

# Alati umjetne inteligencije za izradu video sadržaja

---

**Kušter, Valentina**

**Undergraduate thesis / Završni rad**

**2024**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Organization and Informatics / Sveučilište u Zagrebu, Fakultet organizacije i informatike**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:211:631770>

*Rights / Prava:* [Attribution 3.0 Unported](#)/[Imenovanje 3.0](#)

*Download date / Datum preuzimanja:* **2024-12-02**



*Repository / Repozitorij:*

[Faculty of Organization and Informatics - Digital Repository](#)



SVEUČILIŠTE U ZAGREBU  
FAKULTET ORGANIZACIJE I INFORMATIKE  
VARAŽDIN

Valentina Kušter

ALATI UMJETNE INTELIGENCIJE ZA  
IZRADU VIDEO SADRŽAJA

ZAVRŠNI RAD

Varaždin, 2024.

**SVEUČILIŠTE U ZAGREBU**  
**FAKULTET ORGANIZACIJE I INFORMATIKE**  
**V A R A Ž D I N**

**Valentina Kušter**

**JMBAG: 0016148020**

**Studij: Informacijski i poslovni sustavi**

**ALATI UMJETNE INTELIGENCIJE ZA IZRADU VIDEO SADRŽAJA**

**ZAVRŠNI RAD**

**Mentorica:**

Izv. prof. dr. sc. Dijana Plantak Vukovac

**Varaždin, rujan 2024.**

*Valentina Kušter*

### **Izjava o izvornosti**

Izjavljujem da je moj završni rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristila drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

*Autorica potvrdila prihvaćanjem odredbi u sustavu FOI-radovi*

---

## Sažetak

Ovaj završni rad bavi se alatima umjetne inteligencije za izradu video sadržaja. Prije opisa i predstavljanja alata umjetne inteligencije za izradu video sadržaja, bilo je potrebno upoznati se s pojmom umjetne inteligencije i njezinim tehnikama. U radu je posebno istaknuta generativna umjetna inteligencija jer se upravo na nju oslanjaju alati umjetne inteligencije za izradu video sadržaja. Opisane su vrste generativne umjetne inteligencije i način na koji alati generiraju video sadržaj. Nakon toga, objašnjena je primjena umjetne inteligencije u pet industrijskih sektora: marketingu, obrazovanju, zdravstvu te filmskoj industriji i industriji vizualnih efekata (VFX). Za svaki od ovih sektora objašnjeno je kako se umjetna inteligencija koristi u kontekstu izrade video sadržaja. Nakon upoznavanja s načinom rada generativne umjetne inteligencije i njezinom primjenom, navedeni su i alati umjetne inteligencije za izradu video sadržaja. Za odabrane alate opisane su njihove osnovne karakteristike i navedene su prednosti i nedostaci ovih alata. U alatu Visla demonstrirana je izrada video sadržaja na različite načine, kao i uređivanje generiranih video sadržaja. Na kraju je provedena analiza učinkovitosti alata Visla i doneseni su zaključci o tome kako umjetna inteligencija unaprjeđuje proces izrade video sadržaja.

**Ključne riječi:** generativna umjetna inteligencija; alati umjetne inteligencije; video sadržaj; tekst-u-video; primjena umjetne inteligencije u izradi video sadržaja; Visla

# Sadržaj

Sadržaj.....	iii
1. Uvod.....	1
2. Umjetna inteligencija.....	2
2.1. Tehnike umjetne inteligencije.....	3
2.1.1. Strojno učenje.....	3
2.1.2. Duboko učenje.....	4
2.1.3. Neuronske mreže.....	4
2.1.4. Obrada prirodnog jezika.....	5
2.1.5. Računalni vid.....	6
3. Generativna umjetna inteligencija.....	8
3.1. Vrste generativnih modela.....	8
3.2. Primjene generativne umjetne inteligencije.....	10
3.3. Vrste generativne umjetne inteligencije.....	10
3.4. Generiranje video sadržaja pomoću umjetne inteligencije.....	11
4. Primjena umjetne inteligencije u procesu izrade video sadržaja.....	15
4.1. Primjena umjetne inteligencije u procesu izrade video sadržaja u marketingu.....	15
4.2. Primjena umjetne inteligencije u procesu izrade video sadržaja u obrazovanju.....	18
4.3. Primjena umjetne inteligencije u procesu izrade video sadržaja u zdravstvu.....	20
4.4. Primjena umjetne inteligencije u procesu izrade video sadržaja u filmskoj industriji i industriji vizualnih efekata (VFX).....	23
5. Alati umjetne inteligencije za izradu video sadržaja.....	25
5.1. InVideo AI.....	25
5.2. Lumen5.....	28
5.3. Visla.....	31
5.4. Runway.....	33
5.5. Sora.....	36
6. Izrada video sadržaja u alatu Visla.....	39
6.1. Korisničko sučelje alata Visla.....	39
6.2. Izrada video sadržaja na temelju ideje.....	42
6.3. Izrada video sadržaja na temelju teksta.....	45
6.4. Izrada video sadržaja na temelju URL-a web stranice.....	47
6.5. Izrada video sadržaja na temelju audiozapisa.....	49
6.6. Izrada video sadržaja na temelju slike.....	51

6.7. Uređivanje video sadržaja .....	53
6.8. Analiza učinkovitosti.....	58
7. Zaključak .....	60
Popis literature.....	61
Popis slika .....	65
Popis tablica .....	67
Prilozi .....	68

# 1. Uvod

Generativna umjetna inteligencija sve se više razvija u tehnološkom svijetu i postaje sve popularnija. Zahvaljujući ovoj vrsti umjetne inteligencije, svatko može kreirati zabavne sadržaje poput slika, tekstualnog sadržaja ili videozapisa, i to samo na temelju jednostavnog upita. Samim time, raste i popularnost alata umjetne inteligencije koji se koriste za stvaranje različitih vrsta sadržaja, a u ovom radu naglasak će biti na alatima umjetne inteligencije namijenjenim za izradu video sadržaja. Motivacija za odabir ove teme završnog rada bila je brzo i jednostavno stvaranje video sadržaja. Dovoljno je unijeti kratak tekst kao upit i umjetna inteligencija će na temelju tog teksta u vrlo kratko vrijeme generirati video sadržaj za čije stvaranje bi inače bili potrebni sati.

Svrha ovog rada je predstaviti primjenu umjetne inteligencije u izradi video sadržaja kroz opis načina na koji se generira video sadržaj pomoću umjetne inteligencije, prikaz konkretnih primjera video sadržaja i pregled te analizu dostupnih alata koji se koriste za stvaranje video sadržaja. U radu će najprije biti općenito opisana umjetna inteligencija, nakon čega će detaljnije biti objašnjena generativna umjetna inteligencija. Nakon toga, bit će navedeni konkretni primjeri primjene umjetne inteligencije u procesu izrade video sadržaja za nekoliko industrijskih sektora. Detaljnije će se opisati i analizirati pet alata umjetne inteligencije za izradu video sadržaja: InVideo AI, Lumen5, Visla, Runway i Sora. U alatu Visla bit će generirano nekoliko video sadržaja korištenjem različitih upita, a bit će demonstriran i proces uređivanja video sadržaja. Na kraju će se analizirati učinkovitost korištenog alata i donijet će se zaključci o tome kako se umjetna inteligencija može koristiti za poboljšanje procesa izrade video sadržaja.



## 2. Umjetna inteligencija

Kako bismo razumjeli način na koji alati umjetne inteligencije generiraju video sadržaje, prvo se moramo upoznati sa samim pojmom umjetne inteligencije. Shi (2020, str. 1) ističe usredotočenost umjetne inteligencije na stvaranje inteligentnih strojeva koji se ponašaju kao ljudi i navodi da se umjetna inteligencija „obično definira kao znanost i inženjerstvo oponašanja, proširenja i povećanja ljudske inteligencije pomoću umjetnih sredstava i tehnika kako bi se stvorili inteligentni strojevi.“ Patel (2023, str. 6) također kao glavni cilj umjetne inteligencije naglašava stvaranje inteligentnih strojeva, odnosno „inteligentnih agenata“ i definira ih kao „bilo koji sustav koji percipira svoju okolinu putem senzora i u interakciji je s njom putem aktuatora, vođen određenim ciljevima.“ Ovakvi sustavi namijenjeni su za interpretaciju podataka, učenje iz iskustava i donošenje odluka poput ljudi. Russell i Norvig (2020) daju čak osam definicija umjetne inteligencije koje raspoređuju u dvije dimenzije. Definicije su raspoređene tako da se one definicije koje se nalaze na vrhu odnose na misaone procese i rasuđivanje, a one na dnu na ponašanje. One definicije koje se nalaze s lijeve strane pritom mjere uspješnost ljudske izvedbe, dok one s desne strane mjere racionalnost. Navedene definicije možemo vidjeti u tablici 1 u nastavku.

Tablica 1: Definicije umjetne inteligencije u dvije dimenzije

<b>Ljudsko razmišljanje</b> „Uzbudljivi novi pokušaj tjeranja računala da misle... <i>strojevi s umovima</i> , u punom i doslovnom smislu.“ (Haugeland, 1985)  „[Automatizacija] aktivnosti koje povezujemo s ljudskim razmišljanjem, aktivnostima kao što su donošenje odluka, rješavanje problema, učenje...“ (Bellman, 1978)	<b>Racionalno razmišljanje</b> „Proučavanje mentalnih sposobnosti kroz korištenje računalnih modela.“ (Charniak and McDermott, 1985)  „Proučavanje izračuna koji omogućuju opažanje, rasuđivanje i djelovanje.“ (Winston, 1992)
<b>Ljudsko djelovanje</b> „Umijeće stvaranja strojeva koji obavljaju funkcije koje zahtijevaju inteligenciju kada ih izvode ljudi.“ (Kurzweil, 1990)  „Proučavanje kako natjerati računala da rade stvari na kojima ljudi trenutno rade bolje.“ (Rich and Knight, 1991)	<b>Racionalno djelovanje</b> "Računalna inteligencija je proučavanje dizajna inteligentnih agenata." (Poole i sur., 1998.)  „AI...se bavi inteligentnim ponašanjem u artefaktima.“ (Nilsson, 1998)

Izvor: (Russell i Norvig, 2020, str. 2.)

Na temelju svih navedenih definicija možemo zaključiti da je umjetna inteligencija grana računalne znanosti čiji je cilj stvaranje inteligentnih agenata koji obavljaju aktivnosti za koje je potrebna ljudska inteligencija. To je vrlo složen pojam koji obuhvaća različite tehnike, a najvažnije tehnike bit će opisane u nastavku.

## 2.1. Tehnike umjetne inteligencije

Način rada umjetne inteligencije može se opisati kroz sljedeće tehnike:

- strojno učenje
- duboko učenje
- neuronske mreže
- obrada prirodnog jezika
- računalni vid.

### 2.1.1. Strojno učenje

Strojno učenje (*eng. Machine Learning - ML*) je grana umjetne inteligencije koja omogućuje umjetnoj inteligenciji da oponaša način na koji ljudi uče pri čemu koristi podatke i algoritme koji s vremenom postaju sve točniji (IBM, bez dat.). Detaljniju definiciju strojnog učenja daje (Kanade, 2022) koji naglašava sposobnost strojnog učenja da omogući strojevima da identificiraju uzorke i rade predviđanja na temelju onog što nauče iz podataka i prošlih iskustava uz minimalnu ljudsku intervenciju. Kako bi strojno učenje moglo identificirati uzorke i obavljati predviđanja, ono koristi algoritme koji izvlače važne informacije iz velikih količina podataka. Ti algoritmi se poboljšavaju kako raste broj dostupnih uzoraka.

Sustav učenja, koji se odnosi na algoritme strojnog učenja, može se podijeliti na tri dijela: proces odlučivanja, funkcija pogreške i proces optimizacije modela (IBM, bez dat. prema UC Berkeley). **Proces odlučivanja** odnosi se na procjenu o uzorku u podacima koju algoritam strojnog učenja radi na temelju neoznačenih ili označenih ulaznih podataka. **Funkcija pogreške** radi usporedbu za procjenu točnosti modela na temelju poznatih primjera. Na kraju, **proces optimizacije modela** odnosi se na smanjivanje razlika između poznatog primjera i procjene modela prilagođavanjem težina.

