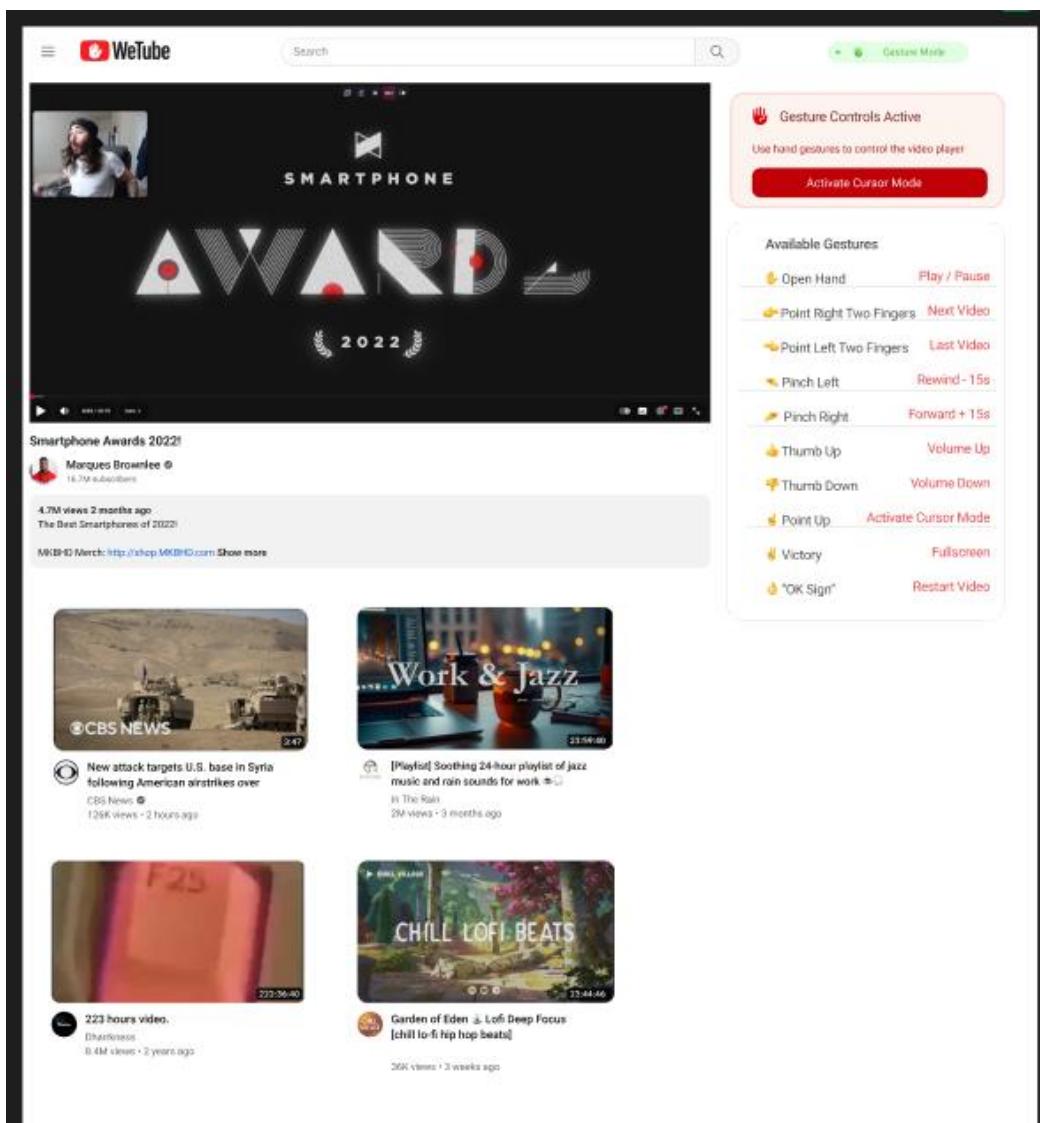




Erwähnenswert ist, dass die Aktion „Auswählen“ nicht durch das Einziehen des Zeigefingers, sondern durch eine Kneifbewegung (Pinch) ausgelöst wird. Diese Entscheidung dient dazu, eine mögliche Zweideutigkeit mit der Faustgeste (Fist), die zum Schließen oder Beenden verwendet wird, zu vermeiden und Fehlinterpretationen zu reduzieren.

