DynamicSlide: Reference-based Interaction Techniques for Slide-based Lecture Videos

Hyeungshik Jung, KAIST
Hijung Valentina Shin, Adobe Research
Juho Kim, KAIST

kixlab

KAIST School of Computing

MOOC and video

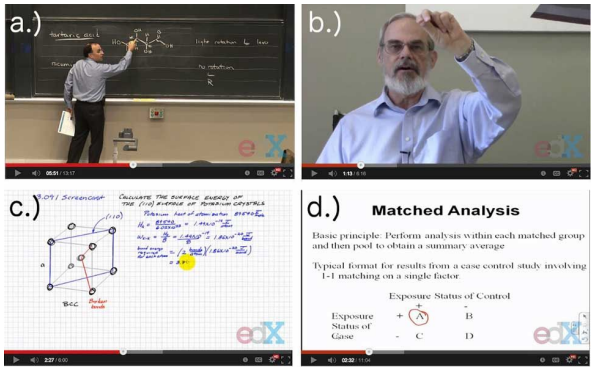
CLASS CENTRAL

78M
Students

800+
Universities

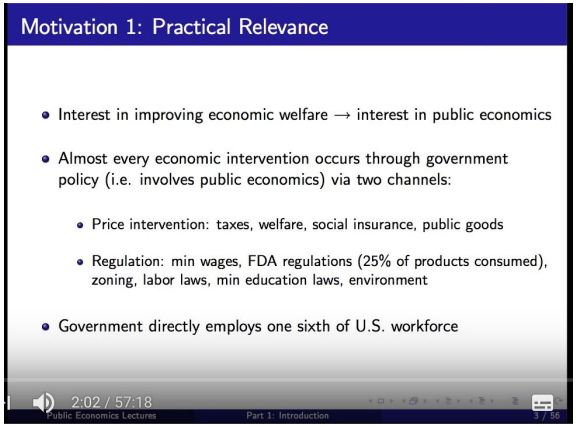
9.4k
Courses

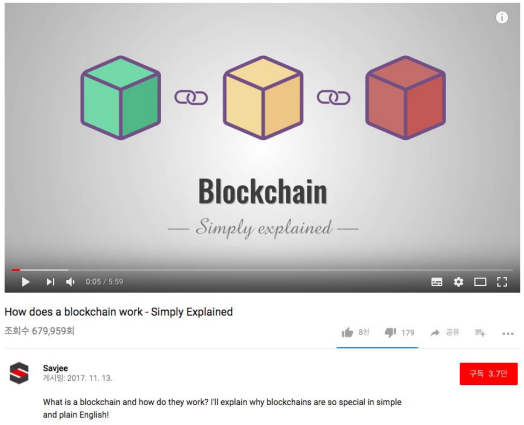
By the Numbers: MOOCs in 2017

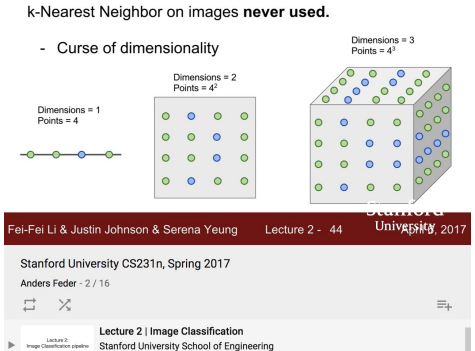


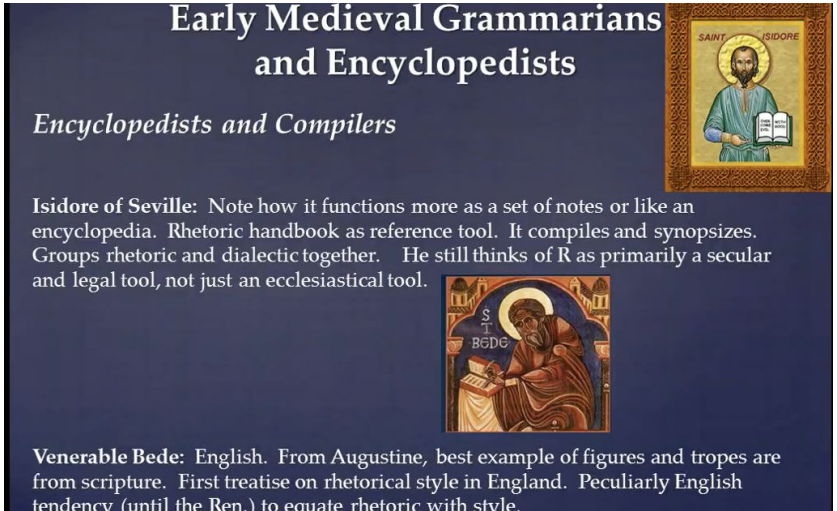
https://www.class-central.com/report/moocs-stats-and-trends-2017/