Postoji nekoliko tipova strojnog učenja: nadzirano strojno učenje, nenadzirano strojno učenje i polu-nadzirano strojno učenje. **Nadzirano strojno učenje** obuhvaća obuku strojeva za klasificiranje i predviđanje rezultata na temelju označenih skupova podataka. Model se trenira s ulazom i odgovarajućim izlazom te prilagođava svoje težine kako se unose ulazni podaci da bi se prilagodio odgovarajućem izlazu. Za razliku od nadziranog strojnog učenja koje koristi označene skupove podataka, **nenadzirano strojno učenje** odnosi se na grupiranje neoznačenih skupova podataka i omogućuje strojevima predviđanje bez nadzora, odnosno bez ljudske intervencije. Algoritmi strojnog učenja grupiraju neoznačene skupove podataka na temelju sličnosti i razlika u informacijama. **Polu-nadzirano strojno učenje** je kombinacija

nadziranog i nenadziranog učenja koje koristi manji označeni i veći neoznačeni skup podataka za obuku svojih algoritama.

### 2.1.2. Duboko učenje

Prema (de Albuquerque, Raj, Yadav, 2024, str. 5) „duboko učenje (*eng. Deep Learning*) je podskup strojnog učenja koji računalima omogućuje učenje iz složenih podataka.“ Glavna komponenta dubokog učenja je korištenje neuronskih mreža. Kako bi pomogli strojevima da uče učinkovitije, algoritmi strojnog učenja koriste različite metode kao što su konvolucijske neuronske mreže, rekurentne neuronske mreže i generativne suparničke mreže. Algoritmi dubokog učenja sastoje se od umjetnih neuronskih mreža prema strukturi ljudskog mozga. Strukturu algoritma dubokog učenja čine neuroni koji djeluju jedni s drugima i s izvornim ulaznim podacima. Svaki neuron prenosi podatke sljedećem neuronu u mreži. Model generalizira korisne uzorke u podacima tako da se optimiziraju težine veza između neurona.

Gillis (2023) opisuje različite metode dubokog učenja: pad stope učenja, prijenos učenja, obuku od nule i odustajanje. Stopa učenja je hiperparametar koji kontrolira koliko promjena doživljava model kao odgovor na procijenjenu grešku. Ako su stope učenja previsoke, to može dovesti do nestabilnih procesa obuke ili učenja skupa težina koji nije optimalan. S druge strane, ako su stope učenja premale, to može dovesti do dugotrajnog procesa obuke za koji postoji opasnost da će zapeti. Kako bi se povećale performanse i smanjilo vrijeme obuke, provodi se prilagođavanje stope učenja pomoću metode koja se naziva **pad stope učenja** ili adaptivna stopa učenja. Sljedeća metoda je **prijenos učenja** koja obuhvaća usavršavanje prethodno uvježbanog modela na način da se postojećoj mreži dodaju novi podaci s nepoznatim klasifikacijama nakon čega se izvode prilagodbe na mreži kako bi se novi podaci mogli izvoditi sa specifičnijim sposobnostima kategorizacije. Metoda **obuke od nule** korisna je za nove aplikacije, a zahtijeva veliki označeni skup podataka i konfiguriranje mrežne arhitekture zbog čega se rjeđe koristi jer ovakva obuka može trajati duže vremena. **Odustajanje** je metoda koja rješava problem prekomjernog opremanja u mrežama s velikom količinom parametara, a poboljšava izvedbu neuronskih mreža u području prepoznavanja govora, klasifikacije dokumenata i računalne biologije.

### 2.1.3. Neuronske mreže

Sljedeća tehnika umjetne inteligencije su neuronske mreže (*eng. Neural networks*). Macukow (2016) opisuje **umjetne neuronske mreže** kao sustave koji se sastoje od jednostavnih elemenata za obradu informacija ili signala. Ti elementi su međusobno povezani i surađuju kako bi riješili željeni računalni zadatak. Umjetne neuronske mreže funkcioniraju na način kao što funkcioniraju biološki živčani sustavi, kao što je ljudski mozak. Sastoje se od

umjetnih neurona ili procesnih elemenata koji su organizirani u slojeve i čine neuronsku strukturu, a svaki neuron ima ponderirane ulaze, prijenosnu funkciju i jedan izlaz. Ponašanje neuronske mreže ovisi o prijenosnim funkcijama neurona, kao i o pravilu učenja i arhitekturi.

SAS (bez dat.) objašnjava način na koji rade neuronske mreže. Jednostavna neuronska mreža sastoji se od tri sloja: ulazni sloj, izlazni sloj i skriveni sloj koji se nalazi između ulaznog i izlaznog sloja. Ti slojevi povezani su preko čvorova i tvore neuronsku mrežu. Čvorovi se aktiviraju kada postoji dovoljno podražaja ili unosa što se zatim širi cijelom neuronskom mrežom pri čemu se stvara odgovarajući izlaz. Veze između neurona funkcioniraju kao sinapse i omogućuju prijenos signala s jednog neurona na drugi pri čemu se signali obrađuju tijekom prijenosa. Kada dobiju zadatak za rješavanje, neuroni putem izračuna provjeravaju ima li dovoljno informacija za prosljeđivanje sljedećem neuronu. Ako je zbroj primljenih podataka veći od granične vrijednosti, neuron se aktivira i također aktivira i ostale neurone s kojima je povezan. S povećanjem broja skrivenih slojeva unutar neuronske mreže, nastaju **duboke neuronske mreže**. Duboke neuronske mreže omogućuju strojno učenje koje može osposobiti računala da oponašaju zadatke koje obavljaju ljudi i omogućiti im da samostalno uče.

Na temelju svega navedenog, možemo dati detaljniji opis funkcioniranja neuronskih mreža. Prvo se podaci unose u neuronsku mrežu kroz ulazni sloj i zatim se obrađuju u skrivenim slojevima kroz sustav ponderiranih veza. U skrivenom sloju čvorovi dodjeljuju odgovarajuće težine podacima ulaznog sloja čiji umnošci se zatim zbrajaju, a zbroj se uspoređuje s određenom graničnom vrijednosti. Na kraju se skriveni slojevi povezuju s izlaznim slojem i dohvaćaju se izlazi.

#### 2.1.4. Obrada prirodnog jezika

Sljedeća tehnika umjetne inteligencije je obrada prirodnog jezika (*eng. Natural Language Processing - NLP*). Prema (Crabtree, 2024) obrada prirodnog jezika je grana umjetne inteligencije koja je usmjerena na korištenje prirodnog jezika za interakciju između ljudi i računala. Pomoću obrade prirodnog jezika računala mogu proizvesti, razumjeti i interpretirati prirodni jezik na smislen način.

DataJobs (bez dat.) navodi nekoliko razina obrade prirodnog jezika. Jedna od ovih razina je fonologija. **Fonologija** se bavi tumačenjem govornih zvukova u riječima pri čemu postoje tri vrste pravila. Fonetska pravila koriste se za glasove unutar riječi, fonemska za varijacije izgovora kada se riječi izgovaraju zajedno i prozodijska pravila za fluktuaciju naglaska i intonacije. U sustavu obrade prirodnog jezika zvučni valovi se analiziraju i pretvaraju u digitalni signal kako bi se izvelo tumačenje pomoću različitih pravila ili usporedba s određenim jezičnim modelom.

Još jedna razina obrade prirodnog jezika je **morfologija**. Morfologija se bavi morfemima, najmanjim jedinicama značenja od kojih su sastavljene riječi. Morfemi ne mijenjaju svoje značenje u drugim riječima pa to koristi sustav obrade prirodnog jezika koji prepoznaje značenje koje prenosi svaki morphem.

**Leksička razina** bavi se tumačenjem značenja pojedinih riječi. Ova razina može zahtijevati korištenje leksikona što ovisi o sustavu obrade prirodnog jezika. Leksikoni mogu biti jednostavni ili složeni. Jednostavni leksikoni sadržavaju samo riječi i njihove dijelove govora, a složeniji leksikoni sadržavaju više informacija kao što su informacije o semantičkoj klasi riječi, semantičko polje u kojem se koristi svaki smisao višeznačne riječi itd.

**Sintaktička razina** obuhvaća određivanje gramatičke strukture rečenice na temelju analize riječi u rečenici. Rezultat ove razine je prikaz rečenice koja otkriva odnose strukturne ovisnosti između riječi.

**Semantička razina** bavi se mogućim značenjima rečenice na temelju interakcija između značenja na razini riječi u rečenici. Riječi se mogu na ovoj razini semantički razjašnjavati s višestrukim smislom. Semantička višeznačnost omogućuje da se odabire samo jedan smisao višeznačnih riječi koji se uključuje u semantičku reprezentaciju rečenice.

**Diskurs** je razina koja se bavi tekstovima dužim od rečenice i proučava njihova svojstva koja prenose značenje stvaranjem veza između sastavnih rečenica. Postoji nekoliko vrsta obrade diskursa, a neke od njih su razlučivost anafore i prepoznavanje strukture diskursa/ teksta. Razrješenje anafore odnosi se na zamjenu semantički praznih riječi odgovarajućim entitetom, a prepoznavanje strukture diskursa/ teksta bavi se određivanjem funkcija rečenica u tekstu.

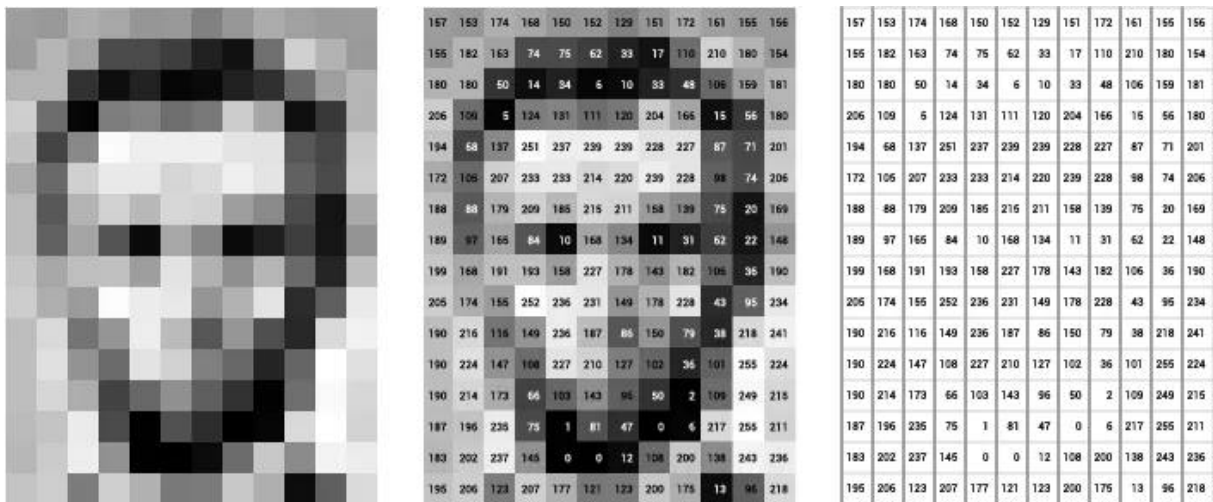
Zadnja razina obrade prirodnog jezika je **pragmatična razina** koja se usredotočuje na svrhovitu upotrebu jezika. Cilj ove razine je objasniti kako se dodatno značenje čita u tekstovima, a da to značenje nije zabilježeno u tim tekstovima.

## 2.1.5. Računalni vid

Posljednja tehnika umjetne inteligencije koju je potrebno objasniti je računalni vid (*eng. Computer Vision*). Prema (Simplilearn, 2024) računalni vid je grana umjetne inteligencije koja omogućuje računalima da korištenjem digitalnih slika i naprednih algoritama dubokog učenja shvate i interpretiraju vizualne podatke. Postoji nekoliko aspekta računalnog vida, a jedan od njih je **prepoznavanje slike** gdje sustav prepoznaje objekt, osobu ili radnju na slici. Još jedan aspekt je **detekcija objekata** koja se odnosi na prepoznavanje i identificiranje lokacije objekata unutar slike. Sljedeći aspekt je **prepoznavanje lica** gdje sustav identificira osobu iz digitalne slike ili video okvira. Još možemo navesti i **analizu kretanja** koja obuhvaća

razumijevanje putanje pokretnih objekata u videu i **strojni vid** koji je kombinacija računalnog vida i robotike za obradu vizualnih podataka i kontrolu kretanja hardvera.