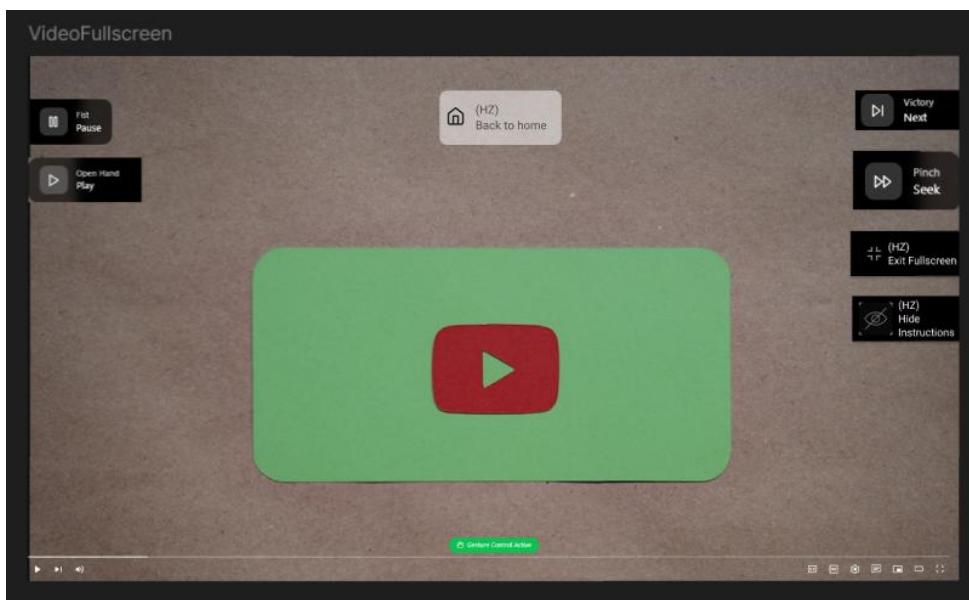
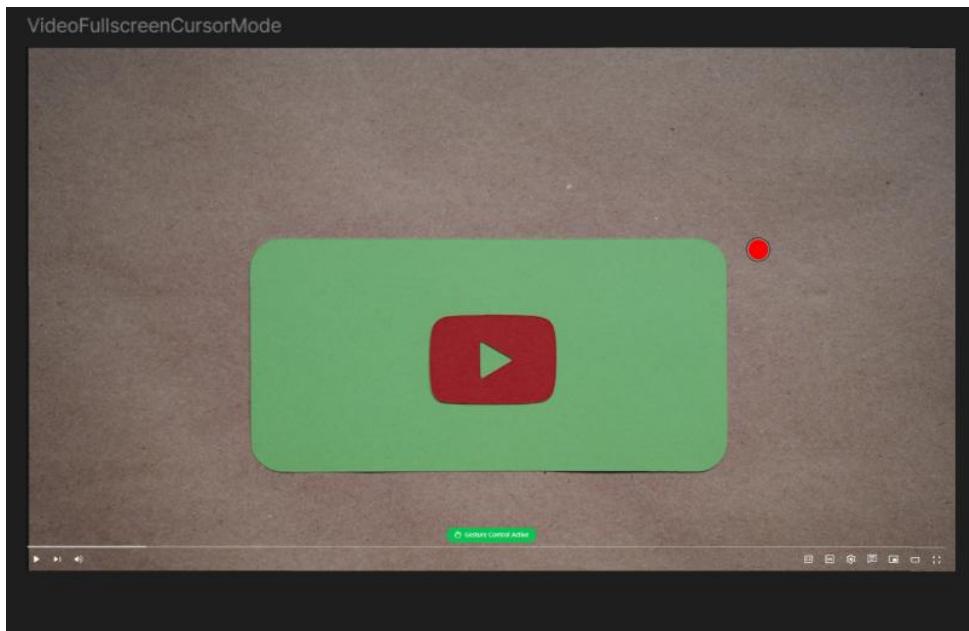
Zusätzlich wurde ein klickbarer Button zum Wechseln der Modi vorgesehen. Diese Ergänzung unterstützt eine intuitive Bedienung und bietet eine alternative, leicht verständliche Möglichkeit zum Modewechsel, ohne die gestenbasierte Interaktion einzuschränken. Im Nutzungskontext stellt dies eine sinnvolle und benutzerfreundliche Ergänzung dar.

### 3.3. Gesture Mode



Der Gesture Mode im Videoplayer stellt den zentralen Bestandteil der Anwendung dar und umfasst entsprechend eine Vielzahl an Gesten. Bei der Auswahl und Gestaltung dieser Gesten standen insbesondere die eindeutige Unterscheidbarkeit sowie die Natürlichkeit der Bewegungen aus einer geschlossenen Handposition im Vordergrund. Ziel war es, Gesten zu definieren, die intuitiv ausführbar sind und sich klar voneinander abgrenzen lassen.