Slides are widely used in lecture videos

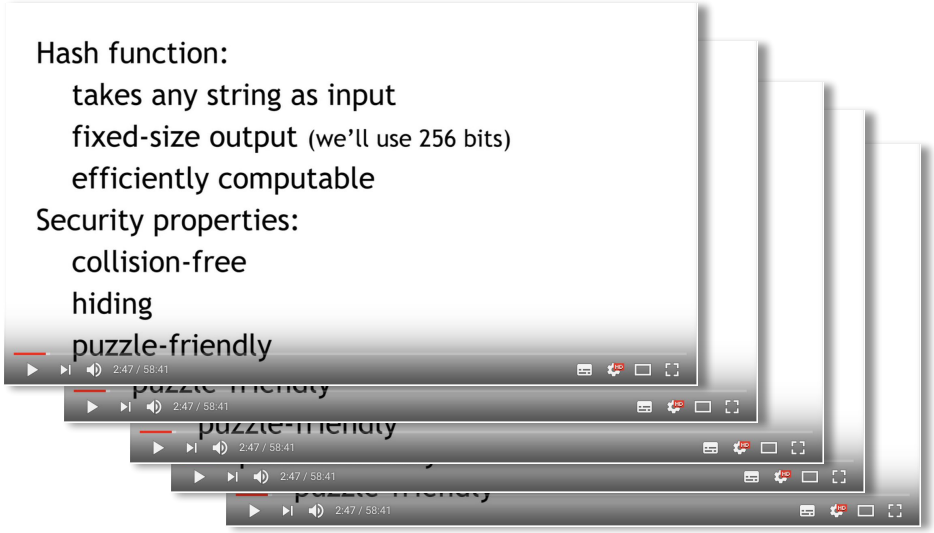


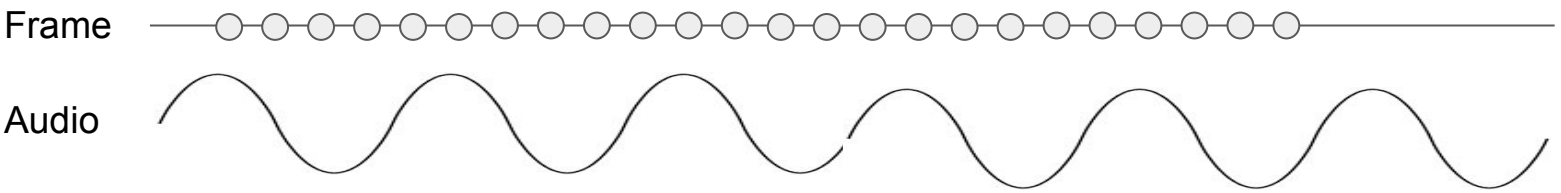




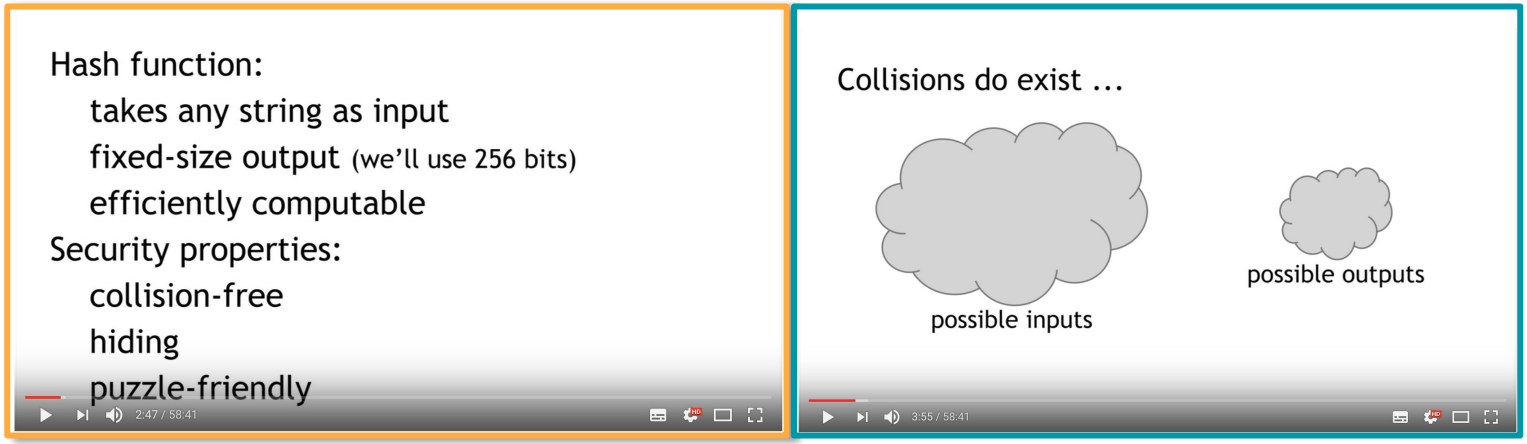


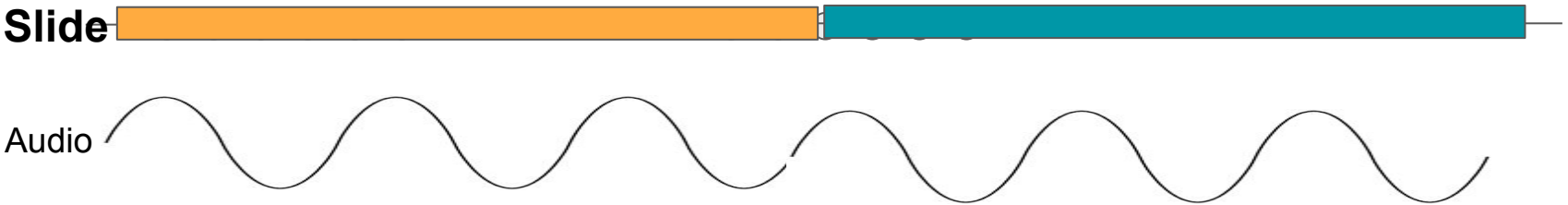
Decomposing Slide-based Video



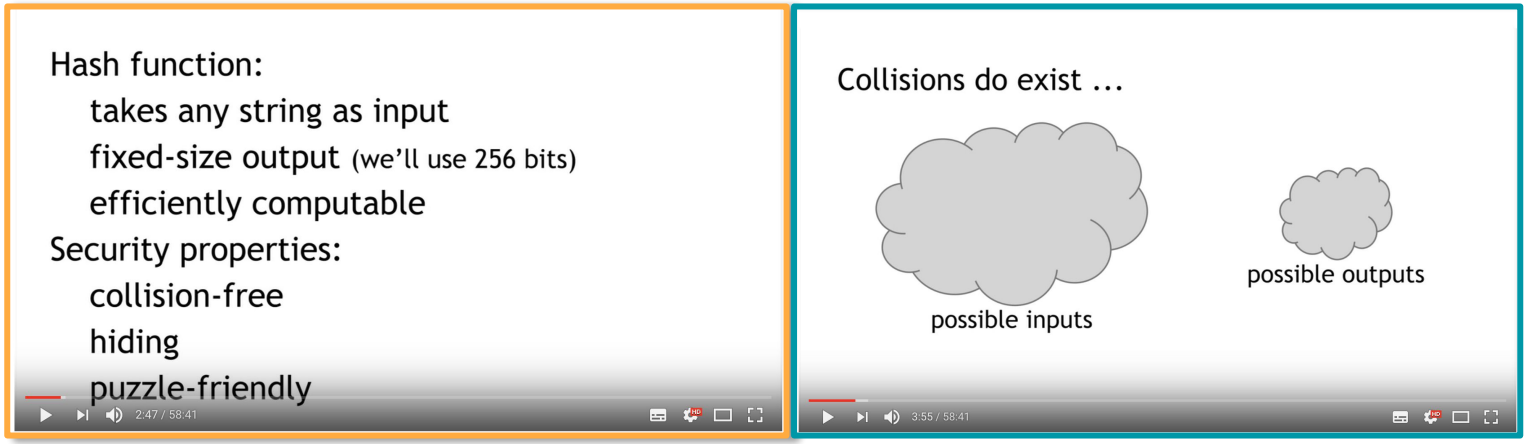


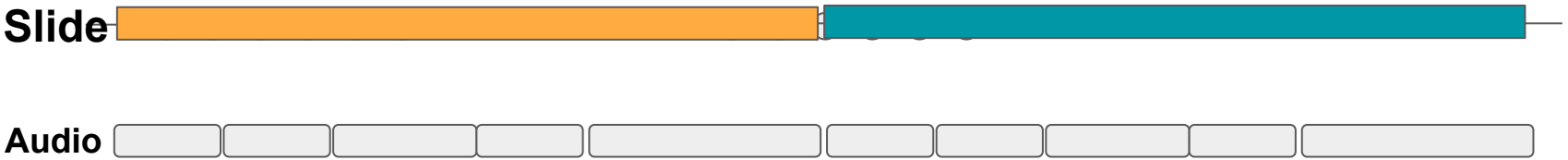
Decomposing Slide-based Video



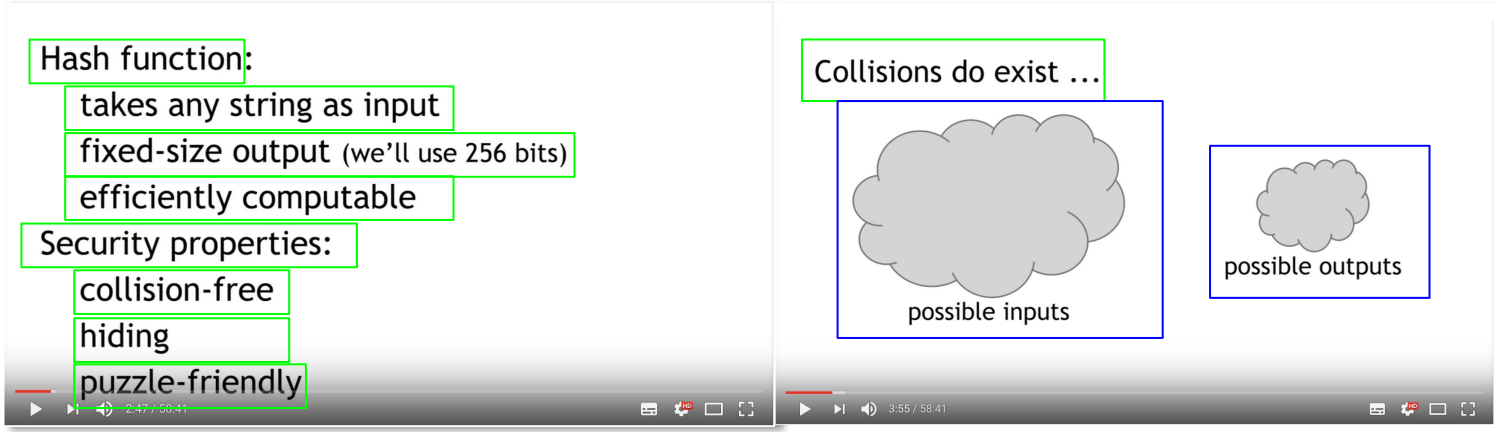


Decomposing Slide-based Video





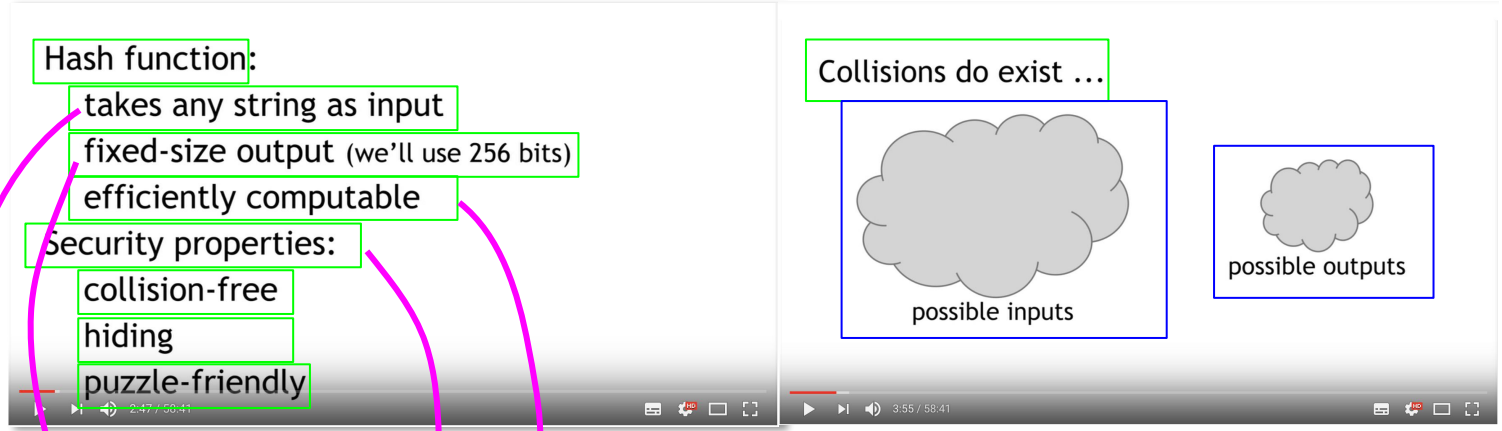
Decomposing Slide-based Video



timeframe

Slide	"Hash function"	"Hash function"
	"Takes any string as input"	Image1
	"Fixed-size output (we'll use 256 bits)"	Image1
Audio		

Decomposing Slide-based Video



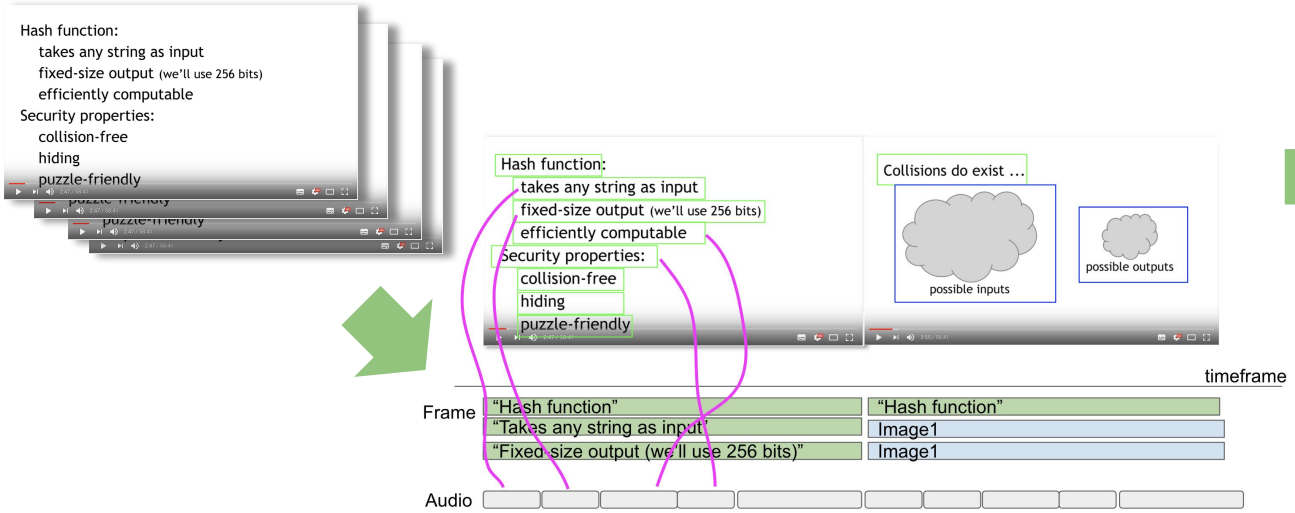
timeframe

Frame

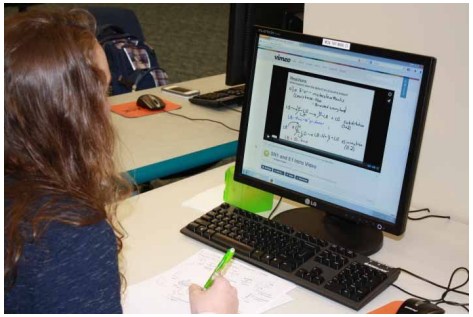
"Hash function"	"Hash function"
"Takes any string as input"	Image1
"Fixed-size output (we'll use 256 bits)"	Image1

Audio

Object-oriented tools for both learners & instructors

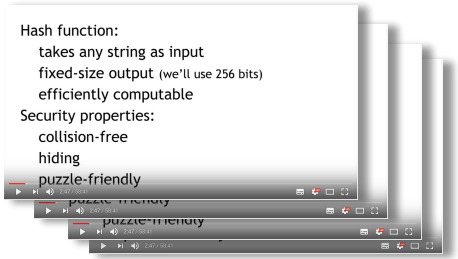


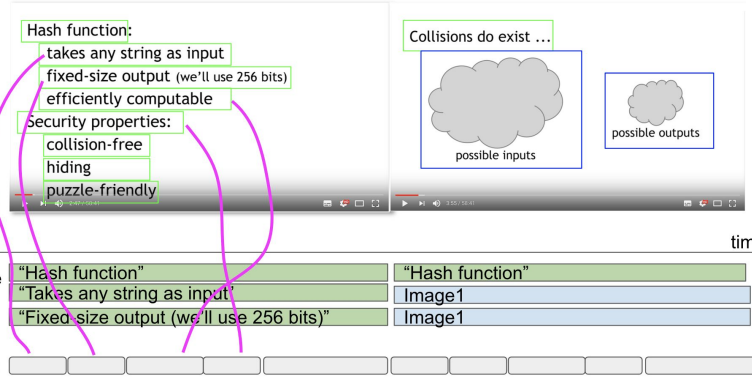
For Learners (Player)



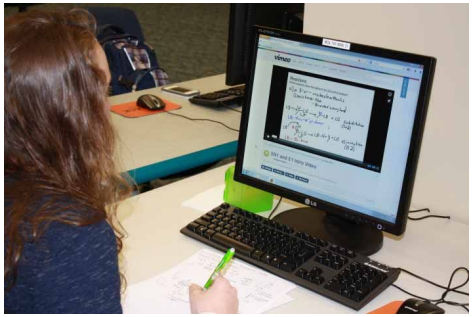
For Instructors
(Lightweight editor)

Object-oriented tools for both learners & instructors



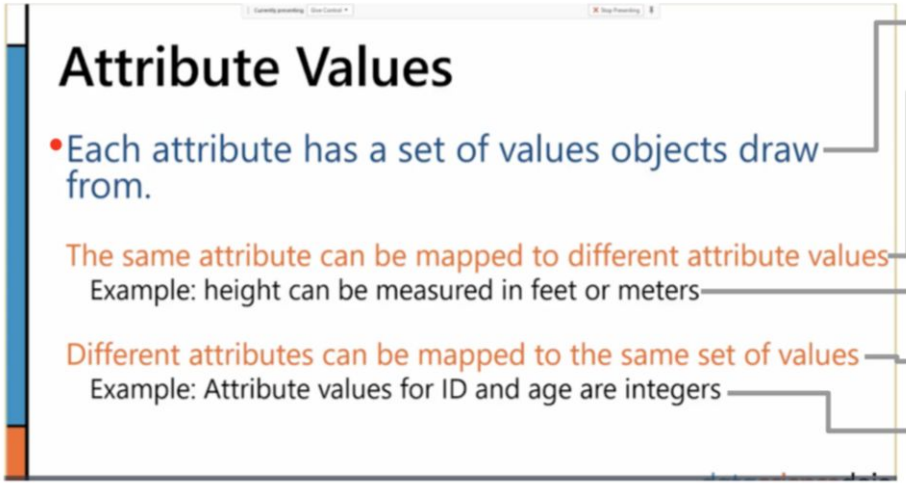


For Learners (Player)



For Instructors
(Lightweight editor)

Link between slide and narration



0:17 So each attribute, each object is defined by a set of attribute values.

0:25 And each attribute we can think of as being defined by the set of values that it can hold.