Mihajlović (2019) objašnjava način na koji radi računalni vid. S obzirom na to da se računalni vid temelji na prepoznavanju uzoraka, računalu je potrebno dati velik broj slika koje su označene i koje će onda biti podvrgnute različitim softverskim tehnikama kako bi računalo na temelju njih moglo stvarati uzorke. Na temelju ovih uzorka iz označenih slika, računalo će moći prepoznati neoznačenu sliku. Na slici 1 možemo vidjeti brojčane vrijednosti za svjetlinu piksela od kojih se sastoji slika Abrahama Lincolna.



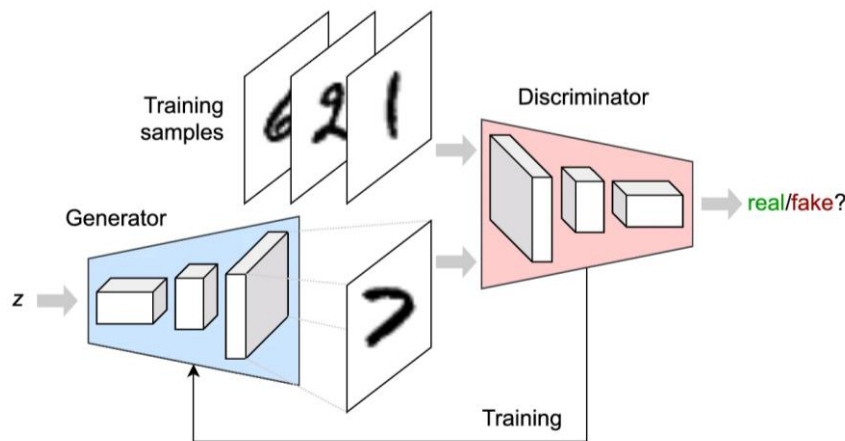
Slika 1: Dijagram podataka piksela (Izvor: Mihajlović, 2019)

## 3. Generativna umjetna inteligencija

Alati umjetne inteligencije za izradu video sadržaja oslanjaju se na vrstu umjetne inteligencije koja se naziva generativna umjetna inteligencija. Prema (Patel, 2023, str. 50) generativna umjetna inteligencija je vrsta umjetne inteligencije koja omogućuje računalima kreiranje novog i jedinstvenog sadržaja koristeći modele strojnog učenja pri čemu kreirani sadržaji izgledaju kao sadržaji koje je kreirao čovjek. Modeli generativne umjetne inteligencije uče iz danog skupa podataka temeljne obrasce i strukture i koriste to razumijevanje za kreiranje novih podataka. Takvi modeli nazivaju se generativni modeli, a treniraju se davanjem skupa podataka iz kojih oni uče obrasce, strukture i varijacije.

### 3.1. Vrste generativnih modela

Postoje različite vrste generativnih modela, a Porter (2023) navodi čak šest vrsta. Prva vrsta generativnih modela je **generativna suparnička mreža** (*eng. Generative Adversarial Network - GAN*). Ove mreže sastoje se od generatora i diskriminatora pri čemu generator kreira sintetičke podatke, kao što su slike i zvuk, iz nasumičnog šuma, a diskriminator treba prepoznati koji su podaci stvarni, a koji generirani. Cilj generatora je stvoriti što realističniji sadržaj kako bi zavarao diskriminator, dok diskriminator u isto vrijeme poboljšava svoju sposobnost prepoznavanja lažnih podataka. Zahvaljujući stvaranju realističnog sadržaja, GAN-ovi se koriste u sintezi slika, stvaranju umjetnosti i video generiranju. Na slici 2 prikazan je način rada generativnih suparničkih mreža koje se sastoje od generatora i diskriminatora. Generator kao ulaz koristi nasumični šum i generira podatke tako da budu što bliži stvarnim podacima, a diskriminator, kao što možemo vidjeti na slici, kao ulaz koristi podatke koje je stvorio generator i uzorke za obuku koji su stvarni podaci, i zatim treba odrediti koji su podaci stvarni, a koji lažni. S vremenom će, tijekom obuke, generator stvarati sve uvjerljivije podatke.



Slika 2: Generativna suparnička mreža (Izvor: Feuerriegel, Hartmann, Janiesch i Zschech, 2024)

Još jedna vrsta generativnih modela su i varijacijski autokoderi (*eng. Variational Autoencoder - VAE*). „Varijacijski autokoderi koji uče kodirati podatke u latentni prostor i zatim ih dekodirati natrag kako bi rekonstruirali izvorne podatke.“ Najčešće se koriste za generiranje slika, ali i teksta i zvuka.

Autoregresivni modeli (*eng. Autoregressive Models*) su modeli koji uzimaju u obzir kontekst prethodnih elemenata kako bi predvidjeli distribuciju vjerojatnosti sljedećeg elementa nakon čega koriste distribuciju za prepoznavanje uzoraka kako bi generirali nove podatke. Primjer autoregresivnog modela je generativni unaprijed trenirani transformator (*eng. Generative Pre-trained Transformer - GPT*) koji može generirati kontekstualno prikladan tekst.

„Ponavljajuće neuronske mreže (*eng. Recurrent Neural Network - RNN*) su vrsta neuronske mreže koja obrađuje sekvencijalne podatke, kao što su rečenice prirodnog jezika ili podaci vremenskih serija.“ Nedostatak ovih neuronskih mreža je da imaju problem kod generiranja dugih nizova pa su razvijene naprednije varijante koje rješavaju ovaj problem.

Modeli temeljeni na transformatorima (*eng. Transformer-based Models*) modeliraju odnose između različitih elemenata u nizu korištenjem mehanizama pažnje, a koriste se za generiranje smislenog i kontekstualno relevantnog teksta.

Učenje putem pojačanja za generativne zadatke (*eng. Reinforcement Learning for Generative Tasks*) je vrsta generativnog modela u kojem agent u interakciji s okolinom uči generirati podatke i zatim dobiva povratne informacije o kvaliteti generiranih uzoraka. Ovakav model obično se koristi za generiranje teksta.



## 3.2. Primjene generativne umjetne inteligencije

Zahvaljujući generiranju novog sadržaja generativna umjetna inteligencija ima široku primjenu u raznim područjima. Generiranje eseja, razvoj koda i prijevod primjeri su upotrebe generativne umjetne inteligencije koja se zasniva na generativnim modelima koji se temelje na jeziku i nazivaju se velikim jezičnim modelima (*eng. Large Language Model - LLM*). Generativna umjetna inteligencija koristi se i za razvoj audio sadržaja pa tako može generirati pjesme i isječke audio zapisa na temelju unesenog teksta, kao i kreirati popratne zvukove za video snimke. Najčešća upotreba generativne umjetne inteligencije je generiranje slika, videa, grafikona, avatara i drugih ilustracija. Prilikom generiranja slika postoji mogućnost odabira različitih stilova i tehnika za uređivanje i modificiranje vizuala. Još jedna primjena umjetne inteligencije je razvoj sintetičkih podataka koji se koriste kada ne postoje podaci, ograničeni su ili nisu u mogućnosti riješiti problem s najvećom točnošću. Prednost generativnih modela koji se koriste za razvoj sintetičkih podataka je u tome što stvaraju dodatne proširene podatke za obuku ili uče unutarnju reprezentaciju podataka što rezultira smanjenjem troškova označavanja jer obuka modela zahtijeva manje označenih podataka (Nvidia, bez dat.).

## 3.3. Vrste generativne umjetne inteligencije

Postoje različite vrste generativne umjetne inteligencije koje se razlikuju po vrsti sadržaja koji mogu biti ulazni i izlazni (Patel, 2023, str. 57-58).

Generativna umjetna inteligencija **tekst-u-tekst** (*eng. Text-to-Text - T2T*) generira tekst na temelju unesenog teksta. Ovakva vrsta modela primjenjuje se kod strojnog prevođenja, sažimanja teksta i chatbotova, a primjeri su ChatGPT, Gemini i drugi slični modeli.

**Tekst-u-audio** (*eng. Text-to-Audio - T2A*) je vrsta generativne umjetne inteligencije koja se koristi za pretvaranje pisanog teksta u izgovorene riječi i koristi se za stvaranje audio knjiga i u sustavima pretvaranja teksta u govor (*eng. text-to-speech - TTS*). Primjeri su Googleov Tacotron i Amazonova Polly.

Generativna umjetna inteligencija **tekst-u-sliku** (*eng. Text-to-Image - T2I*) koristi se za generiranje slike na temelju tekstualnog opisa. Primjeri su DALL-E i MidJourney koji generiraju slike na temelju unesenog teksta.

Generativna umjetna inteligencija **tekst-u-video** (*eng. Text-to-Video - T2V*) koristi se za generiranje videa na temelju unesenog teksta. Ovi alati se još uvijek razvijaju jer je pretvaranje teksta u video složeniji zadatak.

Generativna umjetna inteligencija **tekst-u-glazbu** (eng. *Text-to-Music - T2M*) generira glazbu na temelju tekstualnog unosa. Ovakvi modeli mogu kreirati melodije na temelju tekstualnog opisa ili stvoriti glazbenu skladbu na temelju raspoloženja. Primjer je OpenAI-jev MuseNet.

**Audio-u-tekst** (eng. *Audio-to-Text - A2T*) je vrsta generativne umjetne inteligencije za pretvaranje govornog jezika u pisani tekst. Ovakvi modeli poznati su i kao sustavi za automatsko prepoznavanje govora. Audio-u-tekst nije generativan u pravom smislu te riječi, ali čini značajan dio u području umjetne inteligencije. Najpoznatiji primjeri korištenja su glasovni pomoćnici poput Amazonove Alexe i Appleove Siri.

Generativna umjetna inteligencija **tekst-u-kod** (eng. *Text-to-Code - T2C*) koristi se za pretvaranje prirodnog jezika u programski kod. Ovakvi modeli imaju sposobnost generirati kod, dovršavati kod ili otkrivati greške u kodu. GitHub Copilot je primjer za ovu vrstu generativne umjetne inteligencije.

Generativna umjetna inteligencija **slika-u-video** (eng. *Image-to-Video*) koristi se za generiranje video sadržaja na temelju slike kao ulaznog podatka. Kao ulaz mogu se koristiti jedna ili niz slika, a video generator će kreirati video koji je u skladu sa sadržajem, stilom i formatom slike. Ovakav način generiranja video sadržaja omogućuju modeli dubokog učenja koji omogućuju umjetnoj inteligenciji da analizira unos i stvori realističnu grafiku.

### 3.4. Generiranje video sadržaja pomoću umjetne inteligencije

Generiranje video sadržaja pomoću umjetne inteligencije je stvaranje video sadržaja pomoću tehnologija umjetne inteligencije na temelju tekstualnih opisa ili skripti (AI Video Generation, 2024). AI video generatori koriste algoritme strojnog učenja za analiziranje unesenog teksta na temelju kojeg stvaraju vizualne elemente, prijelaze i zvukove. Osim jednostavnog stvaranja videa na temelju tekstualnog unosa, moguće je i stvaranje složenih videozapisa koji uključuju avatare umjetne inteligencije. Kvaliteta generiranih videozapisa je veća što su podaci koji se unose u algoritme raznolikiji i sveobuhvatniji.