## 4. Full Screen



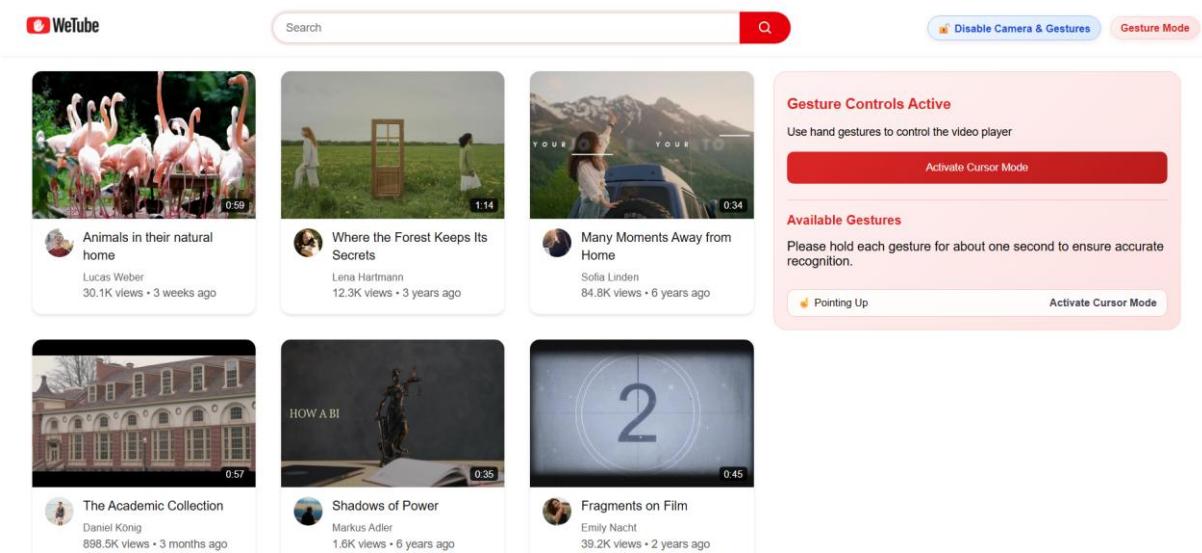
Der Vollbildmodus stellte einige wenige, jedoch zentrale Fragestellungen in den Fokus. Eine wesentliche Herausforderung bestand im Abwegen zwischen Größe und Erkennbarkeit der Bedienelemente und einem möglichst freien, unbeeinträchtigten Bildschirmbereich. Ziel war es, die Steuerungselemente auch im Vollbildmodus klar wahrnehmbar zu gestalten, ohne den Videoinhalt unnötig zu überdecken.

Im Ergebnis wurde festgelegt, dass die Symbole im Vollbildmodus etwas größer dargestellt werden sollten. Die konkrete und optimale Größe wurde bewusst der späteren Umsetzung überlassen, da sie im praktischen Einsatz und unter realen Nutzungsbedingungen besser evaluiert und angepasst werden kann.