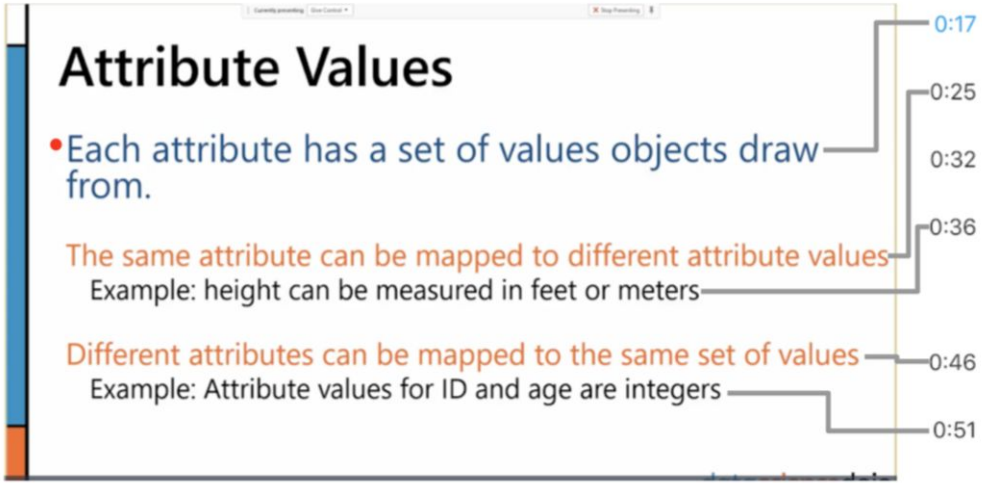
0:32 So we can have the same attribute mapped to different attribute values.

0:36 Height can be measured in meters or feet, temperature can be measured in Celsius, Kelvin, or Fahrenheit, lots of other sorts of things like that.

0:46 And different attributes will often be mapped to the same set of values.

0:51 ID numbers and age are both usually given as integer values.

Link between slide and narration



0:17 So each attribute, each object is defined by a set of attribute values.

0:25 And each attribute we can think of as being defined by the set of values that it can hold.

0:32 So we can have the same attribute mapped to different attribute values.

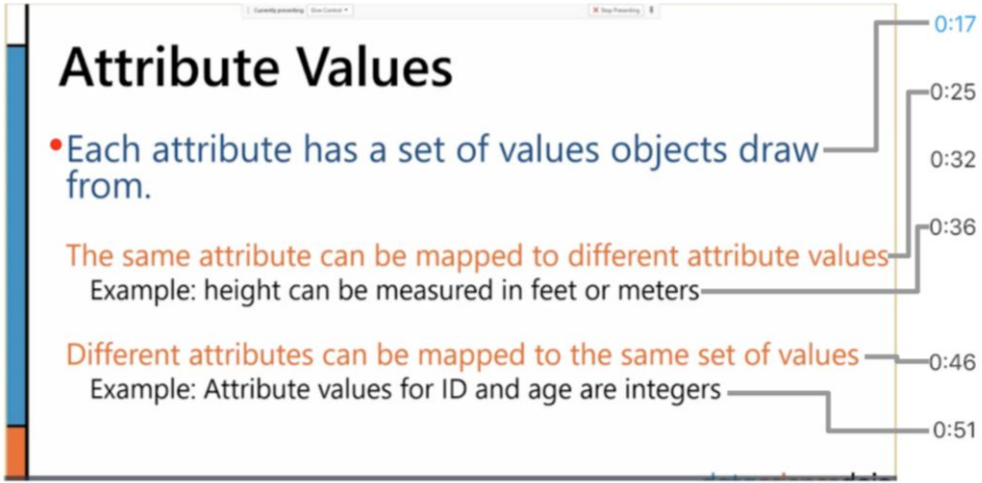
0:36 Height can be measured in meters or feet, temperature can be measured in Celsius, Kelvin, or Fahrenheit, lots of other sorts of things like that.

0:46 And different attributes will often be mapped to the same set of values.

0:51 ID numbers and age are both usually given as integer values.

visual
information hierarchy

Link between slide and narration



0:17 So each attribute, each object is defined by a set of attribute values.

0:25 And each attribute we can think of as being defined by the set of values that it can hold.

0:32 So we can have the same attribute mapped to different attribute values.

0:36 Height can be measured in meters or feet, temperature can be measured in Celsius, Kelvin, or Fahrenheit, lots of other sorts of things like that.

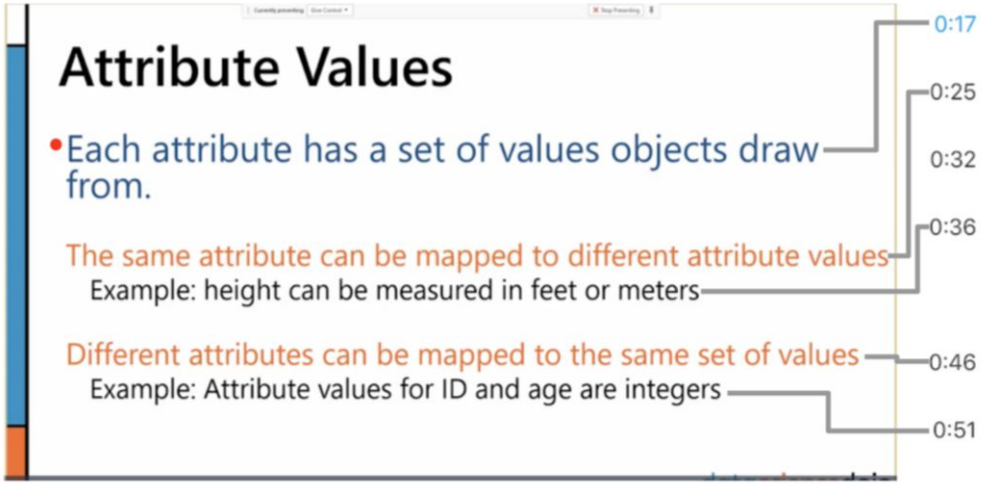
0:46 And different attributes will often be mapped to the same set of values.

0:51 ID numbers and age are both usually given as integer values.

visual
information hierarchy

auditory
detailed explanation

Link between slide and narration



0:17 So each attribute, each object is defined by a set of attribute values.

0:25 And each attribute we can think of as being defined by the set of values that it can hold.

0:32 So we can have the same attribute mapped to different attribute values.

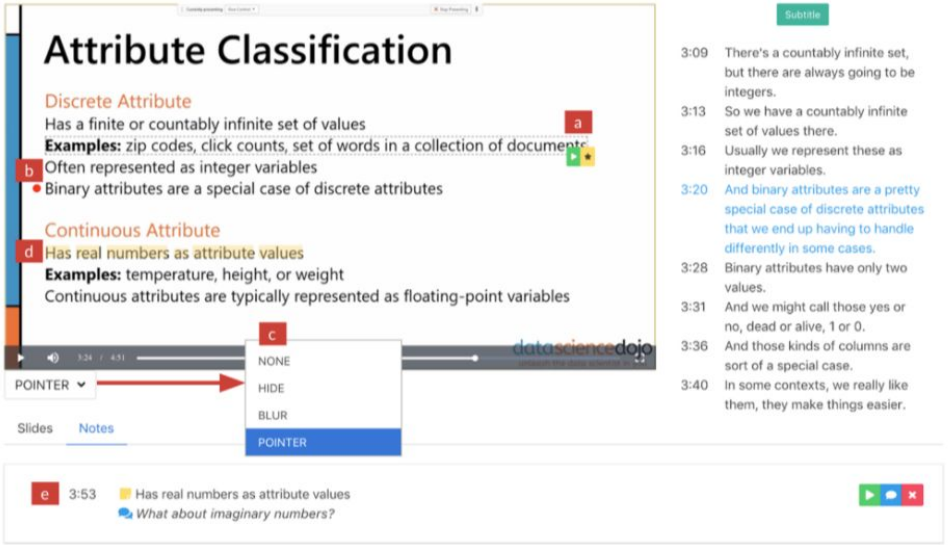
0:36 Height can be measured in meters or feet, temperature can be measured in Celsius, Kelvin, or Fahrenheit, lots of other sorts of things like that.

0:46 And different attributes will often be mapped to the same set of values.

0:51 ID numbers and age are both usually given as integer values.

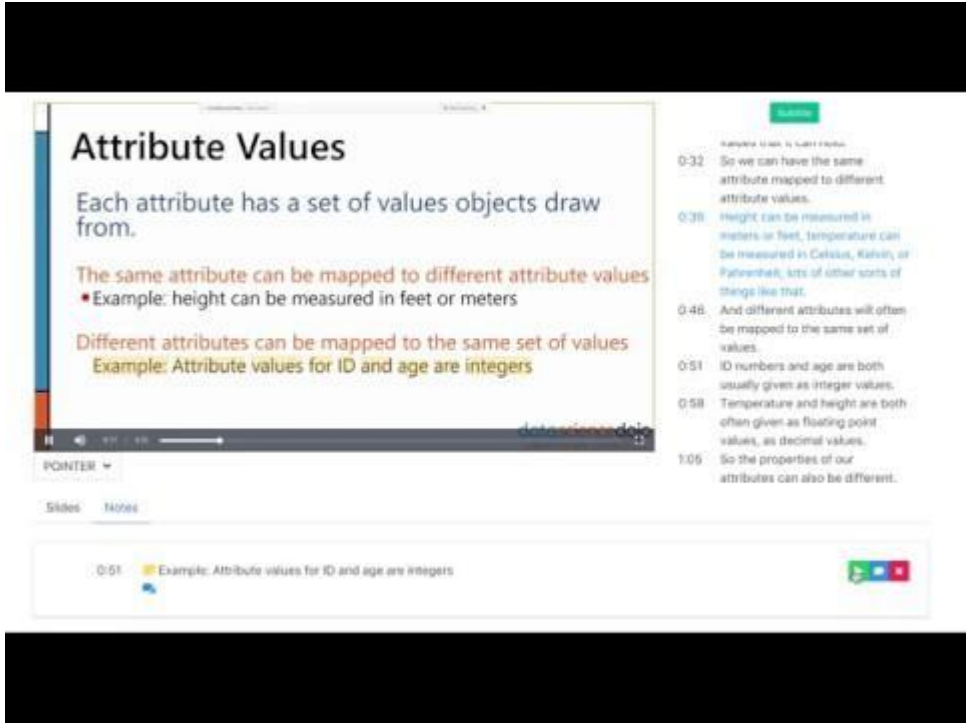
How can we find leverage links between slides and narration for watching video?

Interaction techniques powered by links between slides and narration



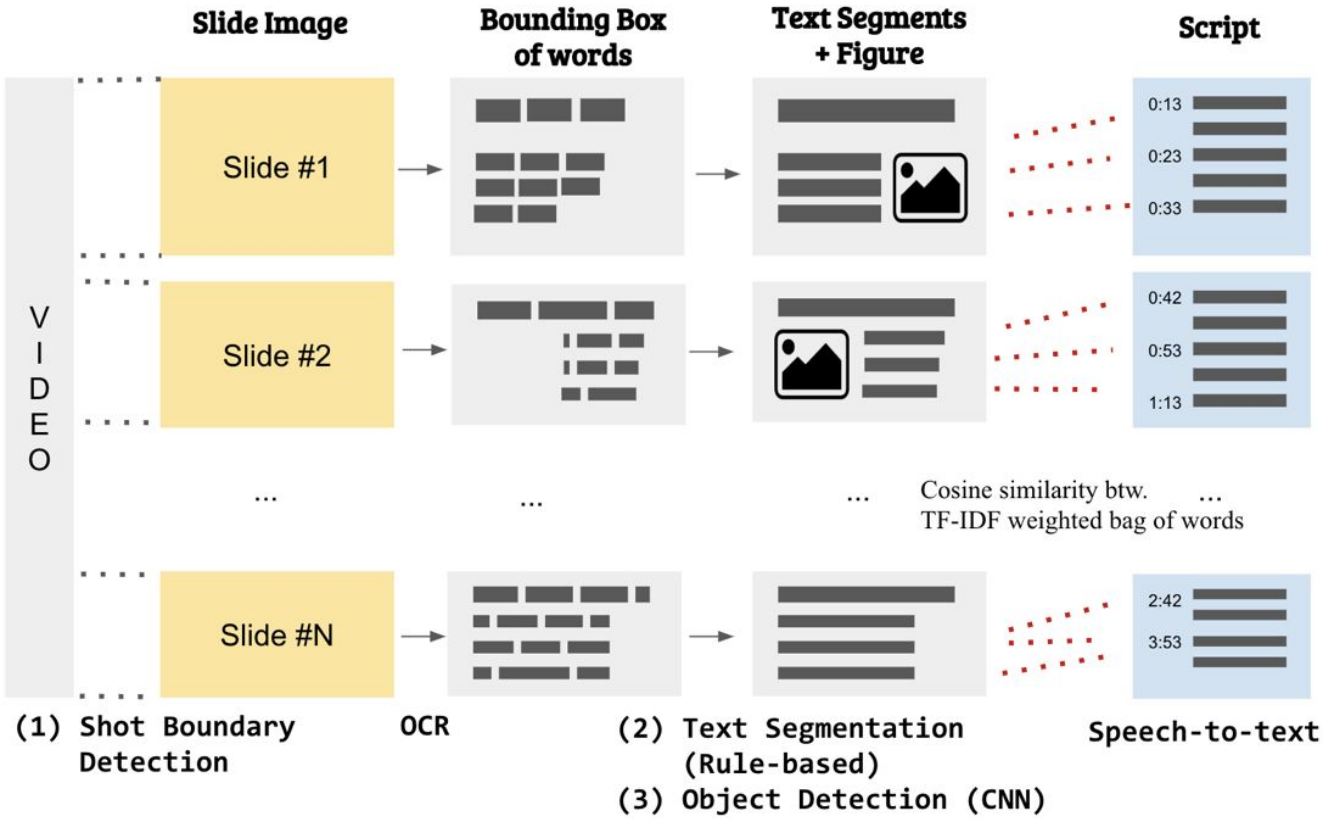
1. Automatic Highlighting
2. Item-based Navigation
3. In-video Bookmarking