VideoZen (2023) opisuje povijesni razvoj umjetne inteligencije u generiranju video sadržaja. Početak razvoja započeo je još 1960-ih i 1970-ih godina kada su se pojavile **računalno generirane slike** (eng. *computer-generated imagery - CGI*) koje su omogućavale stvaranje jednostavnih vizualnih efekata i grafika. Ipak, postojala su određena ograničenja u pogledu računalne snage i nedostatka sofisticiranosti. Značajan preokret dogodio se 1980-ih kada su otkrivene **neuronske mreže** koje su računalima omogućile obradu i prepoznavanje

obrazaca u podacima. To je dovelo do stvaranja složenijeg i nijansiranijeg vizualnog sadržaja. Transformacija umjetne inteligencije u generiranju video sadržaja dogodila se tijekom 2010-ih kada su se pojavili **duboko učenje** i **konvolucijske neuronske mreže**. Zahvaljujući algoritmima dubokog učenja koji imaju sposobnost učenja iz velikih skupova podataka i generiranja zamršenih slika, ostvaren je napredak u obradi videa pa se počinju koristiti detekcija objekata, segmentacija slike i prepoznavanje scene. Konvolucijske neuronske mreže omogućile su analizu i razumijevanje video okvira koji su korisni za automatsko označavanje videa, moderiranje sadržaja i sažimanje videa. 2014. godine Ian Goodfellow predstavio je **generativne suparničke mreže** (eng. *Generative Adversarial Networks - GANs*). Generativne suparničke mreže omogućile su stvaranje vrlo realističnih i uvjerljivih vizualnih sadržaja zbog čega su se počele primjenjivati u raznim poljima kao što su stvaranje realističnih likova i scena u videoigrama, stvaranje realističnih specijalnih efekata u filmovima itd. u budućnosti se očekuje da će se umjetna inteligencija u generiranju video sadržaja još više razvijati, a neka od područja koja imaju potencijal za rast su prijenos stila videozapisa u stvarnom vremenu, generiranje personaliziranog sadržaja i alati za uređivanje videozapisa.

Michelson (2023) opisuje način na koji AI video generatori kreiraju video sadržaje. Prvi korak je prikupljanje velike količine video sadržaja koji se moraju prethodno obraditi kako bi se izdvojile najrelevantnije značajke za prepoznavanje vizualnih elemenata. Sljedeći korak je podučavanje neuronskih mreža. Kako bi umjetna inteligencija razumjela vizualne komponente unutar svakog okvira, koriste se konvolucijske neuronske mreže (eng. *convolutional neural network - CNN*) koje omogućuju analizu slike. S druge strane, ponavljajuće neuronske mreže (eng. *recurrent neural network - RNN*) koriste se za prepoznavanje sekvencijalne prirode video sadržaja. Nakon podučavanja neuronskih mreža potrebno je obaviti ekstrakciju značajki i predstavljanje pri čemu umjetna inteligencija izdvaja iz ulaznih podataka značajke koje su temelj za generiranje novog sadržaja. Nakon ovog koraka slijedi konceptualizacija i razumijevanje scene tijekom koje video generator uči prepoznati scene, interpretirati kontekst i razumjeti narativnu strukturu videa. Na temelju svega naučenog video generator može sintetizirati novi video sadržaj s obzirom na tekstualni ili vizualni unos. Neki video generatori omogućuju prijenos stila oponašanjem umjetničkih stilova ili vizualnih žanrova koje koriste prilikom generiranja video sadržaja. Na kraju je potrebno obaviti pročišćivanje izlaza kako bi se poboljšala kvaliteta konačnog videozapisa.

Generiranje sadržaja pomoću umjetne inteligencije ima mnoge prednosti. Umjetna inteligencija smanjuje troškove i vrijeme potrebno za proizvodnju videa jer automatizira zadatke koje obično obavljaju ljudi. Umjetna inteligencija se također može koristiti za poboljšanje kvalitete konačnog videozapisa poboljšanjem boja i dodavanjem vizualnih elemenata kako bi video bio privlačniji. Umjetna inteligencija također pruža mogućnost manjim

tvrtkama da stvaraju video sadržaje jer za to više nije potreban velik broj ljudi. Dovoljno je imati samo scenarij ili priču na temelju koje će se generirati videozapis. Na ove načine se kreatori video sadržaja mogu više fokusirati na kreativne aspekte procesa (AIContentfy team , 2023).

Generiranje video sadržaja pomoću umjetne inteligencije ima i određene nedostatke. Jedan od nedostataka je stvaranje lažnih ili obmanjujućih videozapisa. Umjetna inteligencija, zahvaljujući novoj razini tehnologije „dubokog lažiranja“ (eng. „*deep fake*“ *technology*), omogućuje bilo kome da stvara realistične video sadržaje koji mogu netočno prikazivati ljude ili događaje mijenjanjem izraza lica i govora ljudi tako da nastane video sadržaj koji izgleda autentično, a prikazuje osobu koja govori nešto što u stvarnosti zapravo nikad nije rekla. Još jedan od nedostataka je i opasnost od gubitka posla za video montažere i animatore. Kako raste potražnja za isplativim alatima umjetne inteligencije koji se koriste za izradu video sadržaja, tako su i poslovi video montažera i animatora sve ugroženiji. Još jedan problem koji možemo navesti su i pristranost i diskriminacija. S obzirom na to da se AI video generatori treniraju na postojećim podacima, mogu sadržavati pristranosti i predrasude zbog čega je važno te pristranosti i predrasude identificirati i trenirati modele protiv njih uvođenjem novih skupova podataka. Kao nedostatak vrlo je važno napomenuti i pravna pitanja i autorska prava jer se sadržaj koji stvara umjetna inteligencija temelji na postojećem sadržaju koji je u nečijem vlasništvu. Stoga može doći do problema oko utvrđivanja vlasništva i atribucije takvog sadržaja (Talvola, 2023).

Najčešći način generiranja video sadržaja pomoću umjetne inteligencije je na temelju tekstualnog upita. Kako bismo dobili videozapise koji odgovaraju onome što smo zamislili, moramo zadati tekstualne upite koje će umjetna inteligencija razumjeti na zadovoljavajući način. Elena (2023) navodi savjete kako napisati najbolje tekstualne upite za generiranje video sadržaja pomoću alata umjetne inteligencije. Prilikom pisanja tekstualnog upita, važno je izbjegavati složene izraze kako umjetna inteligencija ne bi imala problema u rješavanju složenih upita. Osim toga, važno je upitima dodati više detalja. Na primjer, možemo navesti koliko će video trajati, kome je namijenjen video, na kojoj platformi će video biti objavljen, kojeg stila treba biti generirani video itd. Pritom moramo imati na umu da ne unesemo previše informacija jer bi dodatne riječi mogle zbuniti umjetnu inteligenciju. Također je važno ne ostavljati u upitu otvoreno pitanje jer umjetna inteligencija tada može dati samo neutralan odgovor što znači da bi generirani sadržaj tada mogao biti izvan naših očekivanja. Saleem (2024) također navodi određene smjernice za pisanje tekstualnih upita za generiranje video sadržaja pomoću umjetne inteligencije. Prilikom pisanja upita važno je navesti što točno želimo postići videom, odnosno navesti cilj videa. Kada ciljamo na određeni vizualni stil, trebamo uključiti detaljne opise željenih slika, shemu boja i ukupnu estetiku. Umjetnu inteligenciju također možemo koristiti za transformaciju teksta u vizualne ploče scenarija, kao i za

generiranje popisa snimaka. Također je poželjno kreativno eksperimentirati s novim vizualnim stilovima koje je stvorila umjetna inteligencija.

## **4. Primjena umjetne inteligencije u procesu izrade video sadržaja**

Umjetna inteligencija u procesu izrade video sadržaja koristi se u različitim industrijskim sektorima, kao što su marketing, obrazovanje, zdravstvo i filmska industrija, i ubrzava proces izrade novih i jedinstvenih videozapisa. U nastavku će biti navedeni konkretni primjeri upotrebe umjetne inteligencije u izradi video sadržaja za navedene industrijske sektore.

### **4.1. Primjena umjetne inteligencije u procesu izrade video sadržaja u marketingu**

AIContentfy team (2024) navodi nekoliko vrsta sadržaja koji se mogu generirati pomoću umjetne inteligencije za područje marketinga pa je tako moguće generirati 2D ili 3D animacije, kao što su npr. demonstracije proizvoda i vizualni afekti. Animacije se mogu koristiti za izradu videozapisa koji predstavljaju robne marke ili značajke proizvoda na vizualno privlačan način. Osim animacija, umjetna inteligencija može kreirati i virtualne glasnogovornike koji se mogu koristiti u videozapisima čije je svrha prenošenje marketinške poruke ili demonstracija proizvoda. Virtualni glasnogovornici mogu se prilagoditi izgledu robne marke i na taj način stvoriti dojam profesionalnosti i dosljednosti. Algoritmi umjetne inteligencije također se koriste za stvaranje interaktivnih videozapisa koji mogu biti u obliku kvizova, anketa i interaktivnih demonstracija proizvoda. To omogućuje gledateljima da se uključe u sadržaj na inovativan način. Sljedeći primjer sadržaja koji se kreira su sažeci dužih videozapisa. Izradom kratkih i jasnih sažetaka dugih videozapisa, umjetna inteligencija omogućuje korisnicima lakše razumijevanje sadržaja. Generiranje glasa još je jedan primjer sadržaja koji generiraju algoritmi umjetne inteligencije. Umjetna inteligencija na ovoj način olakšava posao marketinškim stručnjacima koji ne moraju angažirati glasovne glumce. Primjeri korištenja generiranja glasa pomoću umjetne inteligencije su video upute, demonstracije proizvoda i videozapisi robne marke.

Kao što smo već mogli primijetiti, sadržaji generirani pomoću umjetne inteligencije mogu se koristiti u različitim područjima marketinga. Jedan od primjera su demonstracije proizvoda koje predstavljaju značajke i prednosti proizvoda ili usluga i mogu se koristiti na web stranicama poduzeća, društvenim medijima ili marketinškim kampanjama. Još jedan primjer su priče o robnoj marki koje se mogu ostvariti na vizualno privlačan način i omogućiti poduzeću da se poveže sa svojom publikom i stvori snažan imidž robne marke. Umjetna inteligencija može se koristiti i za kreiranje video vodiča koji su namijenjeni obrazovanju i informiranju o

proizvodima ili uslugama i omogućuju poboljšanje angažmana kupaca. Također se mogu generirati i privlačni videozapisi za platforme društvenih medija koji se koriste za doseganje šire publike, kao i kratki videozapisi koji se mogu ugraditi u marketinške kampanje putem e-pošte. Umjetna inteligencija može se koristiti i za kreiranje virtualnih događaja poput webinarima i konferencija koji se mogu koristiti za interakciju s publikom u stvarnom vremenu.

Postoje razne prednosti generiranja video sadržaja pomoću umjetne inteligencije u području marketinga, a jedna od tih prednosti je brzina i učinkovitost. Umjetna inteligencija omogućuje izradu visokokvalitetnih sadržaja u vrlo kratkom vremenu pa je to idealno rješenje za poduzeća koja trebaju brzo i učinkovito generirati sadržaj. Jedna od prednosti je i isplativost s obzirom na to da umjetna inteligencija smanjuje potrebu za angažiranjem tima stručnjaka za izradu video sadržaja što je idealno za novoosnovana i manja poduzeća s ograničenim financijskim mogućnostima. Sadržaji generirani umjetnom inteligencijom također omogućuju dosljednost u smislu kvalitete, tona i stila, ali i inovativnost koju možda ne bi bilo moguće postići ručnom izradom sadržaja.

Jedan od primjera videozapisa generiranog pomoću umjetne inteligencije u području marketinga je video koji predstavlja robnu marku Toys „R“ Us. Za izradu ovog videa bilo je potrebno nekoliko tjedana, a video je izrađen pomoću OpenAI-jeve video tehnologije Sora. Osim umjetne inteligencije, u izradi videa su jednim dijelom sudjelovali i ljudi. Video se nalazi na sljedećoj poveznici: <https://www.youtube.com/watch?v=uTVInehpRHQ>. Video je vrlo realističan, ali ima nekoliko pogrešaka. Na primjer, kroz gotovo cijeli video možemo vidjeti da dječak ima pjege, osim na kraju videa gdje vidimo dječaka bez pjega. Na slici 3 tako možemo vidjeti dječaka s pjegama, dok na slici 4 vidimo da dječak nema pjega, a nosi i drugačije naočale. Na slici 5 možemo vidjeti još jednu pogrešku koja je vezana uz dječakove naočale koje su nepravilnog oblika i nisu simetrične. Ovi pogreške ne narušavaju kvalitetu videa, ali će ih pažljiviji gledatelji sigurno uočiti.