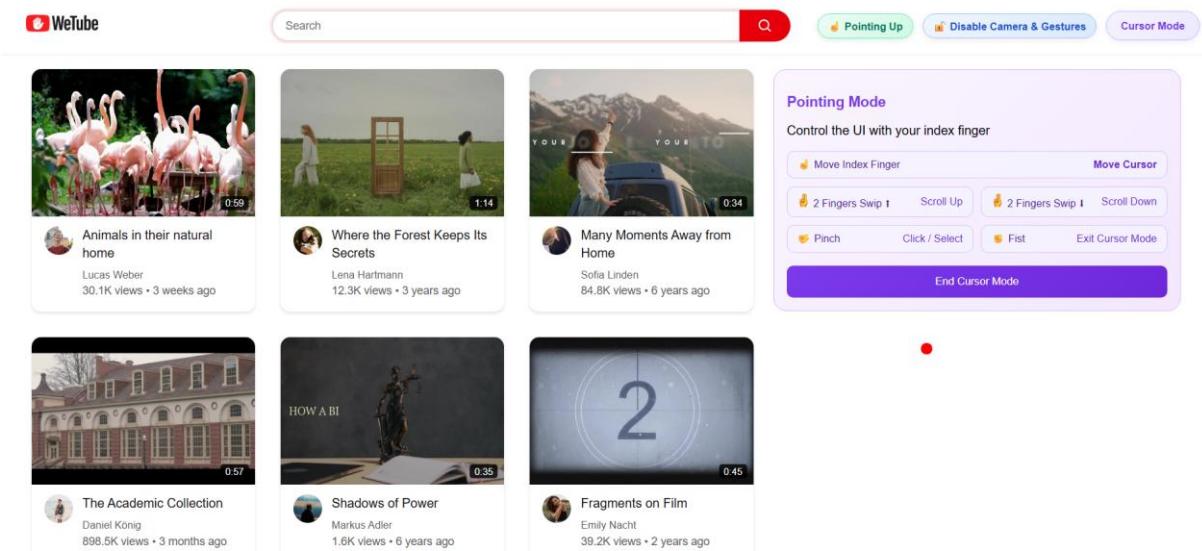
## 5. Finale Projekt

### 5.1. Homepage

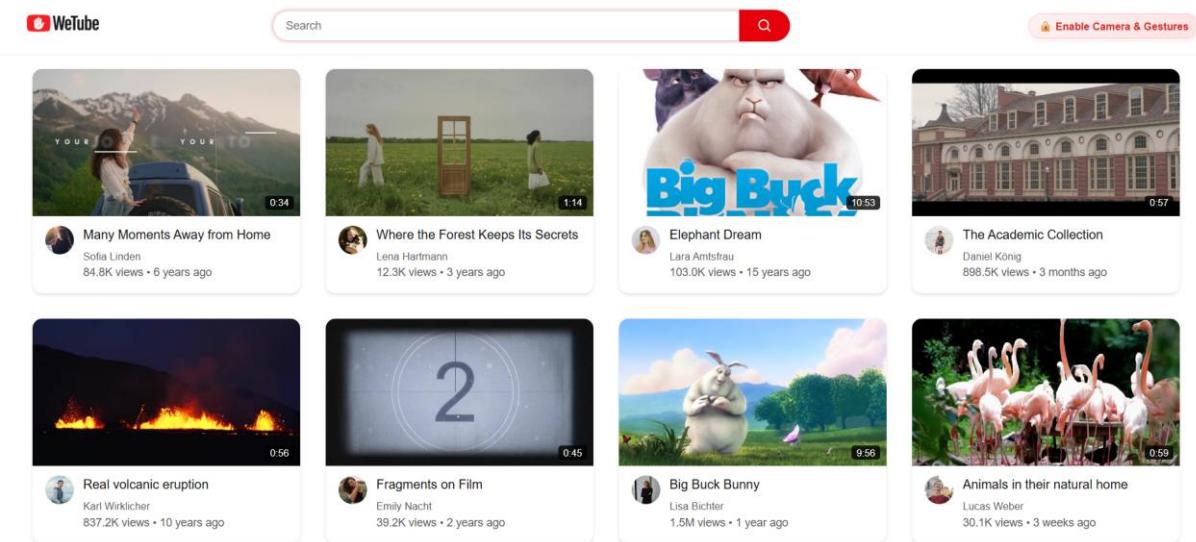
#### 5.1.1. Gesture Mode



#### 5.1.2. Cursor Mode

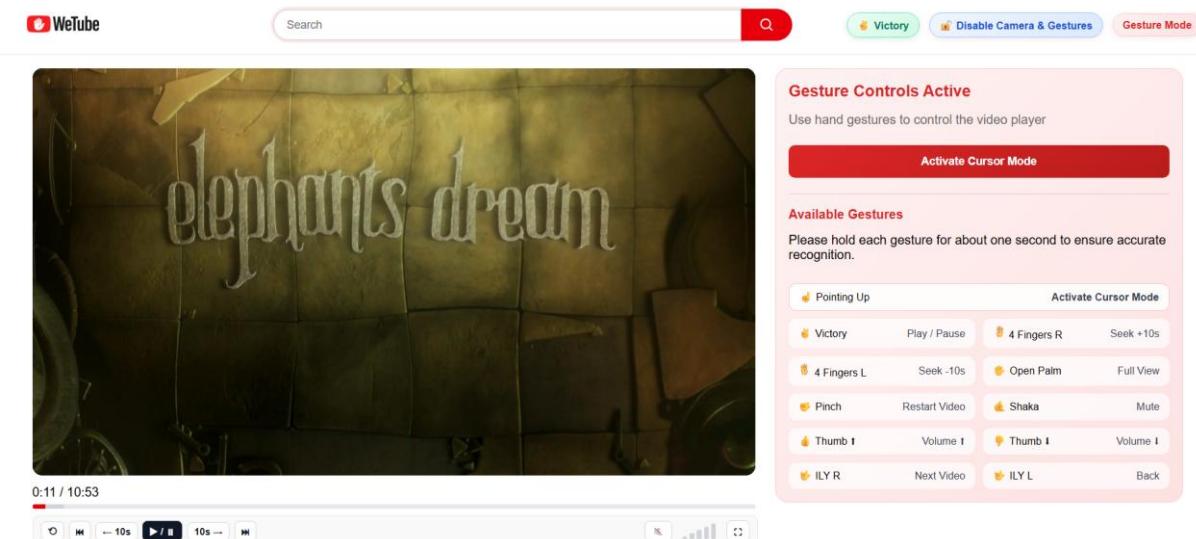


### 5.1.3. Deaktivierte Kamera Mode

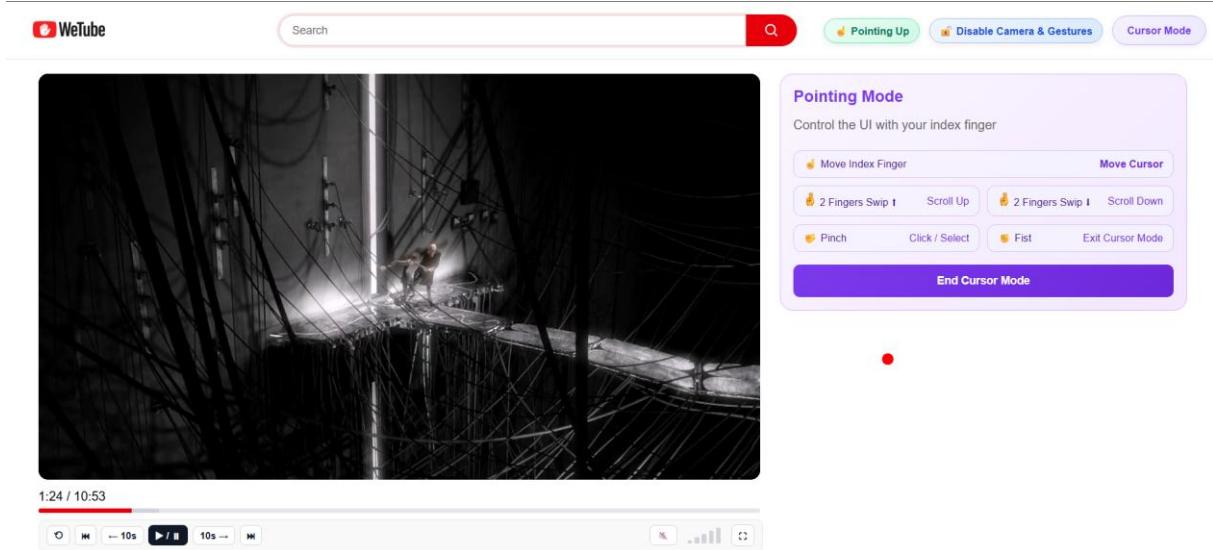


## 5.2. Videoseite

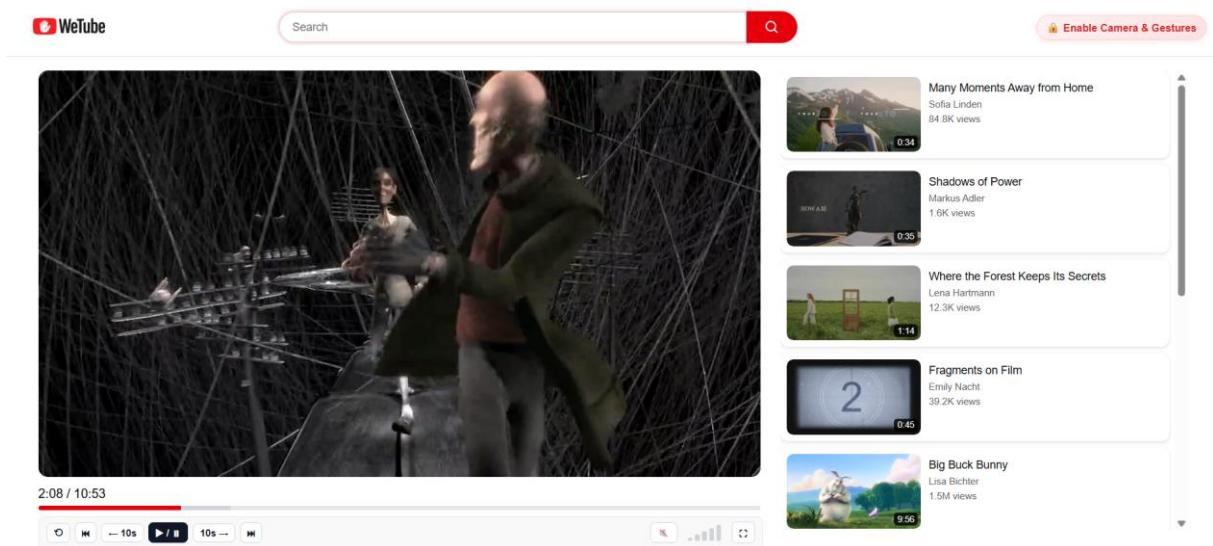
### 5.2.1. Gesture Mode



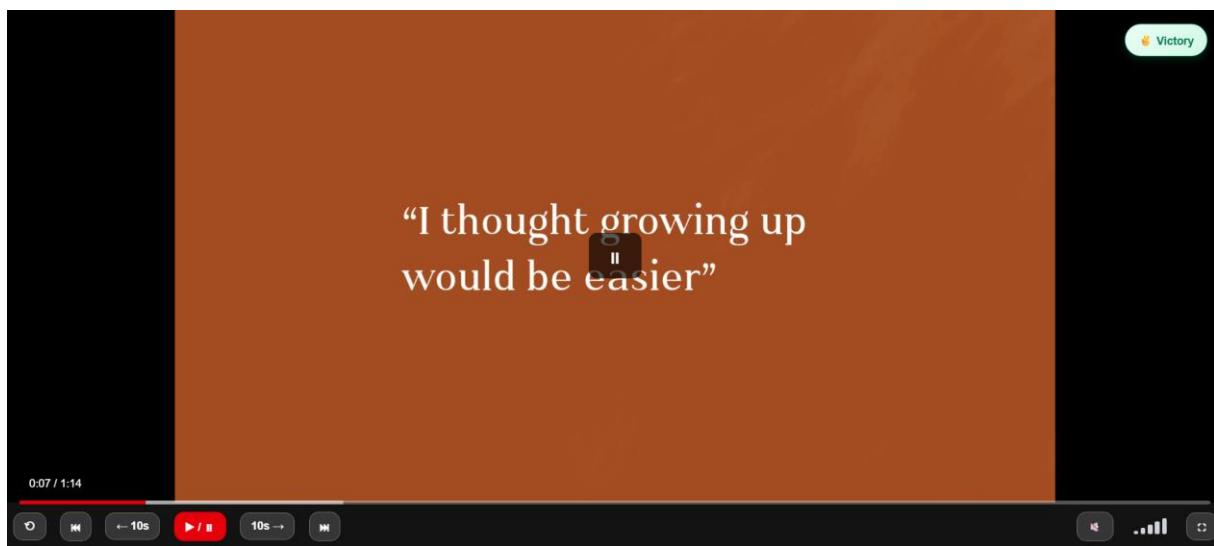
### 5.2.2. Cursor Mode



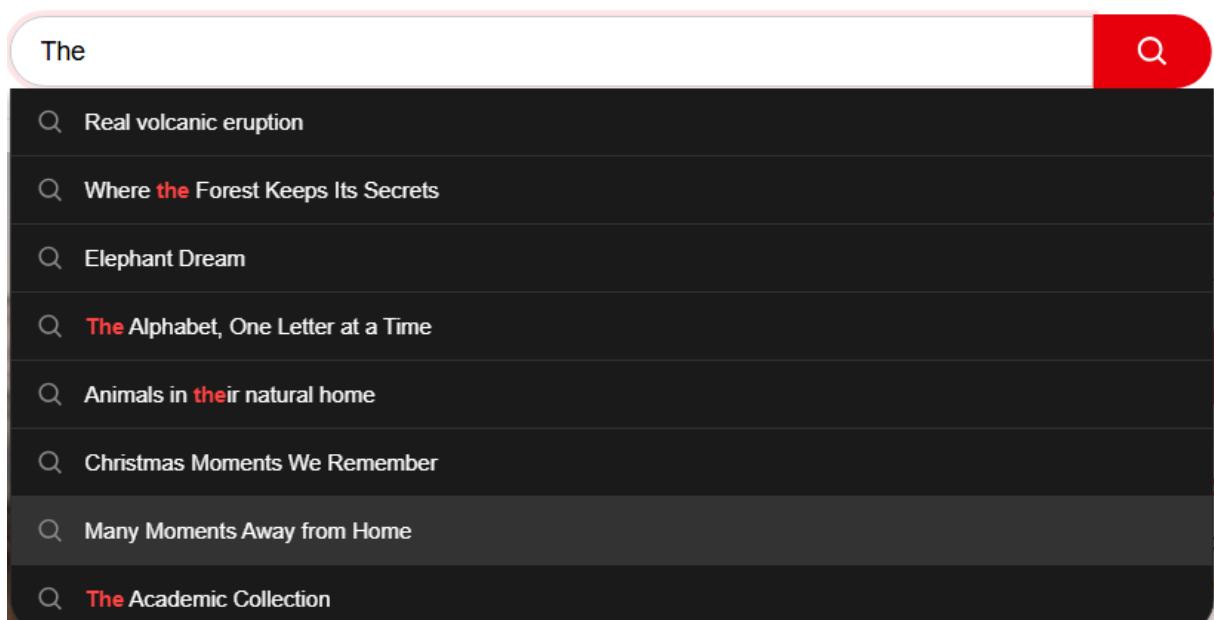
### 5.2.3. Deaktivierte Kamera Mode



## 5.2.4. Full Screen



## 5.3. Search-bar



## 6. Unterschiede zur Mock-Ups

### 6.1. Allgemeine Style

Im Vergleich zu den ursprünglichen Mock-Ups wurde der visuelle Stil der Anwendung weiter verfeinert und optimiert. Dabei wurden vor allem oberflächliche Gestaltungselemente wie Farbgebung, Abstände sowie die Größe und Hervorhebung von Symbolen angepasst. Diese Änderungen dienten in erster Linie der besseren Lesbarkeit, Erkennbarkeit und visuellen Hierarchie der Benutzeroberfläche, insbesondere im Hinblick auf die Nutzung aus größerer Distanz und die gestenbasierte Interaktion.