Interaction techniques powered by links between slides and narration

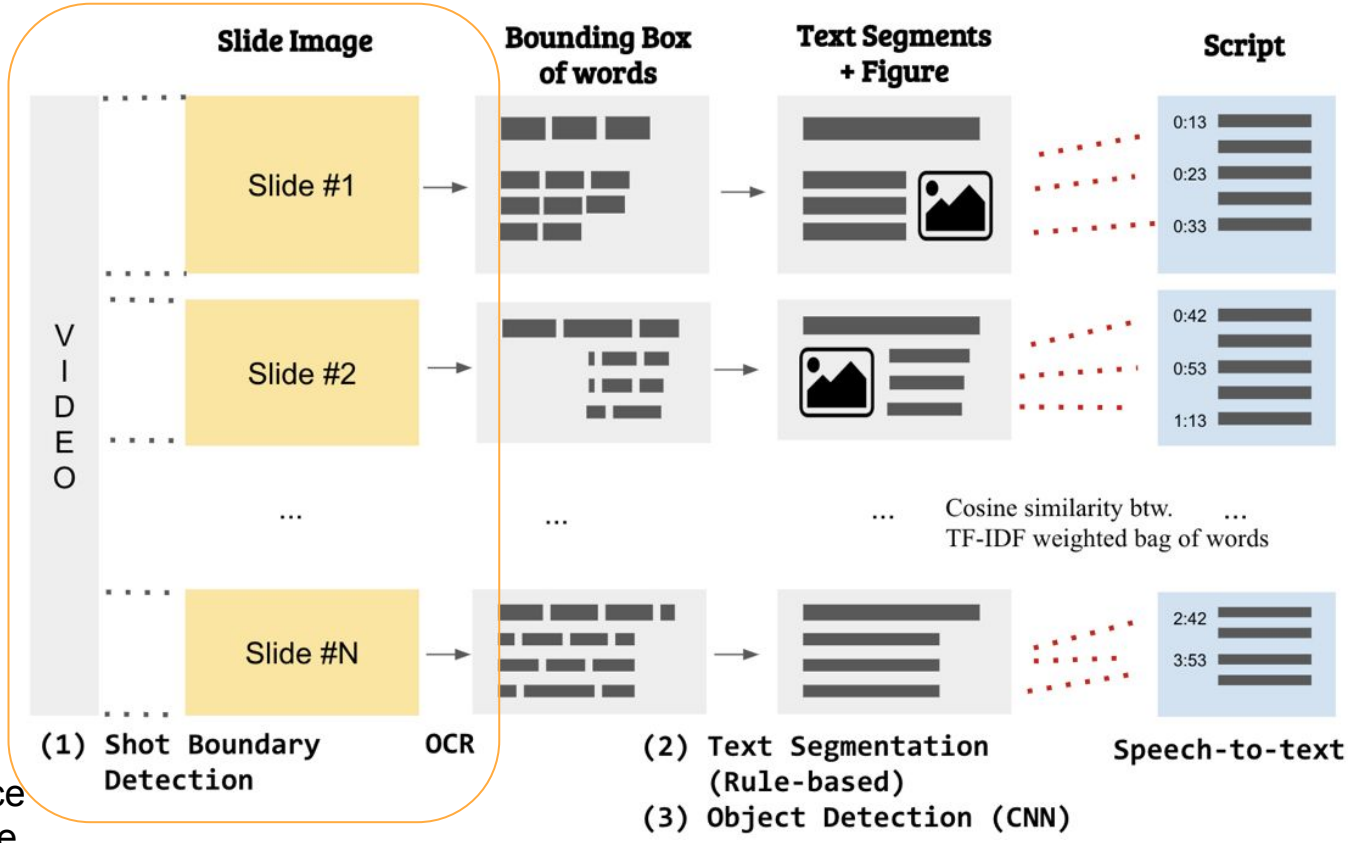


1. Automatic Highlighting
2. Item-based Navigation
3. In-video bookmarking

Computational Pipeline

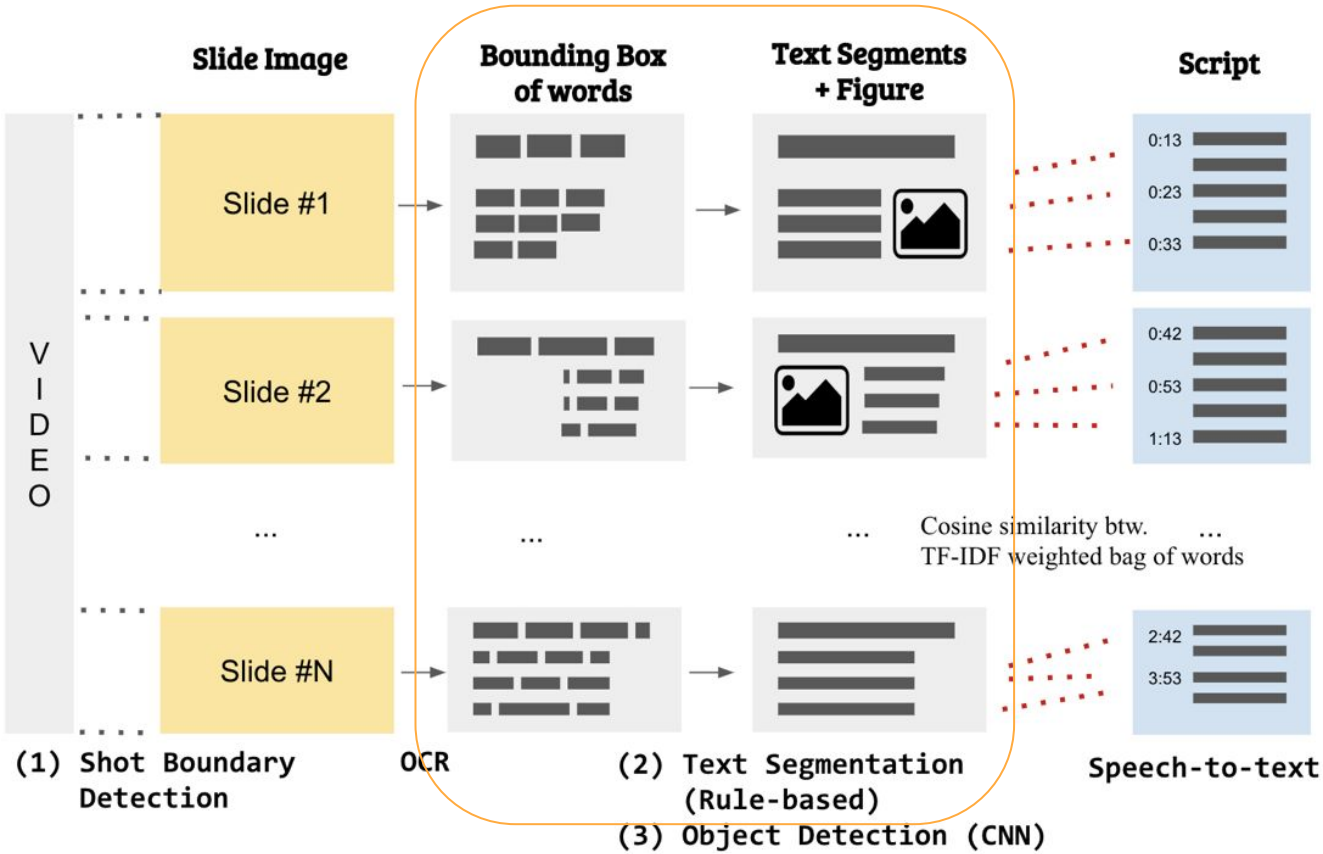


Computational Pipeline

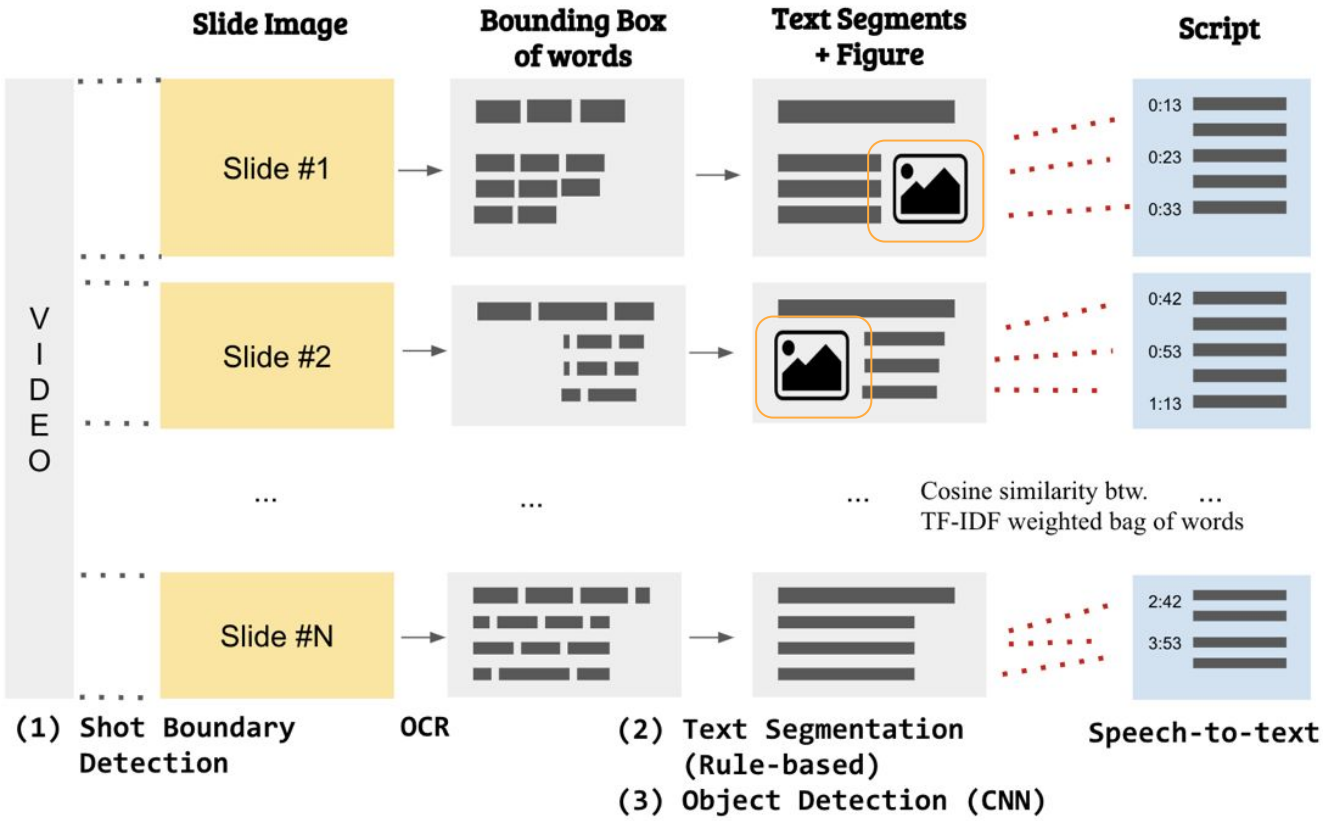


Pixel difference
Text difference

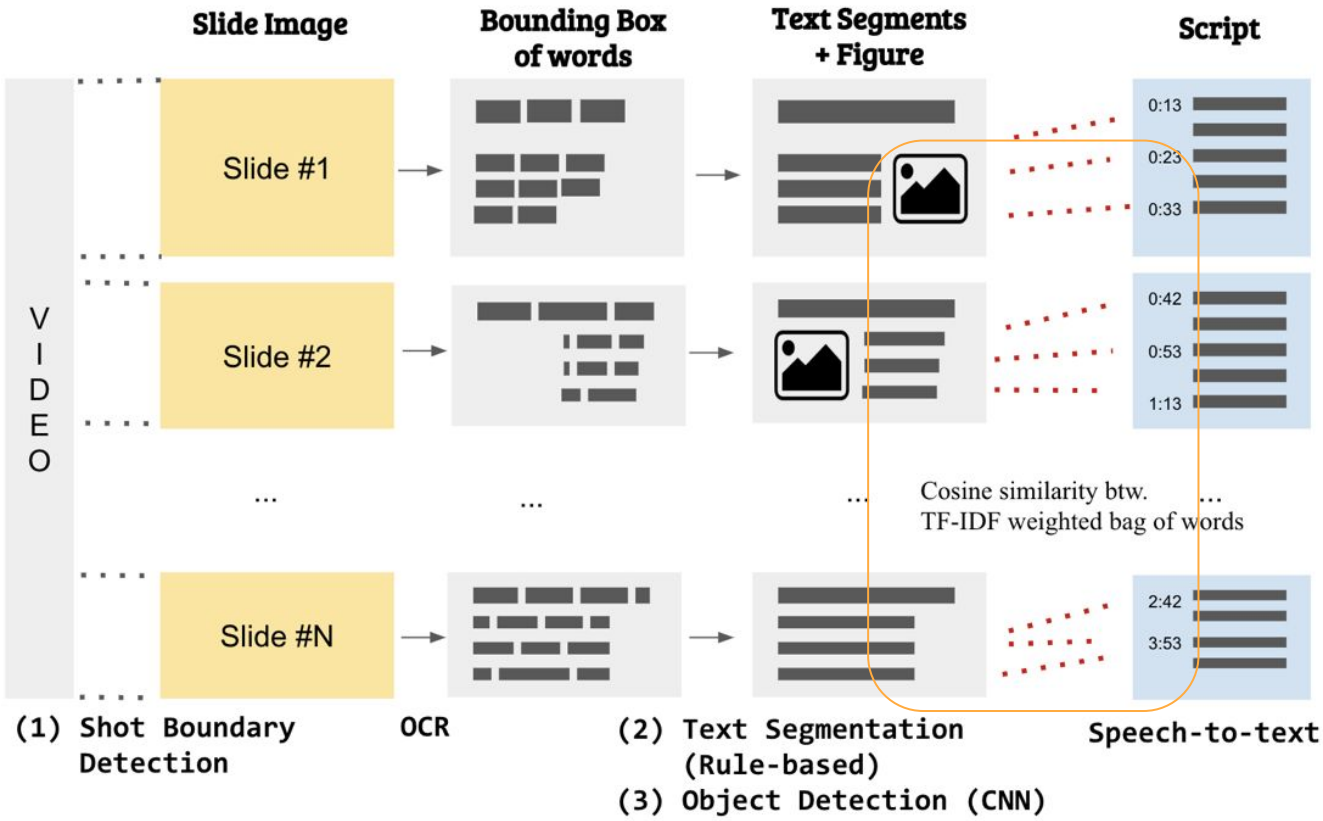
Computational Pipeline



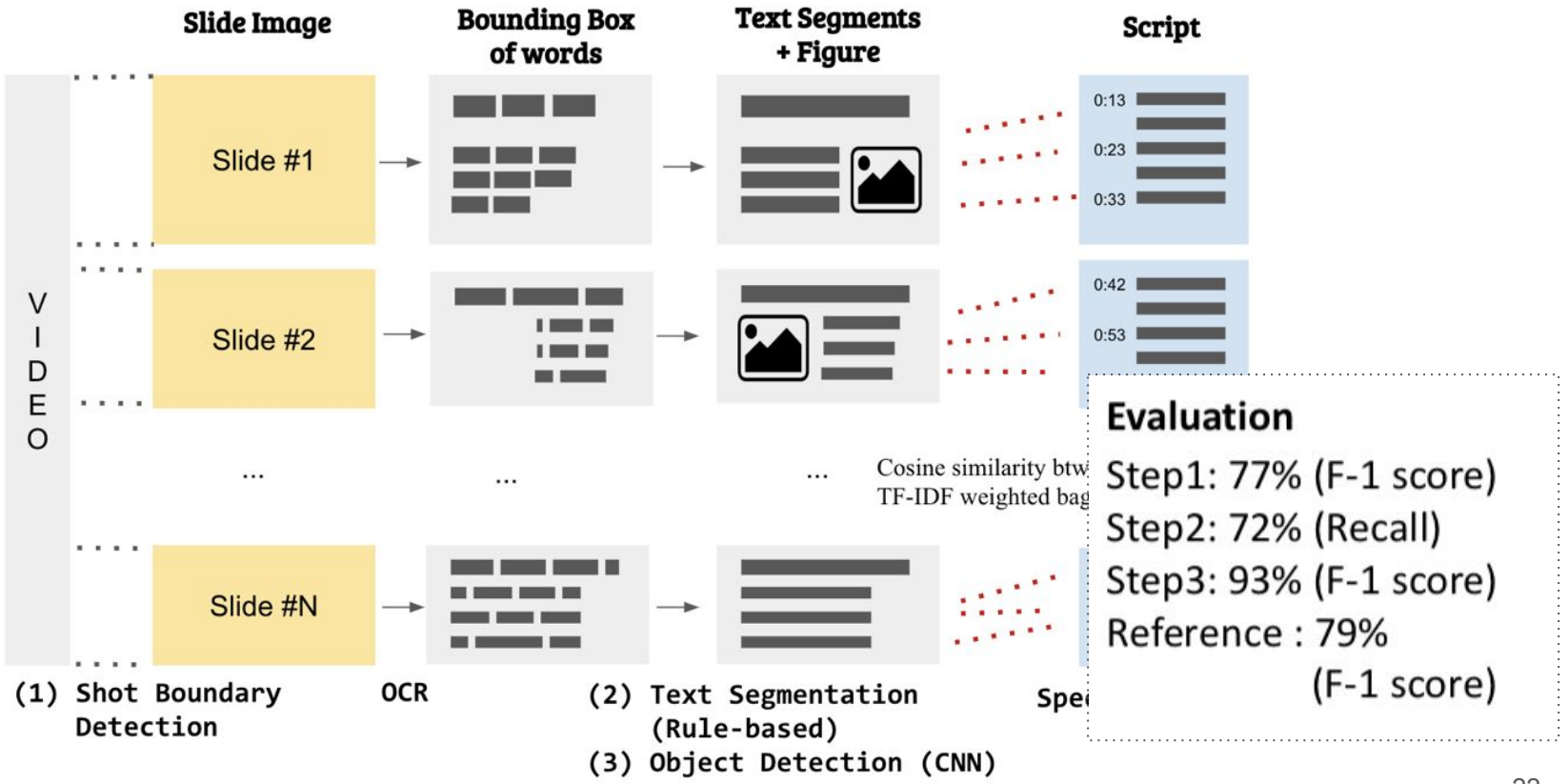
Computational Pipeline



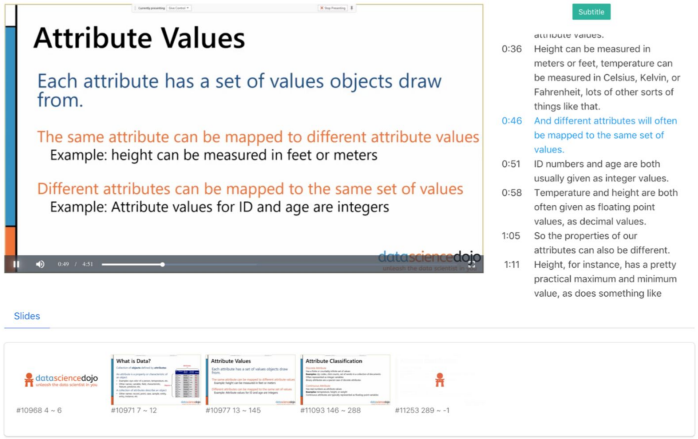
Computational Pipeline



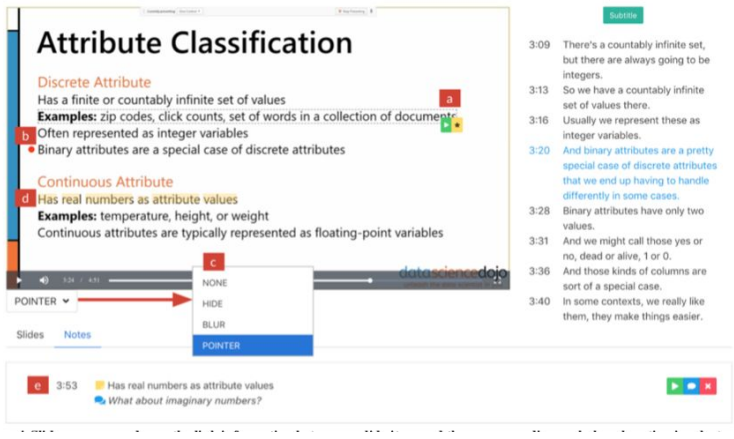