Slika 3: Dječak s pjegama iz videa Toys „R“ Us (Izvor: YouTube)



Slika 4: Dječak bez pjega iz videa Toys „R“ Us (Izvor: YouTube)





Slika 5: Nepravilne naočale iz videa Toys „R“ Us (Izvor: YouTube)

## 4.2. Primjena umjetne inteligencije u procesu izrade video sadržaja u obrazovanju

Obrazovanje je još jedan industrijski sektor u kojem se primjenjuje umjetna inteligencija u procesu izrade video sadržaja. Veliki interes za online tečajevima i diplomama potaknuo je potražnju za visokokvalitetnim obrazovnim resursima, uključujući velik broj videozapisa s uputama za učenje. Da bi ovakvi videozapisi bili učinkoviti, ne smiju koristiti nepotrebne i pretjerane elemente koji odvrćaju pažnju od učenja i preopterećuju kognitivno opterećenje. Ove zahtjeve mogu ispuniti sintetički videozapisi namijenjeni učenju i kreirani pomoću alata umjetne inteligencije. Jedan od primjera korištenja video sadržaja generiranih umjetnom inteligencijom u obrazovanju su i virtualni instruktori ili animirani pedagoški agenti. Virtualni instruktori su likovi na ekranu realističnog prikaza koje je utjelovilo računalo kako bi omogućilo učenje davanjem uputa i smjernica. Korištenje virtualnih instruktora pozitivno djeluje na motivaciju i stavove učenika tijekom učenja (Sharma i Achar, 2024).

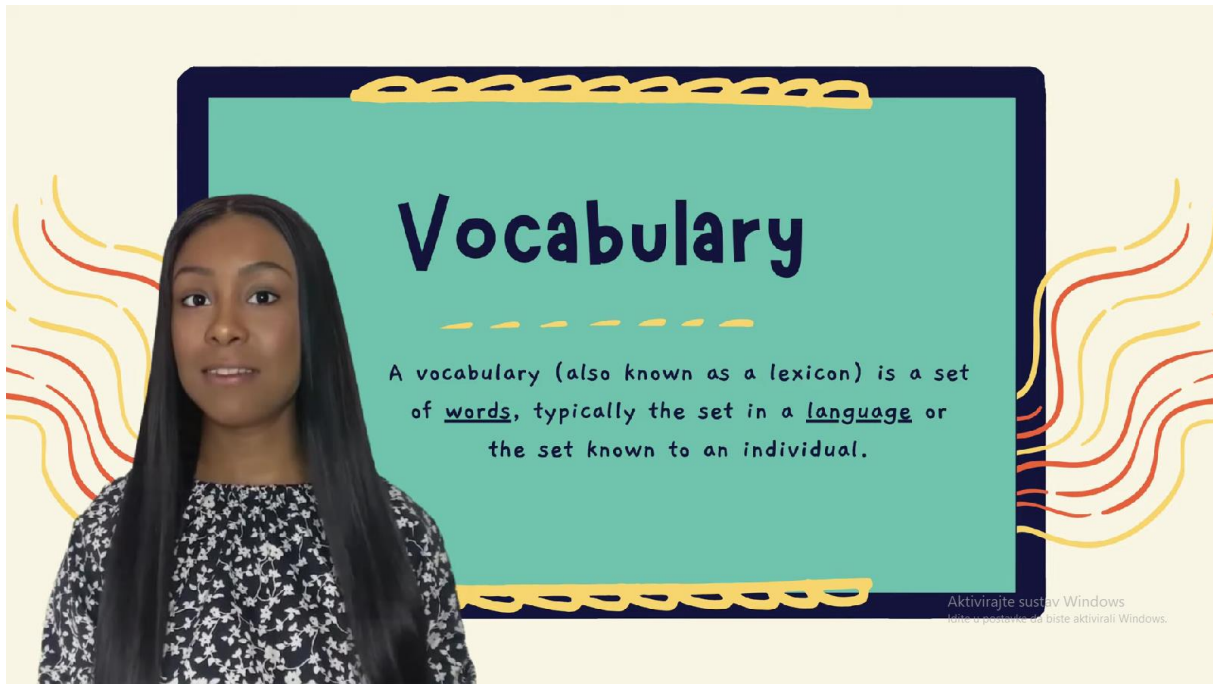
Veliku popularnost u sektoru obrazovanja stekli su video avatari. Tekst „Revolutionising Education: AI Avatar Videos for Engaging Learning“ (2023) objašnjava prednosti korištenja video avatara u učenju. Prva prednost koju navodi je interaktivnost i privlačan sadržaj. Za razliku od tradicionalnih udžbenika, AI video avatari mogu sadržavati multimedijske elemente, igre i aktivnosti iz stvarnog života čineći materijale za učenje privlačnijima za učenike. AI video

avatari dostupni su kroz cijelo vrijeme pa učenici mogu pristupiti obrazovnom materijalu kad god im to odgovara. Ovakav pristup posebno je koristan obrazovnim ustanovama s ograničenim resursima i financijskim mogućnostima jer smanjuje potrebu za fizičkim materijalima kao što su udžbenici. AI video avatari osiguravaju dosljednu kvalitetu nastave jer ne pate od umora i drugih problema s kojima bi se mogli suočiti ljudi. Još jedna prednost AI video avatara je poboljšanje pristupačnosti na način da avatari ponude sadržaj na više jezika, prilagode se poteškoćama u učenju ili daju tekstualne opise za učenike s oštećenjima sluha.

Jedan od primjera primjene umjetne inteligencije u procesu izrade video sadržaja u obrazovanju je obrazovni video o gramatici engleskog jezika koji se nalazi na sljedećoj poveznici: <https://youtu.be/zweEQAFSV3Y?si=uM1bCBmlrTjIDnMN>. Ovaj video je nastao korištenjem platforme Synthesys AI Studio. Općenito gledajući, video je zadovoljavajuće kvalitete, ali ponegdje možemo primijetiti neprirodne treptaje očima i micanje ustima video avatara, kao i neke neprirodne pokrete tijela. Na slici 6 možemo vidjeti dobar primjer upotrebe video avatara koji položajem tijela djeluje vrlo prirodno, slično pravom čovjeku. S druge strane, na slici 7 možemo vidjeti primjer loše upotrebe video avatara koji je svojim položajem tijela usmjeren malo prema desno pri čemu avatar djeluje ukočeno i neprirodno.



Slika 6: Primjer dobro generiranog video avatara u obrazovnom videu (Izvor: YouTube)



Slika 7: Primjer loše generiranog video avatara u obrazovnom videu (Izvor: YouTube)

### 4.3. Primjena umjetne inteligencije u procesu izrade video sadržaja u zdravstvu

Zdravstvo je još jedan primjer gdje se koristi umjetna inteligencija u procesu izrade video sadržaja. Kako bi se izvršila edukacija pacijenata koristi se generativna umjetna inteligencija tekst-u-video koja pretvara složene medicinske izraze u prilagođene videozapise koji daju objašnjenja bolesti, planova liječenja i učinke lijekova i na taj način zamjenjuje teško razumljive dokumente. U budućnosti se očekuje da će umjetna inteligencija moći realno simulirati kirurške postupke za poboljšano učenje ili objasniti složene biokemijske koncepte što bi ubrzalo shvaćanje složenih medicinskih koncepata i nadopunilo ključna praktična iskustva na inovativan način. Umjetna inteligencija tekst-u-video može se koristiti i u terapiji mentalnog zdravlja u sklopu terapije izloženosti generiranjem vizualizacije pokretačkih scenarija. Stvaranjem kontroliranih imerzivnih okruženja, pacijenti se mogu suočiti sa svojim strahovima u sigurnoj okolini. Videozapisi koje je generirala umjetna inteligencija mogu se koristiti i za poboljšanje telemedicine na način da, tijekom konzultacija liječnika s pacijentom, liječnik unosi opise simptoma, a umjetna inteligencija generira vizualna pomagala kako bi demonstrirala postupke samopregleda i rutine vježbanja ili kako bi ilustrirala moguće dijagnoze (Odabashian, 2024).

Videozapisi generirani pomoću umjetne inteligencije koji se koriste za poboljšanje medicinskog obrazovanja imaju određene prednosti, ali i mane ("Revolutionizing Learning & Training in Hospitals: The Pros and Cons of Using Deepfake Generators for AI-Powered Videos", bez dat.). Jedna od prednosti je stvaranje realističnih i sveobuhvatnih scenarija iz stvarnog života što omogućuje zdravstvenim djelatnicima da steknu praktično iskustvo bez izlaganja pravih pacijenata riziku. Velika su prednost i personalizirana iskustva učenja. Umjetna inteligencija može prilagoditi sadržaj za svakog učenika ovisno o njegovim vještinama i specifičnim potrebama pa se tako svaki zdravstveni djelatnik može usredotočiti na ona područja u kojima mu je potrebno poboljšanje. Skalabilnost videozapisa također je jedna od prednosti korištenja umjetne inteligencije u kreiranju video sadržaja. Za razliku od tradicionalnog načina obuke koji zahtijeva određene resurse kao što su stručni treneri i fizička oprema, videozapisi generirani pomoću umjetne inteligencije omogućuju zdravstvenim djelatnicima pristup u bilo koje vrijeme i s bilo kojeg mjesta, bez zahtijevanja fizičkih resursa.

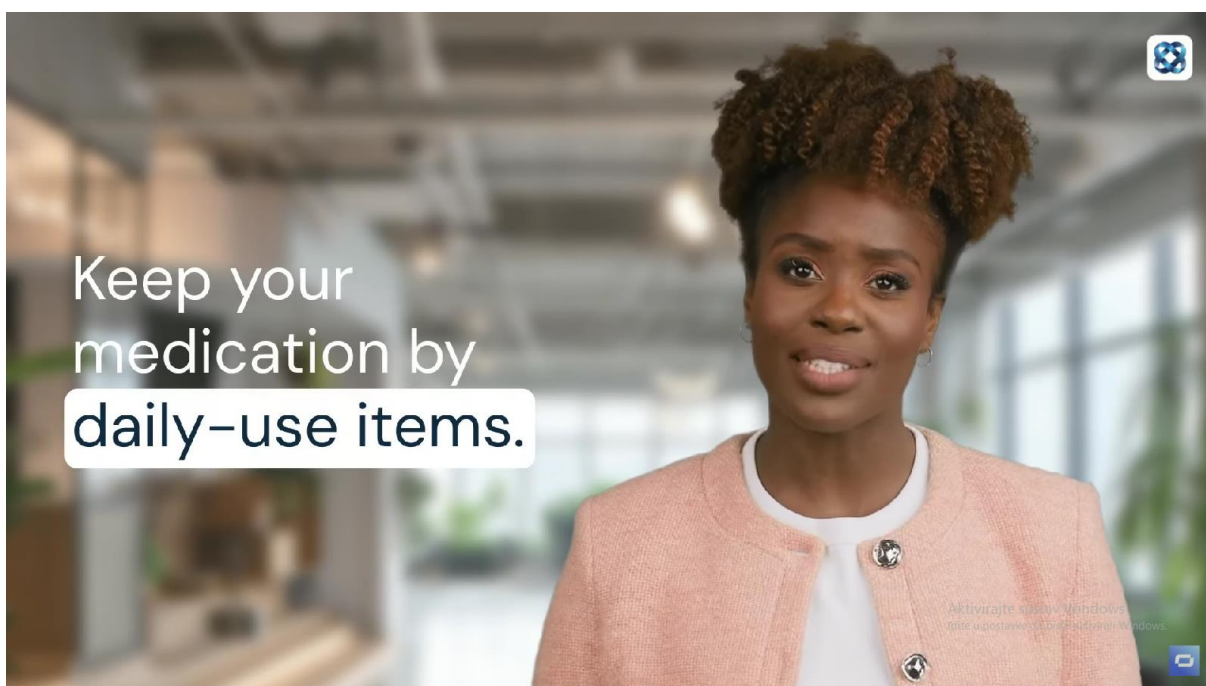
Osim navedenih prednosti, potrebno je opisati i nedostatke videozapisa generiranih pomoću umjetne inteligencije koji se koriste za medicinsku obuku. Jedan od nedostataka su etički problemi koji se mogu javiti prilikom korištenja ove tehnologije. Postoji veliki rizik da bi zlonamjerni pojedinci mogli koristiti ovu tehnologiju kako bi širili dezinformacije s obzirom na to da umjetna inteligencija može stvarati vrlo realistične prikaze. Još jedan od nedostataka je problem pouzdanosti i točnosti videozapisa. S obzirom na to da se modeli moraju obučavati na velikim količinama podataka, ti podaci ne smiju biti nepotpuni ili pristrani jer bi to moglo dovesti do ozbiljnih medicinskih posljedica. Također se javlja i problem ulaganja jer bolnice moraju uložiti financijska sredstva u nabavu hardvera i softvera, obuku zaposlenika i održavanje infrastrukture.