Die gestalterischen Anpassungen haben keinen Einfluss auf die grundlegende Funktionsweise, tragen jedoch wesentlich zu einer klareren, konsistenteren und nutzerfreundlicheren Darstellung der Oberfläche bei.

### 6.2. Verfügbare Gesten

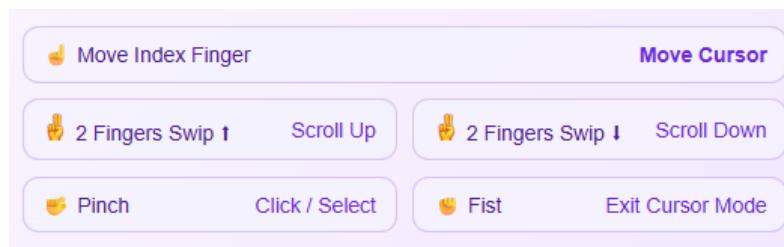
Während der Entwicklung wurden teilweise andere Gesten eingesetzt als ursprünglich geplant. Im Laufe der Implementierung zeigte sich, dass einige Gesten in der Praxis nur schwer zuverlässig voneinander zu unterscheiden waren. Diese Gesten wurden daher angepasst oder durch besser erkennbare Varianten ersetzt. Ein Beispiel dafür ist die Änderung einzelner OK-Zeichen, bei denen es häufig zu Fehlinterpretationen durch die Erkennung von Pinch kam. Bei der Gestenerkennung wurde außerdem festgestellt, dass eine kurze Haltezeit von etwa einer Sekunde notwendig ist, damit eine Geste zuverlässig erkannt und ausgelöst wird. Bei besonders sensiblen Gesten wie dem Neustart des Videos durch eine Pinch-Geste wurde die erforderliche Haltezeit erhöht, um unbeabsichtigte Auslösungen zu reduzieren.

Ursprünglich war ein zweistufiges Erkennungssystem mit einer vorgeschalteten Aktivierungs-geste mit vier Fingern geplant. Dieses Konzept erwies sich jedoch als wenig benutzerfreundlich, da sich Nutzerinnen und Nutzer stets an eine bestimmte Reihenfolge erinnern müssten und die Bedienung dadurch anstrengender wird. Zusätzlich würde ein solches Verfahren immer zu einer spürbaren Verzögerung führen, beispielsweise beim Pausieren eines Videos. Es ist uns bewusst, dass Gesteuerung grundsätzlich das Risiko unbeabsichtigter Auslösungen mit sich bringt. Das Problem besteht darin, dass das System die tatsächliche Intention der Nutzenden nie wirklich zuverlässig erkennen kann, sondern ausschließlich auf Basis der erkannten Handbewegungen reagiert.