Computational Pipeline



User Study (n=12)



Baseline



DynamicSlide

1. Navigation -> Searching relevant part of the question
2. Cognitive load -> NASA-TLX

User Study (n=12)

Navigation task

(Time taken, seconds, shorter the better)

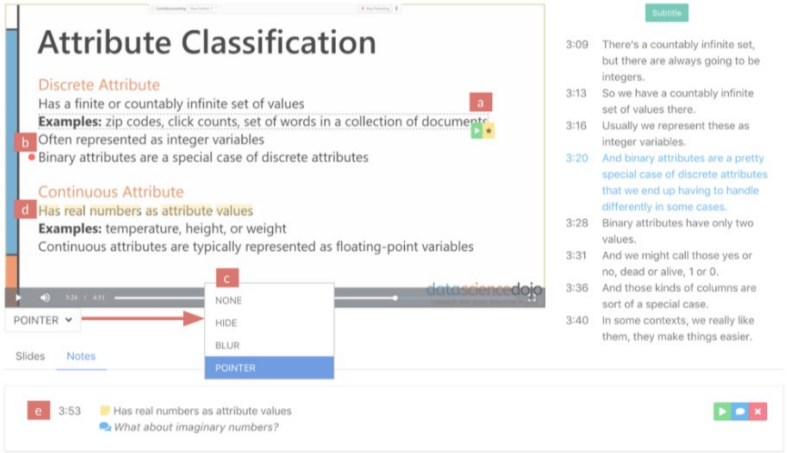
		Baseline	Ours
Answer is in the slide	Q1	26.3	15.9
	Q2	14.7	14.7
Answer is NOT in the slide	Q3	26.1	49.2
	Q4	29.5	16

Cognitive Load

(NASA-TLX, 1~5, smaller the better)

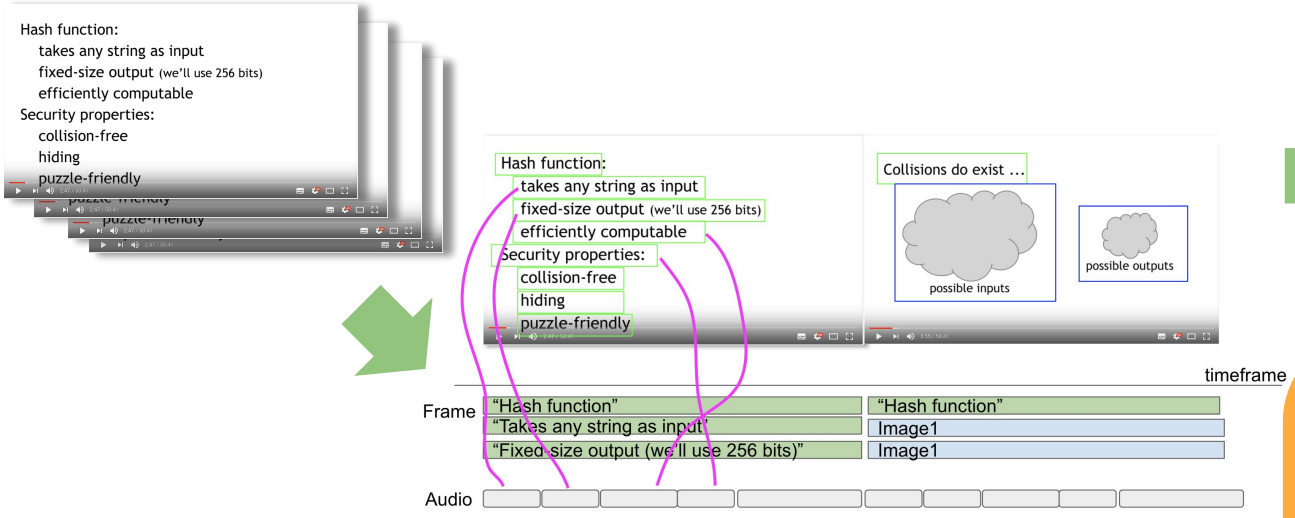
	Baseline	Ours
Overall	3.2	3.1
Mental	3.5	3.2
Performance	3.0	3.0
Effort	3.5	3.2
Frustration	3.1	2.6