Jedan od primjera umjetne inteligencije u procesu izrade video sadržaja u zdravstvu je video u kojem video avatar pacijentu daje savjete o uzimanju lijekova. Video se nalazi na sljedećoj poveznici: <https://www.youtube.com/watch?v=rIpVTpGpiPY>. Video avatar izgleda vrlo realistično, a razlika između avatara i pravog čovjeka je gotovo neprimjetna. Ono što bi moglo biti bolje je to što u videu piše da će se spomenuti tri načina na koji će uzimanje lijekova postati dio pacijentove rutine, a spominju se samo dva načina, odnosno savjeta. Slika 8 prikazuje video avatara s tekstualnim opisom teme videa, a slika 9 video avatara koji objašnjava pacijentu jedan od načina redovitog uzimanja lijekova.





Slika 8: Video avatar s tekstualnim opisom teme videa (Izvor: YouTube)



Slika 9: Video avatar objašnjava pacijentu jedan od načina uzimanja lijekova (Izvor: YouTube)

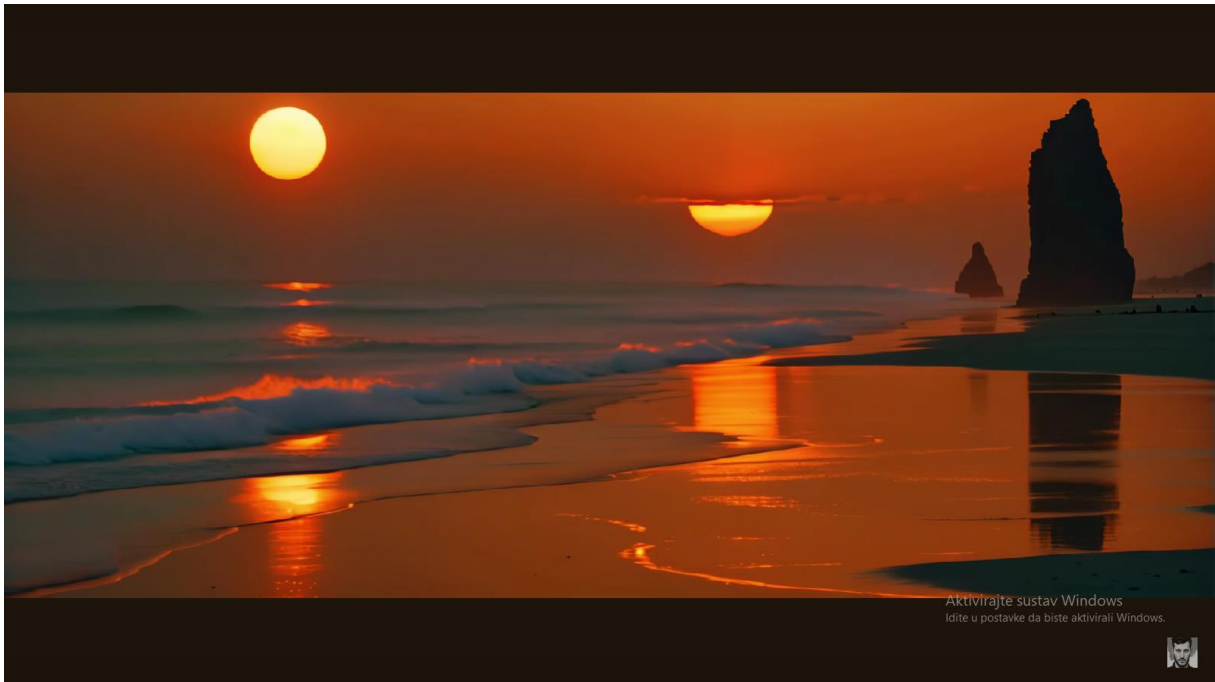
## **4.4. Primjena umjetne inteligencije u procesu izrade video sadržaja u filmskoj industriji i industriji vizualnih efekata (VFX)**

U filmskoj industriji umjetna inteligencija u procesu izrade video sadržaja, odnosno filmova, također ima široku primjenu. Može se koristiti za pisanje scenarija na način da se model trenira velikom količinom podataka u obliku scenarija. Model zatim uči iz tih podataka kako bi na kraju mogao generirati jedinstveni scenarij. Osim stvaranja, umjetna inteligencija također može analizirati scenarije i dati povratne informacije u obliku pitanja i prijedloga. Umjetna inteligencija koristi se i u pretprodukciji kao pomoć u planiranju rasporeda, pronalaženju lokacija i drugim pripremnim aktivnostima. U montaži filmova umjetna inteligencija se koristi na način da prepoznaje akcijske i emotivne filmske scene koje se onda koriste za izradu najave za film. Osim toga, umjetna inteligencija je korisna i za uređivanje cjelovečernih filmova, zahvaljujući svojoj sposobnosti da odredi glavne likove i scene koje uključuju glavnu radnju koristeći tehnologiju prepoznavanja lica. Postoje razni alati umjetne inteligencije koji se koriste za generiranje glazbe i mogu se iskoristiti za stvaranje glazbe koja će biti uključena u film. Umjetna inteligencija koristi glazbene obrasce kako bi generirala glazbu prilagođenu filmskom žanru i situacijama koje se očekuju u filmu. Stvaranje vlastitih filmova također je jedna od sposobnosti umjetne inteligencije. Primjer je Benjamin AI koji je u 48 sati stvorio film „Zone Out“. Na kraju, umjetna inteligencija može se koristiti i za predviđanje uspjeha filma (Andriasyan, bez dat.).

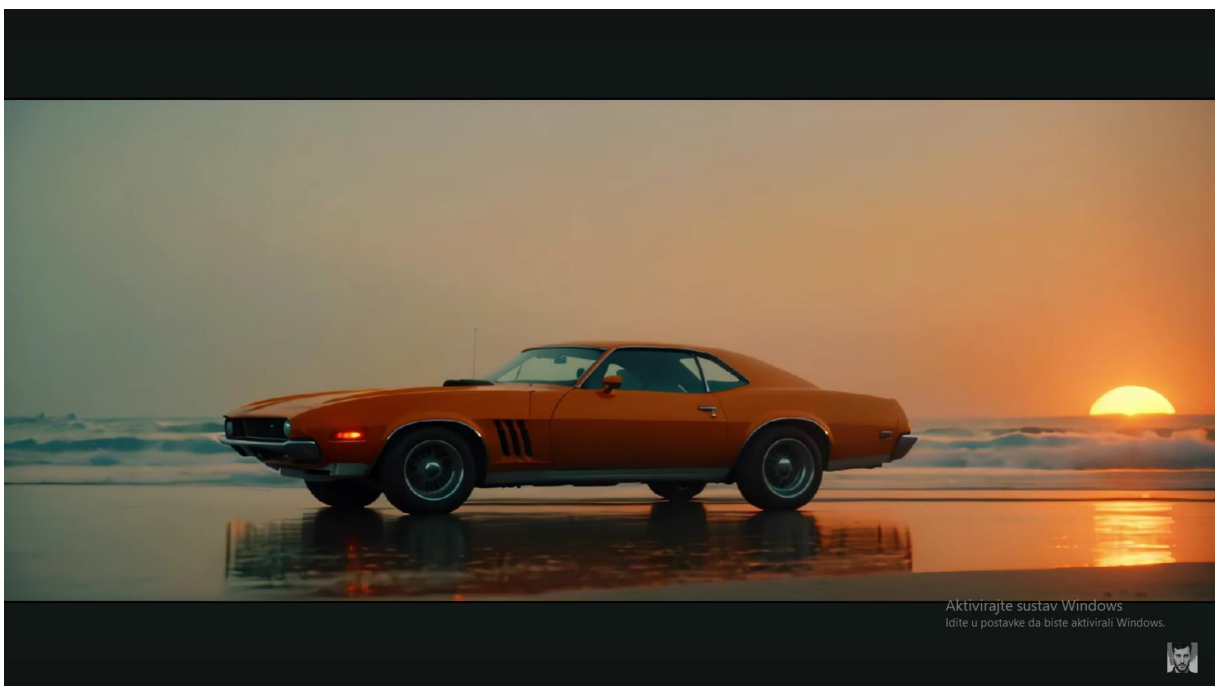
U industriji vizualnih efekata (VFX) također se primjenjuje umjetna inteligencija u izradi video sadržaja. Jedan od primjera je CGI modeliranje i animacija. Program dubokog učenja likova omogućuje kretanje i stvaranje likova na temelju skupa slika ili filmova ljudi pri čemu analizira značajke odabrane osobe. Na temelju identificiranih obilježja subjekta program stvara generiranu teksturu lica i tijela zajedno sa zadanim pozama osobe i počinje ih animirati pri čemu oponaša okolno osvjetljenje subjekata za stvaranje realističnijeg prikaza. Umjetna inteligencija koristi se i kod rotoskopiranja. Rotoskopiranje se odnosi na izdvajanje pojedinačnih elemenata iz pozadine kadra. Za odvajanje elemenata prije su se koristili zeleni i plavi zasloni, a sada računalni programi izdvajaju elemente i likove iz kadra na način da skeniraju snimku kako bi pronašli lica i na temelju toga zaključili što je, a što nije osoba. Nakon toga, program prati osobu izvan snimke (Xagoraris, 2020).

Jedan od primjera primjene umjetne inteligencije u procesu izrade video sadržaja u VFX i filmskoj industriji je trailer za film The Outworld koji se nalazi na sljedećoj poveznici: <https://www.youtube.com/watch?v=s1AXccDlm6A>. Vidljivo je da ljudi nisu stvarni, već

generirani umjetnom inteligencijom, ali je videozapis vrlo realističan. Osim prikaza ljudi, postoje još neke manje pogreške koje bi se mogle poboljšati, kao što je npr. prikaz zalaska sunca koji možemo vidjeti na slici 10 gdje je uz postojeće sunce prikazana još jedna polovica drugog sunca. S druge strane, neke stvari su vrlo kvalitetno prikazane pa tako na slici 11 možemo vidjeti prikaz automobila na plaži koji izgleda kao da je iz stvarnog svijeta.



Slika 10: Prikaz zalaska sunca iz trailera za film The Outworld (Izvor: YouTube)



Slika 11: Prikaz automobila na plaži iz trailera za film The Outworld (Izvor: YouTube)

## 5. Alati umjetne inteligencije za izradu video sadržaja

### sadržaja

Postoji već velik broj alata umjetne inteligencije za izradu video sadržaja koji se razlikuju po značajkama, cjenovnim planovima, korisničkom sučelju itd. Neki od dostupnih alata nude mogućnost generiranja videozapisa samo iz tekstualnog unosa, dok drugi nude naprednije mogućnosti generiranja na temelju slike ili videa. U tablici 2 možemo vidjeti popis nekih alata umjetne inteligencije za izradu video sadržaja zajedno s poveznicama na svaki alat.