Obwohl die Software technisch in der Lage ist, beide Hände gleichzeitig zu erkennen, wurde überwiegend eine handunabhängige Bedienung vorgesehen. Die meisten Gesten können sowohl mit der linken als auch mit der rechten Hand ausgeführt werden, um Barrierefreiheit und Nutzungskomfort zu erhöhen. Eine bewusste Ausnahme bilden die Gesten für Vor- und Zurückspulen sowie für den Wechsel zum nächsten bzw. vorherigen Video. In diesen Fällen wird zwischen linker und rechter Hand unterschieden, um eine klare Richtungszuordnung zu ermöglichen und die Navigation intuitiver zu gestalten.

Finale verfügbare Gesten:

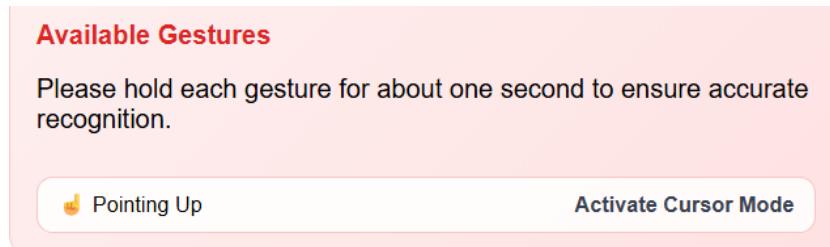
- Cursor Mode



- Gesture Mode (Home)

### Available Gestures

Please hold each gesture for about one second to ensure accurate recognition.

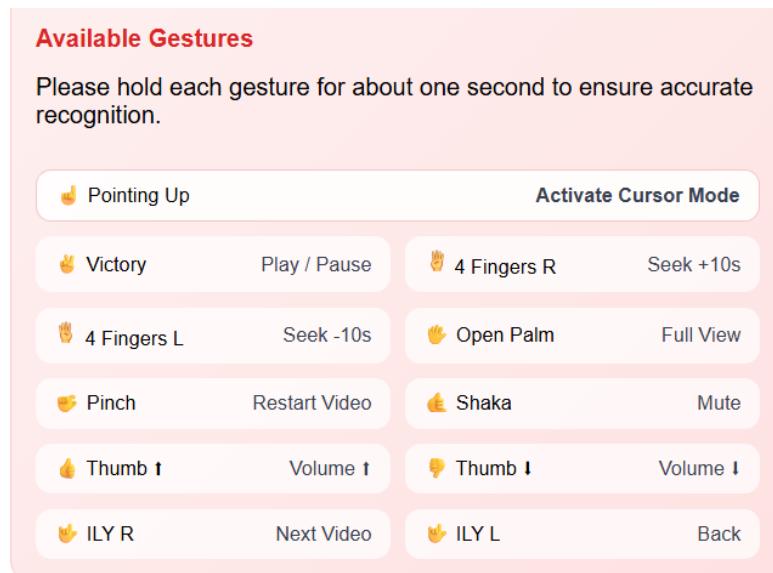


A screenshot showing a single gesture icon for Pointing Up, followed by the text "Activate Cursor Mode".

- Gesture Mode (Videoseite)

### Available Gestures

Please hold each gesture for about one second to ensure accurate recognition.



A screenshot showing a grid of gesture icons and their corresponding functions. The icons are color-coded: blue for Pointing Up, orange for Victory, 4 Fingers L, Pinch, Thumb ↑, ILY R, and Play / Pause; red for 4 Fingers R, Open Palm, Shaka, Volume ↑, ILY L, and Seek +10s; green for Seek -10s, Full View, Mute, and Back. The functions are: Activate Cursor Mode, Play / Pause, Seek +10s, Seek -10s, Full View, Mute, and Back.

## 6.3. Videoplayer Controls und Full Screen

Für den Videoplayer wurden die Control-Buttons sowie die Vollbildansicht manuell implementiert. Diese Entscheidung basiert auf zwei wesentlichen Gründen. Zum einen erfordert der native Vollbildmodus des Browsers eine wiederholte Abfrage von Nutzerberechtigungen, was den Interaktionsfluss insbesondere bei gestenbasierter Bedienung unterbricht. Zum anderen sind die standardmäßigen Videosteuerungen nicht zuverlässig mit einem virtuellen Cursor bedienbar, da sie primär für Maus- oder Touch-Interaktionen ausgelegt sind.

Durch die eigene Implementierung der Bedienelemente konnte eine konsistente, gesten- und cursorfreundliche Steuerung realisiert werden, die unabhängig von Browser-Restriktionen funktioniert und eine nahtlose Nutzererfahrung ermöglicht.

## 6.4. Feedback

Das Gesture Feedback Badge gibt dem Nutzer eine unmittelbare visuelle Rückmeldung, sobald eine Geste erkannt wurde. Da bei gestenbasierter Interaktion kein physisches Feedback wie ein Mausklick oder Tastendruck vorhanden ist, ist diese Rückmeldung besonders wichtig. Sie reduziert Unsicherheit, verhindert mehrfaches oder übertriebenes Ausführen von Gesten und stärkt das Vertrauen des Nutzers in das System. Dadurch wird die Interaktion insgesamt verständlicher, präziser und angenehmer.