Discussion

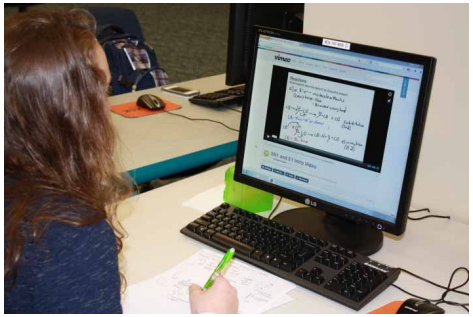


- How to robustly find link between slides and narrations?
- Do links commonly exist in lecture videos?
- The effect of finding and leveraging wrong link

Object-oriented tools for both learners & instructors



For Students (Player)



For Instructors
(Lightweight editor)

Lifespan of a lecture

t

Lifespan of a lecture

- Update the content of slide
- Try different versions of explanation

Lifespan of a lecture

- Update the content of slide
- Try different versions of explanation

- Hard to update the published video

Lifespan of a lecture

- Update the content of slide
- Try different versions of explanation

- Hard to update the published video

How can we build a tool to help instructors directly update their video without having to patch and re-publish?

Formative study

Interviews with

7 **Instructors** who made lecture video
(Professors + Part time)

3 **Video editors**
(Content team at the university)

Formative study

Interviews with

7 **Instructors** who made lecture video
(Professor + Part time)

3 **Video editors**
(Content team at the university)

1. Instructors want to edit the videos before and after publications
2. Editing cost is prohibitive
3. Editing often involves substituting the original (slide) contents

Formative study

Interviews with

7 **Instructors** who made lecture video
(Professor + Part time)

3 **Video editors**
(Content team at the university)

1. **Instructors want to edit the videos before and after publications**
2. Editing cost is prohibitive
3. Editing often involves substituting the original (slide) contents

Before
- Add supplementary content
- Highlight the text
- Cut part with mistakes

After
- Fix typo
- Fix figure

Formative study

Interviews with

7 **Instructors** who made lecture video
(Professor + Part time)

3 **Video editors**
(Content team at the university)

1. **Instructors want to edit the videos before and after publications**
2. Editing cost is prohibitive
3. Editing often involves substituting the original (slide) contents

Before
- Add supplementary content
- Highlight the text
- Cut part with mistakes

After
- Fix typo
- Fix figure

Formative study

Interviews with

7 **Instructors** who made lecture video
(Professor + Part time)

3 **Video editors**
(Content team at the university)

1. Instructors want to edit the videos before and after publications
2. **Editing cost is prohibitive**
3. Editing often involves substituting the original (slide) contents

Formative study

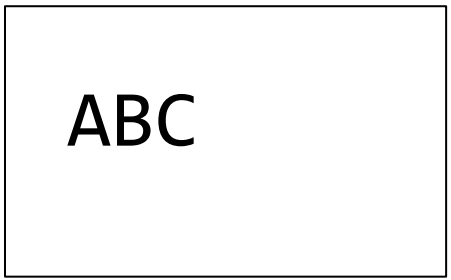
Interviews with

7 **Instructors** who made lecture video
(Professor + Part time)

3 **Video editors**
(Content team at the university)

1. Instructors want to edit the videos before and after publications
2. Editing cost is prohibitive
3. **Editing often involves substituting the original (slide) contents**

Changing text (ABC -> CBA)



Formative study

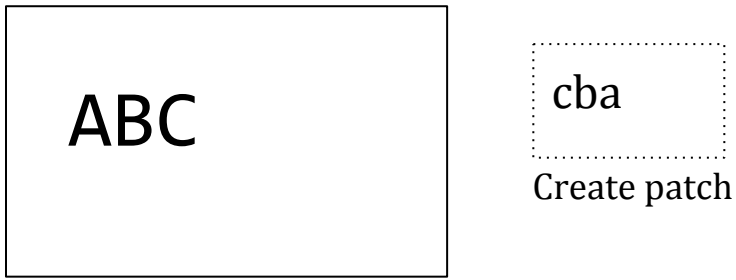
Interviews with

7 **Instructors** who made lecture video
(Professor + Part time)

3 **Video editors**
(Content team at the university)

1. Instructors want to edit the videos before and after publications
2. Editing cost is prohibitive
3. **Editing often involves substituting the original (slide) contents**

Changing text (ABC -> CBA)



Formative study

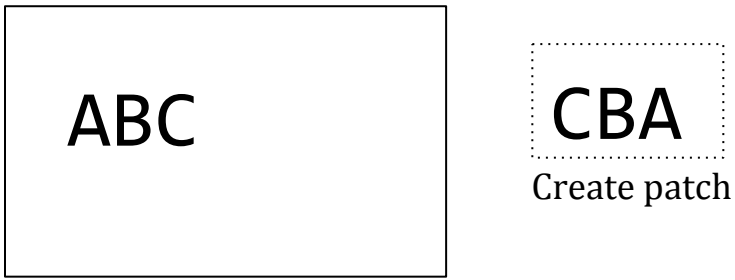
Interviews with

7 **Instructors** who made lecture video
(Professor + Part time)

3 **Video editors**
(Content team at the university)

1. Instructors want to edit the videos before and after publications
2. Editing cost is prohibitive
3. **Editing often involves substituting the original (slide) contents**

Changing text (ABC -> CBA)



Formative study

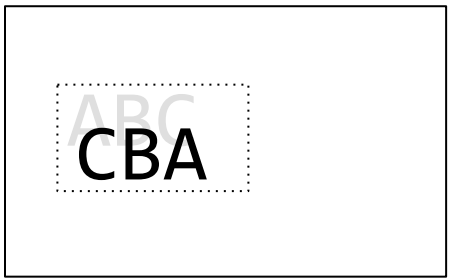
Interviews with

7 **Instructors** who made lecture video
(Professor + Part time)

3 **Video editors**
(Content team at the university)

1. Instructors want to edit the videos before and after publications
2. Editing cost is prohibitive
3. **Editing often involves substituting the original (slide) contents**

Changing text (ABC -> CBA)



Overwrite + Adjust timing

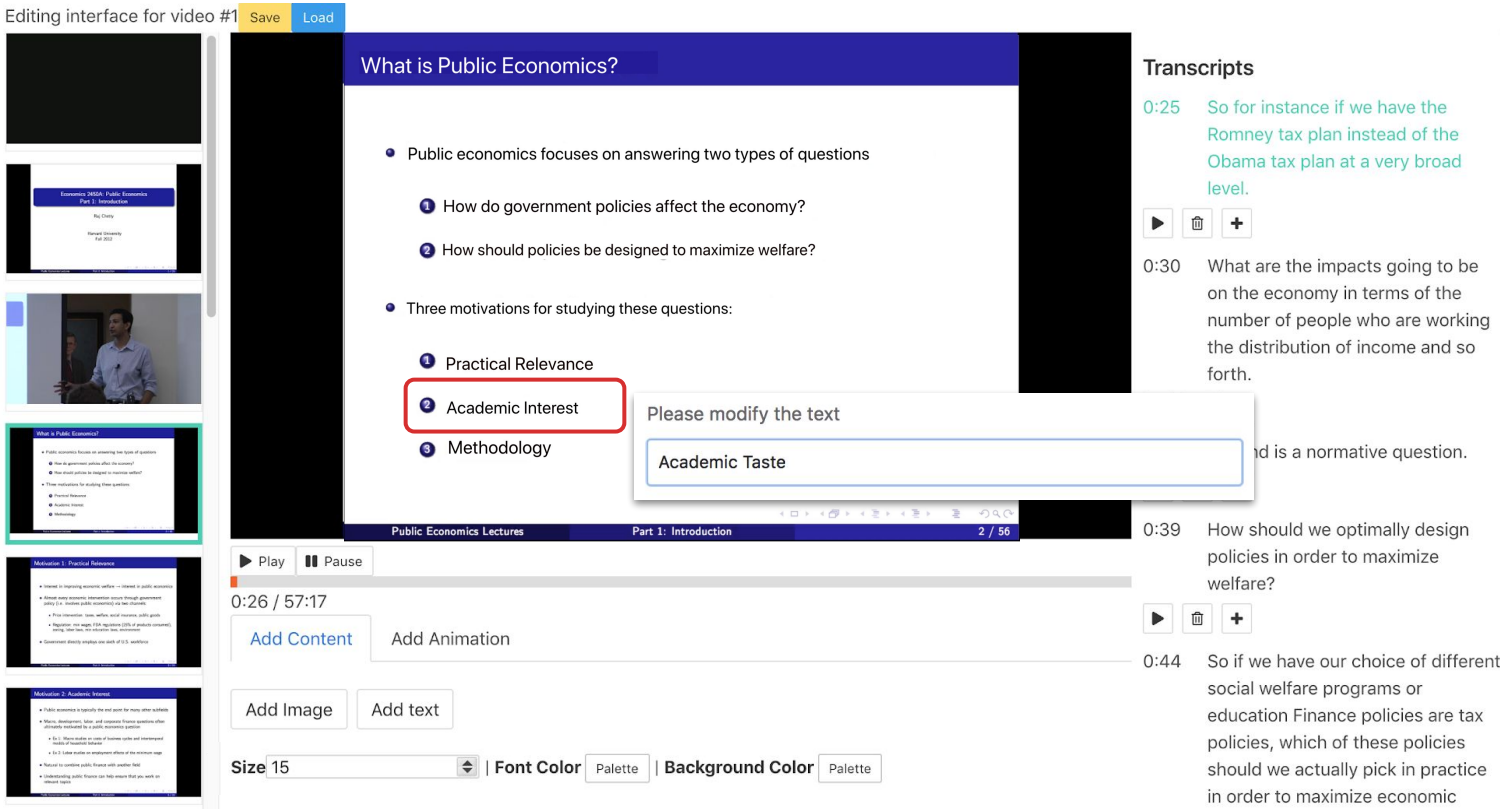
Design Goals

1. Provide an intuitive interface for instructors to directly edit their lecture videos

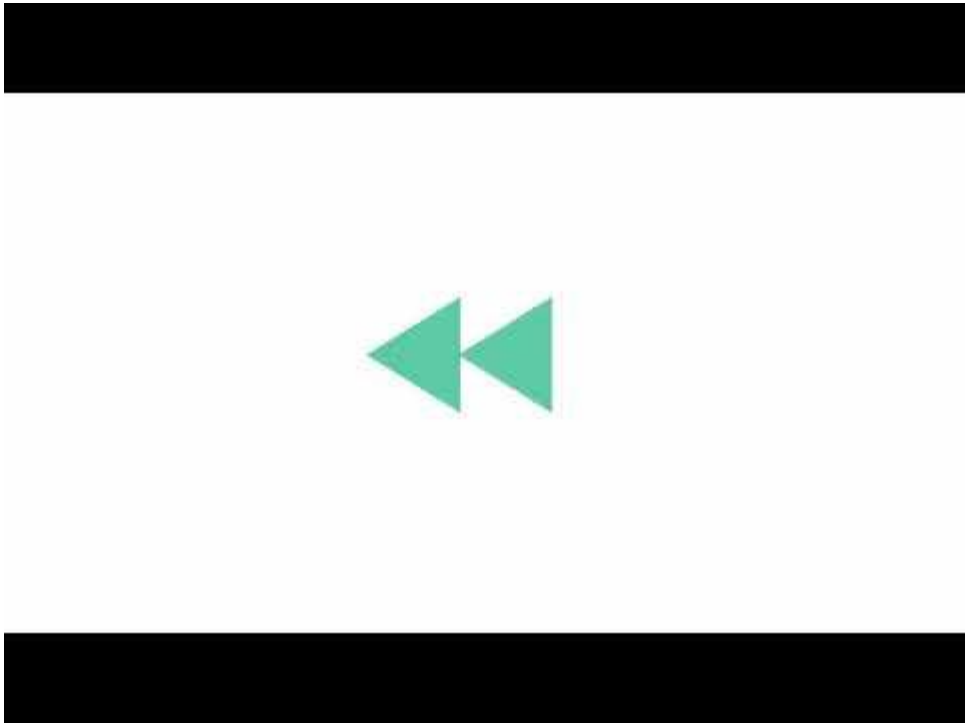
2. Enable users to update elements of the lecture slide within the video