Tablica 2: Alati umjetne inteligencije za izradu video sadržaja

Naziv alata	Poveznica na alat
Lumen5	<a href="https://lumen5.com/">https://lumen5.com/</a>
Visla	<a href="https://www.visla.us/">https://www.visla.us/</a>
Runway	<a href="https://runwayml.com/">https://runwayml.com/</a>
Sora	<a href="https://www.openaisora.video/">https://www.openaisora.video/</a>
InVideo AI	<a href="https://invideo.io/">https://invideo.io/</a>
Synthesia	<a href="https://www.synthesia.io/">https://www.synthesia.io/</a>
Pictory	<a href="https://pictory.ai/?el=2000&amp;htrafficsource=pictoryblog">https://pictory.ai/?el=2000&amp;htrafficsource=pictoryblog</a>
Deepbrain AI	<a href="https://www.deepbrain.io/?r=0">https://www.deepbrain.io/?r=0</a>
Colossyan	<a href="https://www.colossyan.com/ai">https://www.colossyan.com/ai</a>
Fliki	<a href="https://fliki.ai/">https://fliki.ai/</a>

Alati koji će biti opisani u sklopu ovog rada su: InVideo AI, Lumen5, Visla, Runway i Sora. Prva četiri alata odabrana su jer su vrlo jednostavna za korištenje i mogu se koristiti besplatno uz određena ograničenja, dok neki od ostalih alata ne nude mogućnost besplatnog korištenja. Iako još nije dostupna javnosti, Sora će također biti opisana u sklopu ovog rada. Sora može generirati videozapise koji su vrlo realistični tako da se vrlo teško može raspoznati da je riječ o videozapisu generiranom pomoću umjetne inteligencije, a ne stvarnom videozapisu.

### 5.1. InVideo AI

InVideo AI je alat umjetne inteligencije za izradu video sadržaja na temelju tekstualnog unosa, a namijenjen je stvaranju marketinških videa i videa s objašnjavanjem. Pruža velik broj

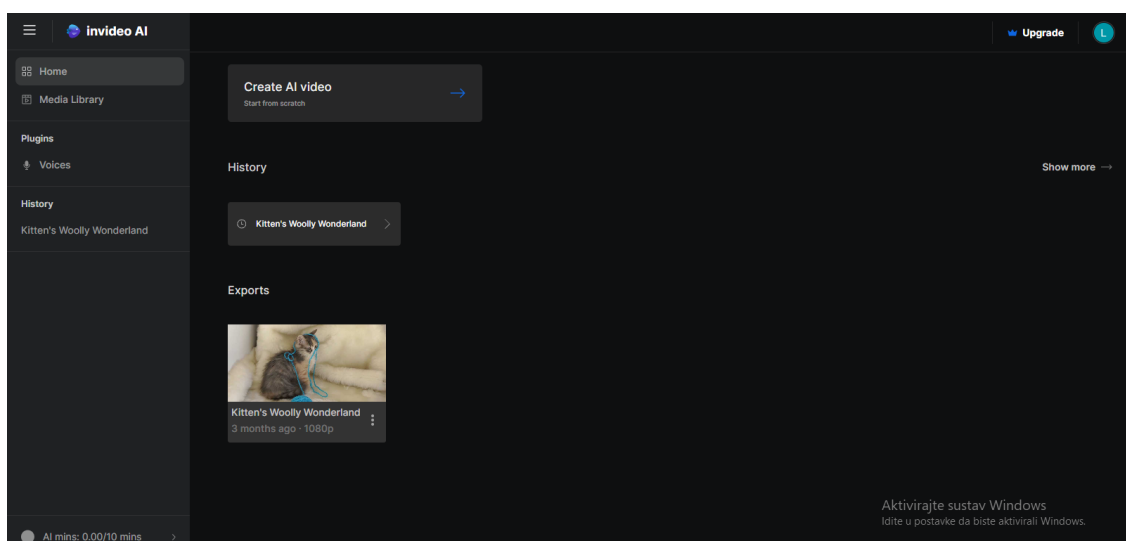


profesionalno dizajniranih predložaka, a za stvaranje videozapisa nije potrebno prethodno iskustvo. Dovoljno je da korisnik unese željeni tekst i zatim odabere ili prilagodi predložak nakon čega može preuzeti video i dijeliti ga na platformama kao što Facebook, Instagram i YouTube (Santiago, 2024). Neke od značajki alata InVideo AI su sljedeće: audio alati, medijska biblioteka, mogućnost podešavanja brzine, mogućnost dijeljenja/spajanja, podržava HD rezoluciju, posebni tekstualni okviri, glazba bez autorskih prava i neograničena duljina znakova (McCallister, 2023). InVideo AI može se koristiti besplatno uz određena ograničenja kao što je eksportiranje s vodenim žigom i nemogućnost korištenja glasovnih klonova. Detaljni cjenovni planovi prikazani su u tablici 3. Na slici 12 prikazan je početni ekran alata InVideo AI.

Tablica 3: Cjenovni planovi za InVideo AI

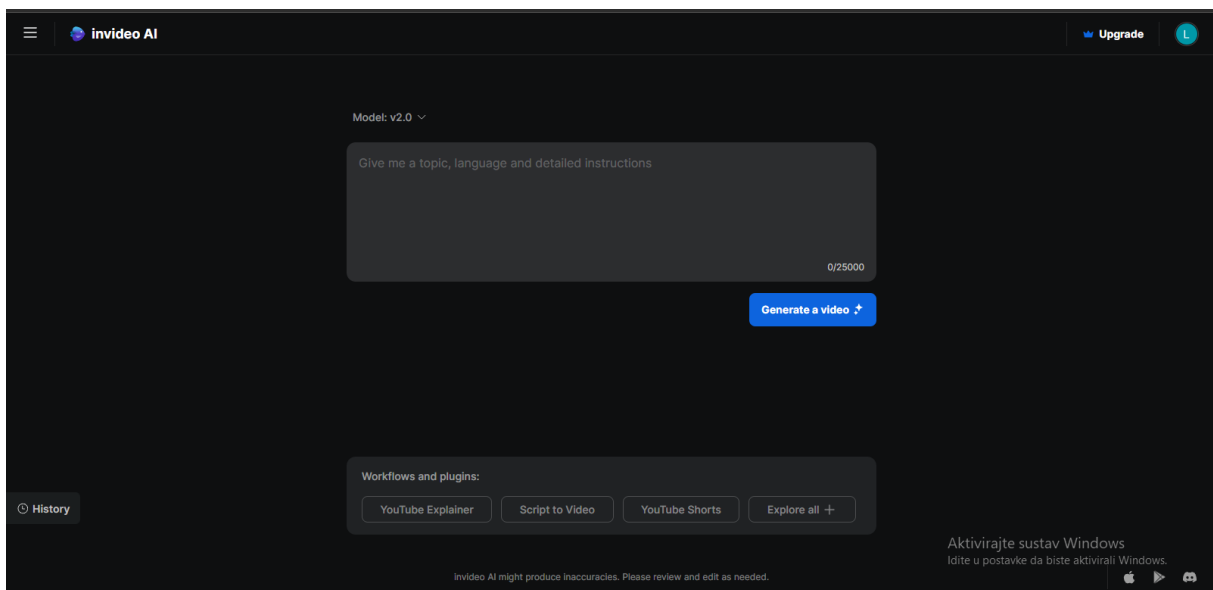
Free	Plus	Max
\$0	\$20	\$48
10 min/tj. AI generiranja	50 min/mj. AI generiranja	200 min/mj. AI generiranja
Nema iStock	80/mj. iStock	320/mj. iStock
10 GB prostora za pohranu	100 GB prostora za pohranu	400 GB prostora za pohranu
4 izvoza/tj. s InVideo logom	Neograničeni izvozi	Neograničeni izvozi
2.5M + standardni mediji	2 glasovna klona	5 glasovnih klonova
Bez glasovnih klonova	Do 1 korisnik	Do 1 korisnik

(Prema: InVideo, bez dat.)

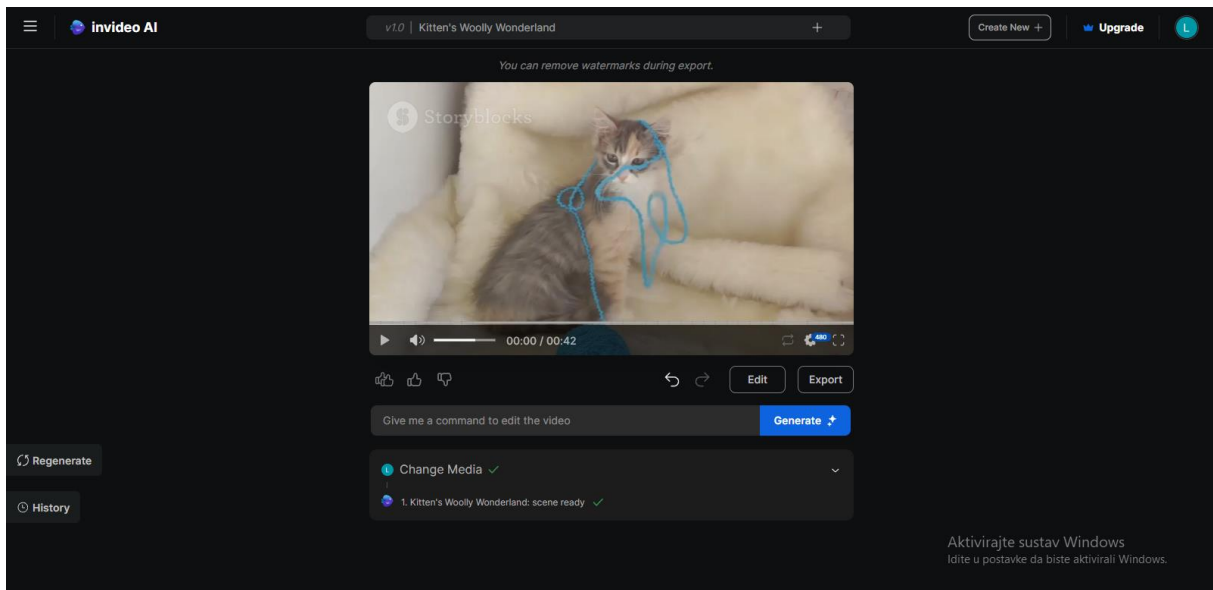


Slika 12: Izgled početnog ekrana alata InVideo AI (Izvor: autorski primjer)

Jedna od prednosti alata InVideo AI je jednostavno korisničko sučelje prilagođeno korisnicima koji se s lakoćom mogu kretati kroz platformu. Sučelje je također prilagođeno početnicima tako da i oni mogu stvarati videozapise. Velika prednost je i cjelodnevna korisnička podrška. Tim za podršku dostupan je u bilo kojem trenutku kako bi riješio tehničke probleme ili dao smjernice za korištenje određenih značajki. Međutim, InVideo AI ima i određene nedostatke. Jedan od nedostataka je ogroman broj dostupnih mogućnosti pa se korisnici mogu osjećati preopterećeno prilikom odabira određenih postavki i značajki što također može usporiti proizvodnju videozapisa. Još jedan nedostatak je i taj što nije moguće izravno objaviti video na društvenim medijima. Korisnici moraju najprije preuzeti videozapis i tek onda mogu objaviti videozapis na platforme društvenih medija što usporava proces objave i uključuje dodatne korake za korisnike. Još jedan od problema je i mogućnost nedosljednog prikazivanja videa u smislu brzine i kvalitete pa je moguće da neki rezultati neće biti zadovoljavajući (Harry, 2024). Na slici 13 možemo vidjeti sučelje za unos teksta na temelju kojega InVideo generira videozapis, a na slici 14 sučelje za uređivanje generiranog videozapisa.