Zusätzlich zum Gesten-Feedback wird bei ausgeführten Videoaktionen ein visuelles Video-Action-Feedback eingeblendet, beispielsweise beim Ändern der Lautstärke oder beim Wechsel zwischen Wiedergabezuständen. Dieses Feedback erscheint zentral über dem Videointhalt und informiert die Nutzerinnen und Nutzer unmittelbar über die ausgeführte Aktion sowie deren aktuellen Zustand (z. B. Lautstärke in Prozent).

Das Video-Action-Feedback dient der Bestätigung erfolgreicher Eingaben und erhöht die Transparenz der Interaktion, insbesondere bei gestenbasierter Steuerung ohne haptisches Feedback. Dadurch wird die Bedienung nachvollziehbarer und Fehlbedienungen können schneller erkannt und korrigiert werden.



## 6.5. Aktivierung / Deaktivierung von Kamera und Gestures

Die Möglichkeit, Kamera und Gestensteuerung gezielt zu aktivieren oder zu deaktivieren, erhöht die Flexibilität und Benutzerfreundlichkeit des Systems. Ist die Gestensteuerung aktiv, wird die Benutzeroberfläche an die gestenbasierte Interaktion angepasst und unterstützt die Bedienung ohne Maus oder Tastatur. Wird die Kamera bzw. Gestensteuerung deaktiviert, wechselt das Layout bewusst in eine klassische, mausbasierte Ansicht, vergleichbar mit herkömmlichen Video-Plattformen wie YouTube.

Um eine unbeabsichtigte Deaktivierung während der gestenbasierten Nutzung zu vermeiden, ist der entsprechende Button nicht über Gesten auslösbar. Eine Deaktivierung ist ausschließlich über den Cursor Mode möglich, in dem der Button gezielt angeklickt werden kann. Da ein versehentliches Auslösen per virtuellem Cursor als sehr unwahrscheinlich gilt, stellt diese Lösung einen sinnvollen Kompromiss zwischen Sicherheit und Bedienbarkeit dar.

 **Disable Camera & Gestures**

 **Enable Camera & Gestures**

## 6.6. Videoliste & Search Bar

Die Videoliste wird bei jedem Laden der Seite in zufälliger Reihenfolge generiert. Dieses Verhalten orientiert sich an bekannten Videoplattformen und unterstützt die Entdeckung neuer Inhalte. Die Search Bar ist in der Anwendung dort verfügbar, wo sie im Interface angezeigt wird, und ermöglicht eine schnelle Suche nach Videos. Während der Eingabe greift die Search Bar auf und stellt eine Autocomplete-Funktion bereit, die passende Vorschläge basierend auf Videotiteln, Beschreibungen und Creator-Namen anzeigt. Dies verbessert die Benutzerfreundlichkeit und Effizienz der Suche, da Suchanfragen schneller und mit weniger Tippaufwand durchgeführt werden können.

## 7. Verwendetes API

Für die Handerkennung wurde Google MediaPipe Hands (HandLandmarker) eingesetzt. Dies ermöglicht eine zuverlässige Erkennung von Händen in Echtzeit. Ein großer Vorteil von MediaPipe ist, dass die Verarbeitung direkt im Browser erfolgt und somit keine spezielle Hardware erforderlich ist. Die von MediaPipe erkannten Hand-Landmarks bilden die Grundlage für die Gestenerkennung. Einige Gesten basieren auf der integrierten Erkennung von MediaPipe (GestureRecognizer), während weitere Gesten auf Basis der Landmark-Daten individuell definiert wurden.

## 8. Fazit

Mit WeTube wurde eine Videoplattform konzipiert und prototypisch umgesetzt, die eine kontaktlose, gestenbasierte Bedienung in den Mittelpunkt stellt und sich dabei bewusst an etablierten Videoplattformen wie YouTube orientiert. Durch diese Kombination aus vertrautem Interface und innovativem Interaktionskonzept wird eine niedrige Einstiegshürde geschaffen und gleichzeitig ein deutlicher Mehrwert für Nutzungsszenarien ohne Maus, Tastatur oder Touchscreen geboten.

Die Trennung in Gesture Mode und Cursor Mode ermöglicht unterschiedliche Formen der Interaktion und erlaubt es, sowohl einfache als auch präzisere Bedienhandlungen abzudecken. Besondere Aufmerksamkeit wurde der Gestenerkennung, der Vermeidung von Fehlbedienungen sowie der visuellen Rückmeldung durch Feedback-Elemente gewidmet, um Transparenz und Vertrauen in die Interaktion zu gewährleisten. Auch Entscheidungen wie die manuelle Implementierung der Videosteuerung oder die bewusste Einschränkung bestimmter Funktionen im MVP tragen zu einer konsistenten und nutzerfreundlichen Umsetzung bei.

Insgesamt zeigt das Projekt, dass gestenbasierte Steuerung eine sinnvolle Ergänzung zu klassischen Eingabemethoden darstellen kann, insbesondere im Hinblick auf Barrierefreiheit, Komfort und alternative Nutzungssituationen. WeTube bildet damit eine solide Grundlage für weiterführende Entwicklungen, beispielsweise im Bereich der Gestenverfeinerung, Personalisierung oder erweiterten Accessibility-Funktionen.