3. Facilitate synchronization of visual and audio events in the video

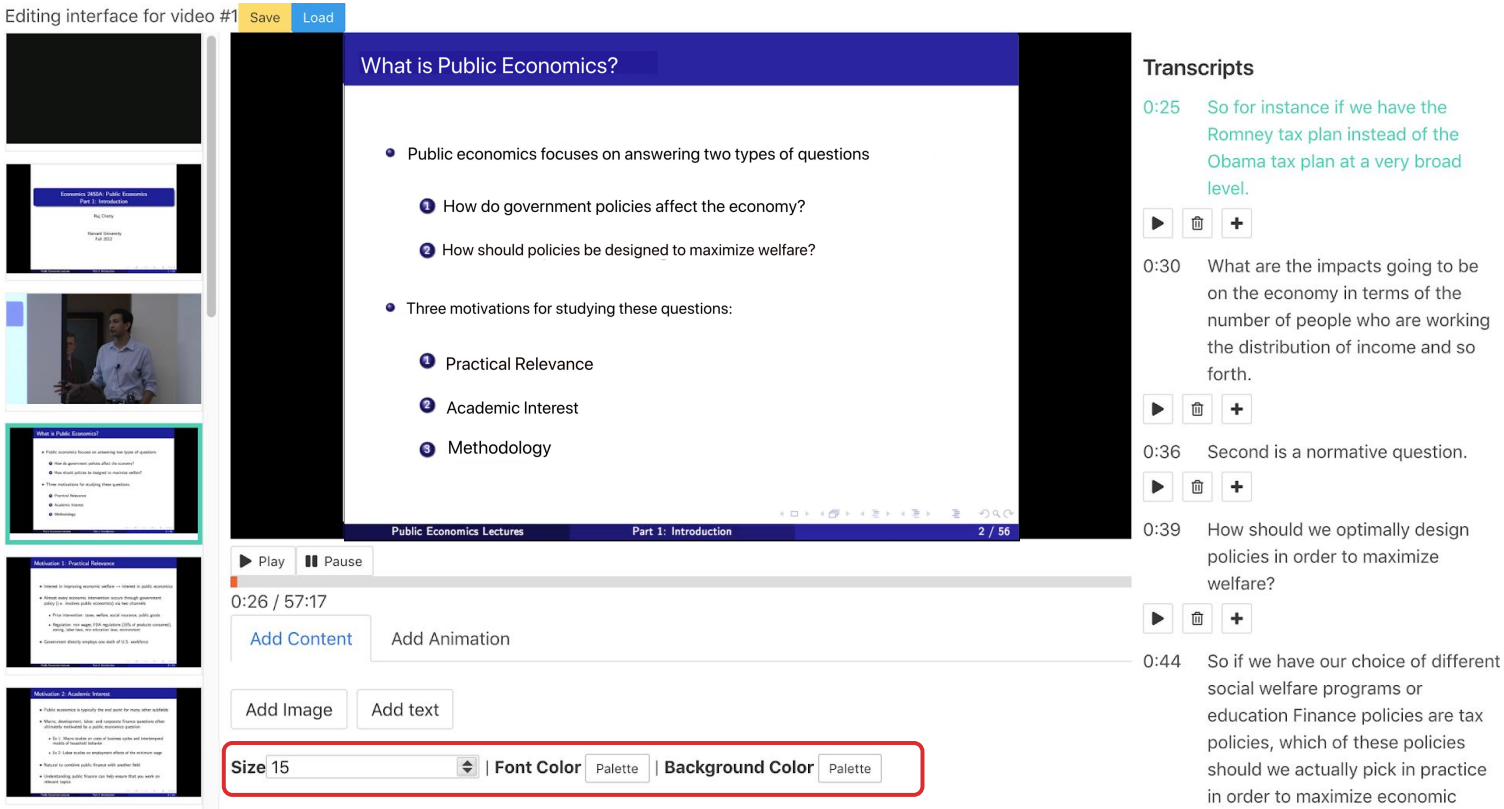
DynamicLecture



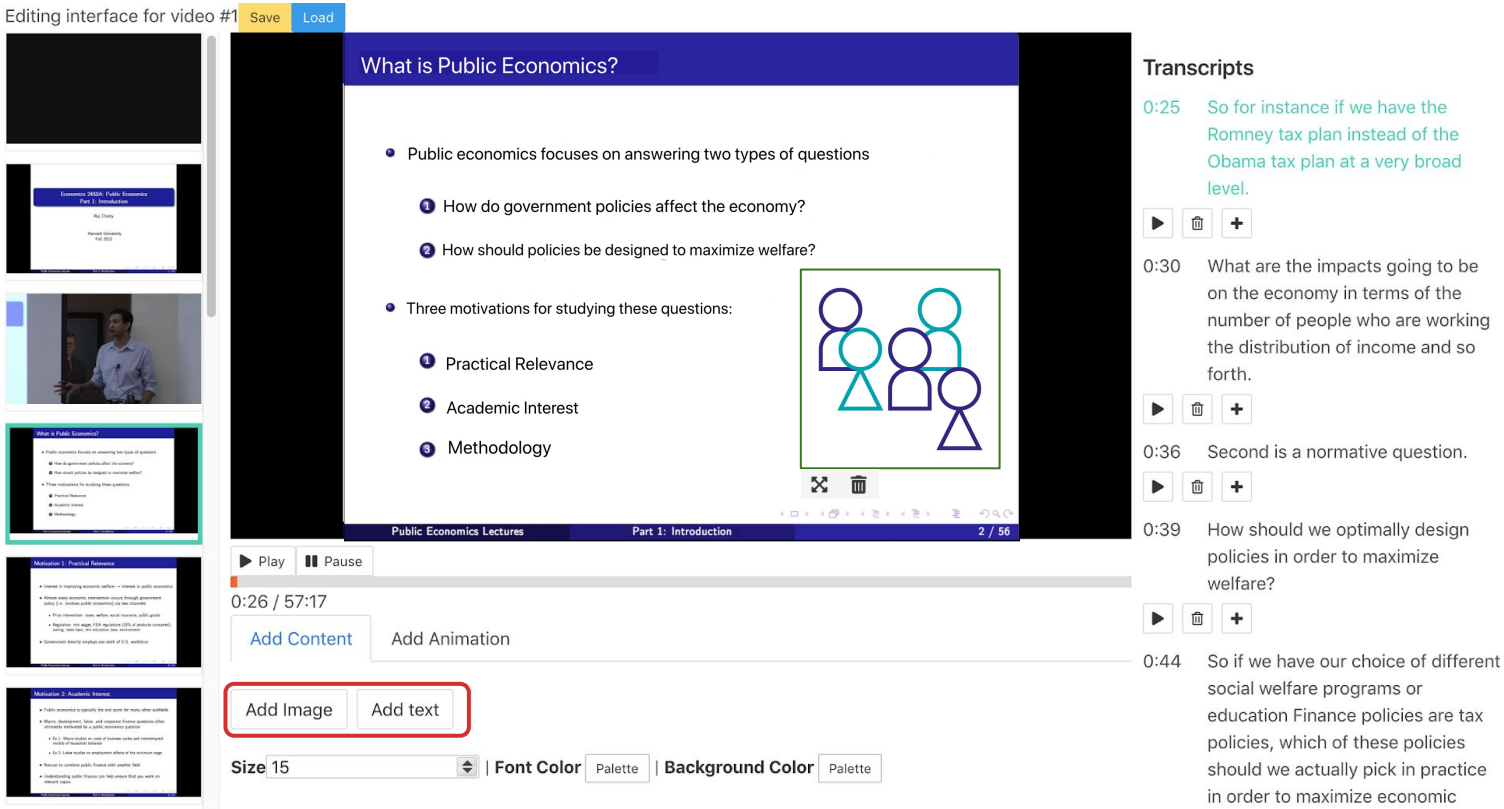
DynamicLecture



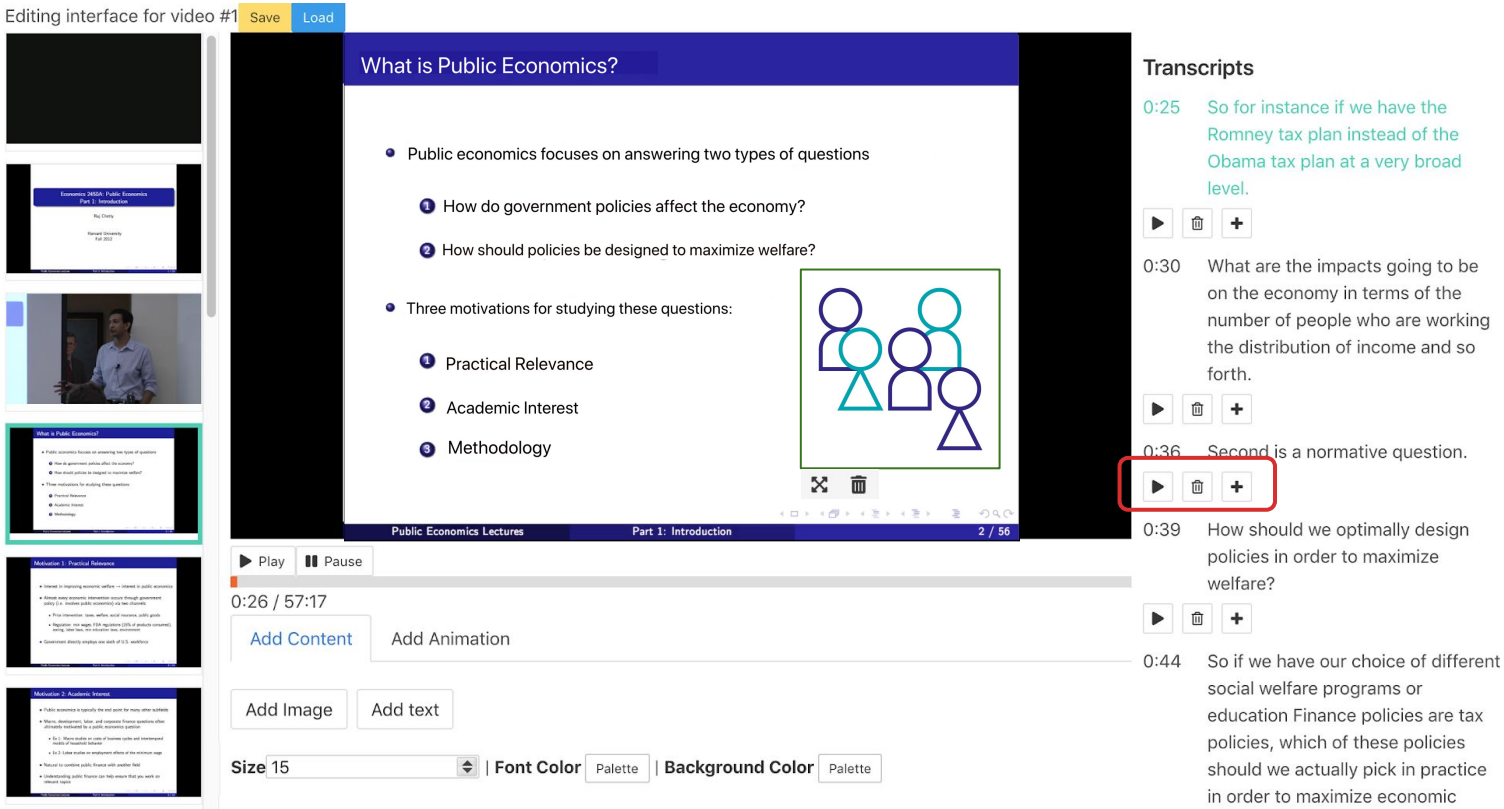
DynamicLecture



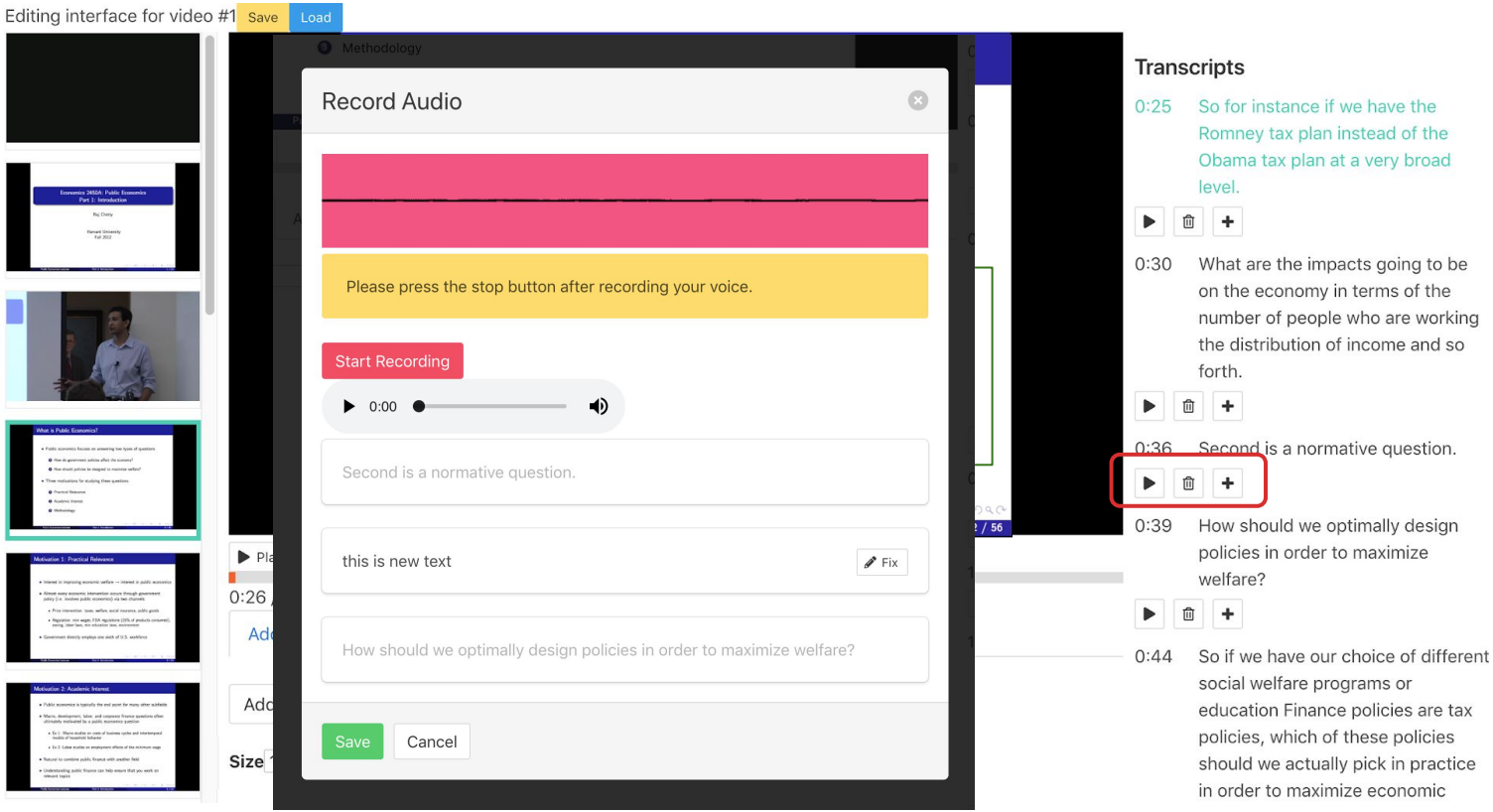
DynamicLecture



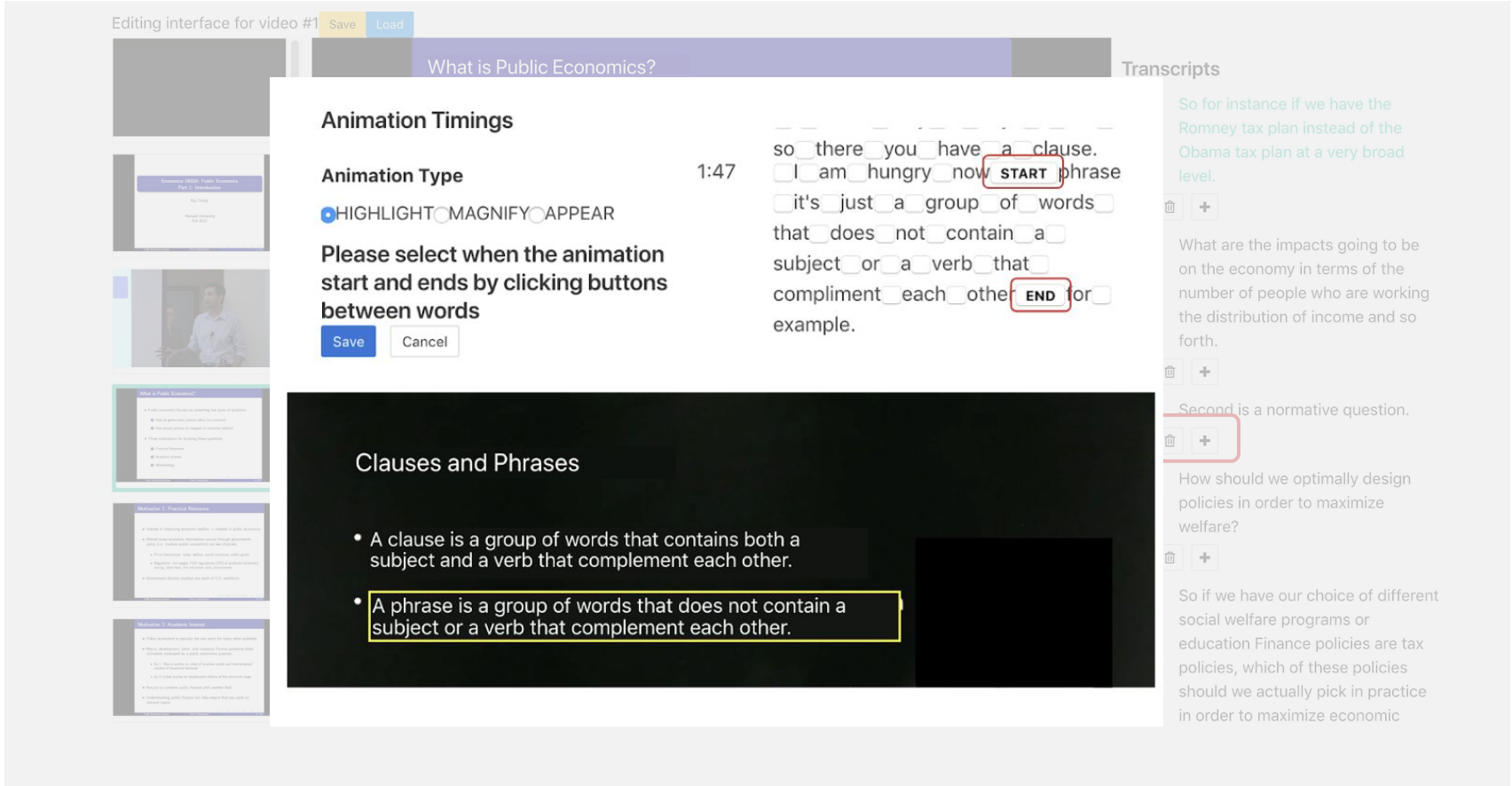
DynamicLecture



DynamicLecture



DynamicLecture

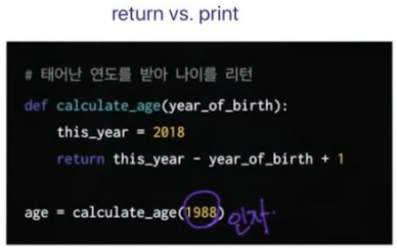


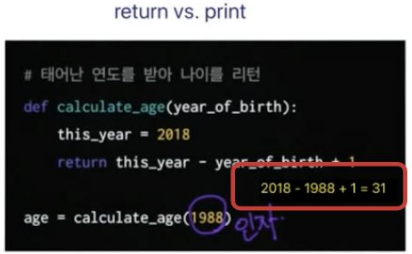
Evaluation

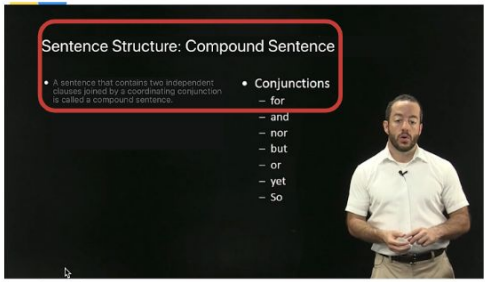
1. Instructors review their lecture videos, then list up improvement items
2. Inststructors fix their video using the system

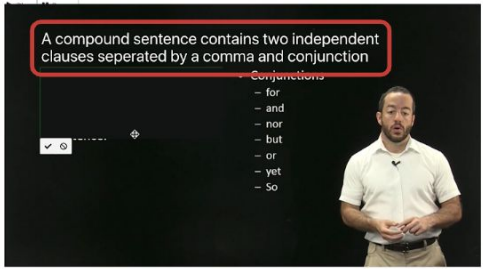
Result

1. Instructors review their lecture videos, then list up improvement items
2. Inststructors fix their video using the system

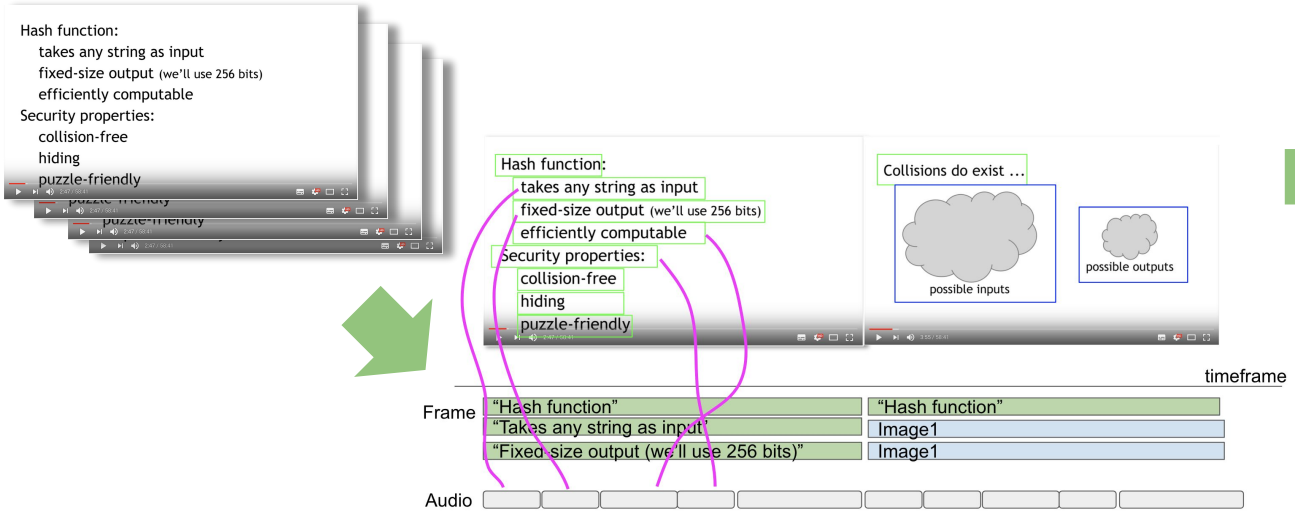




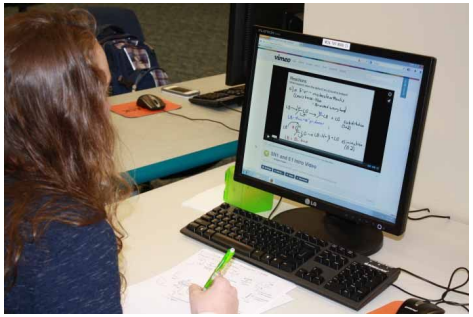




Conclusion



For Students (Player)



For Instructors
(Lightweight editor)

1. *Parsing* videos into objects can improve the way we interact, consume, & create lecture videos.
2. We need to explore novel object-based video interaction techniques. (e.g., participatory video improvement by learners)