Slika 13: Sučelje za unos teksta u alatu InVideo AI (Izvor: autorski primjer)



Slika 14: Sučelje za uređivanje generiranog videa u alatu InVideo AI (Izvor: autorski primjer)

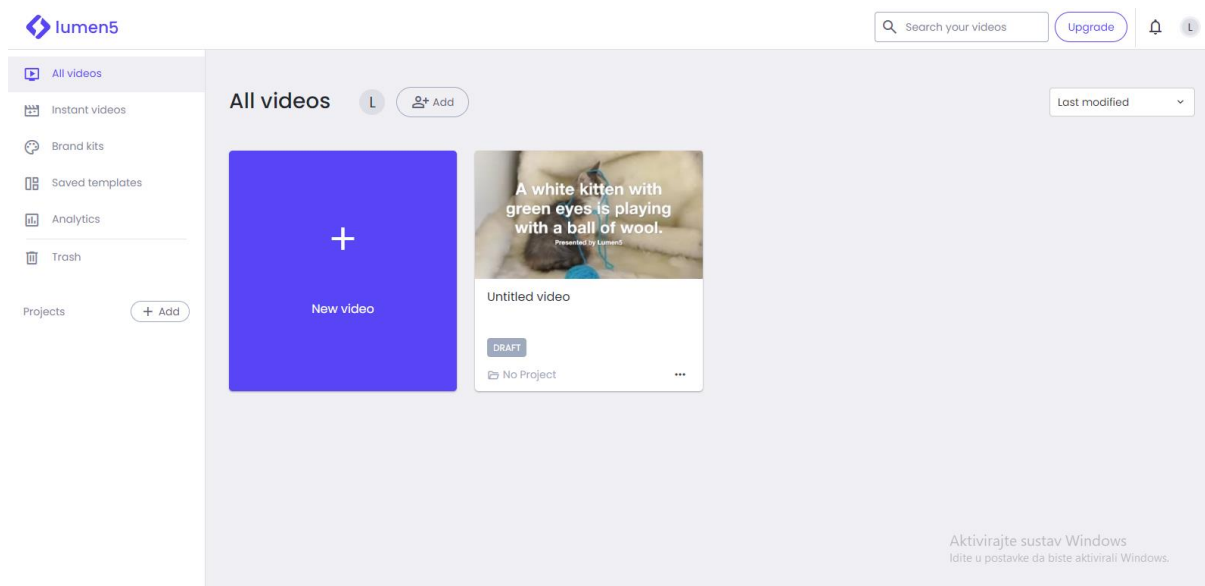
## 5.2. Lumen5

Lumen5 je alat umjetne inteligencije za izradu video sadržaja koji također generira videozapis na temelju tekstualnog unosa. Za korištenje alata Lumen5 nije potrebno prethodno iskustvo uređivanja i stvaranja videozapisa. Lumen5 se zasniva na algoritmima umjetne inteligencije koje su obučeni na velikim količinama podataka kako bi mogli generirati videozapise u skladu s tekstualnim unosom. Jedna od značajki ovog alata je biblioteka s video isječcima i slikama koje korisnici mogu besplatno koristiti, kao i mogućnost prenošenja vlastitih slika na servis (Jalli, 2023). Osim toga, Lumen5 nudi mogućnost dodavanja glasa videozapisima, a ima i veliku zbirku predložaka s različitim dimenzijama za sadržaj društvenih medija. Također postoji mogućnost dodavanja različitih animacija za tekstualne efekte i prilagodba duljine zaslona, pozicije i veličine teksta, promjena fonta i boje („Lumen5 review: Is this Video Maker Right for you? Features, Pro’s & Con’s“ , bez dat.). Ovaj alat nudi besplatni plan koji omogućuje stvaranje do pet videa mjesečno, a videozapisi sadrže vodeni žig. Detaljniji cjenovni planovi prikazani su u tablici 4. Na slici 15 prikazan je izgled početnog ekrana alata Lumen5.

Tablica 4: Cjenovni planovi za Lumen5

Community	Basic	Starter	Professional	Enterprise
\$0	\$19	\$59	\$149	Prilagođena cijena
Sadrži Lumen5 vodeni žig	Sve iz Community paketa	Sve iz Basic paketa	Sve iz Starter paketa	Sve iz Professional paketa
720p video rezolucija	Bez Lumen5 brendiranja	1080p video rezolucija	500M stock fotografija i videozapisa	Lumen5 dizajnerski tim
Do pet videozapisa mjesečno	Napredne značajke AI Script Composer-a	50M+ stock fotografija i videozapisa	Prilagođeni vodeni žigovi	Prilagođeni brendirani predlošci
2 minute AI govorne sinteze	Duži AI Voiceover videozapisi i više glasova	Odabir vlastitih fontova	Učitavanje vlastitih fontova	Posvećena podrška za uspjeh korisnika
		Prilagodljive boje	Više spremljenih predložaka	Timovi i suradnja

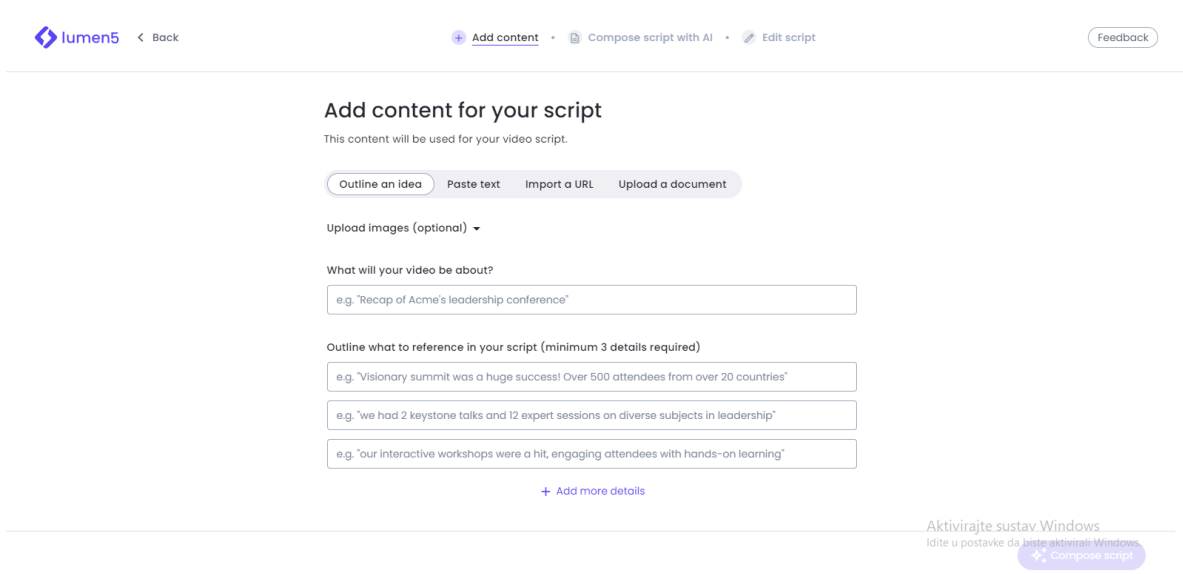
(Prema: Lumen5, bez dat.)



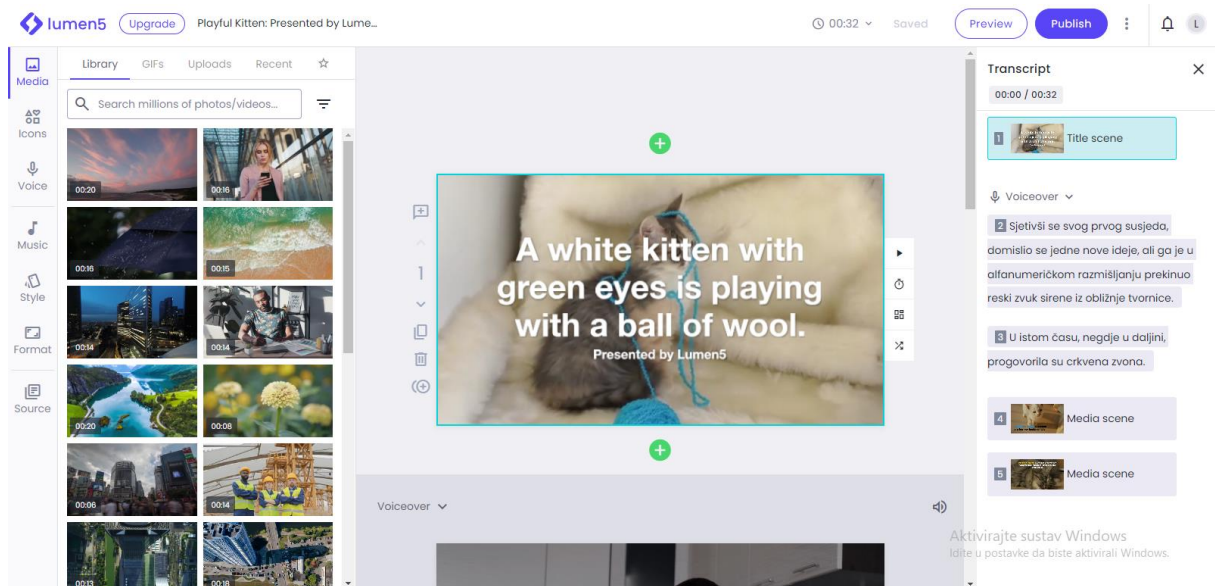
Slika 15: Izgled početnog ekrana alata Lumen5 (Izvor: autorski primjer)

U tekstu („Lumen5“ , bez dat.) opisane su prednosti i nedostaci alata Lumen5. Kao prednost navedeno je da je alat jednostavan za korištenje i ne zahtijeva prethodno iskustvo u obradi videa, sposoban je automatski generirati zadatke iz teksta ili web stranica, ima veliku biblioteku s video isječcima, slikama i predlošcima, korisnici imaju mogućnost prilagođavanja videozapisa dodavanjem različitih elemenata i funkcija i omogućuje stvaranje videozapisa za različite platforme. S druge strane, nije u potpunosti besplatan, već nudi različite planova cijena ovisno o dostupnim značajkama, ne pruža dovoljno kreativne kontrole za detaljnije uređivanje videozapisa i moguće je da nije kompatibilan s nekim uređajima i preglednicima koje bi neki korisnici htjeli koristiti za gledanje videozapisa. Na slici 16 možemo vidjeti sučelje alata Lumen5

za unos potrebnih informacija o videu koji će se generirati, a na slici 17 sučelje za uređivanje generiranog videa.



Slika 16: Sučelje alata Lumen5 za unos informacija o videu (Izvor: autorski primjer)



Slika 17: Sučelje alata Lumen5 za uređivanje generiranog videa (Izvor: autorski primjer)

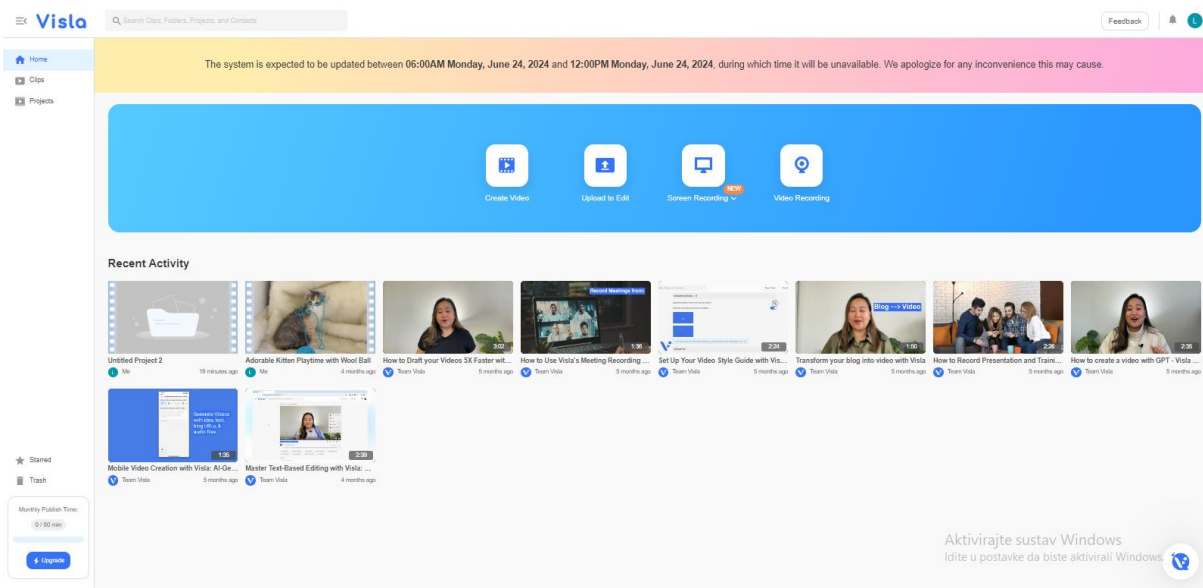
### 5.3. Visla

Visla je alat umjetne inteligencije za izradu raznolikog video sadržaja, kao što su videozapisi za društvene mreže, glasovne snimke i dr. Visla radi na način da korisnici unesu svoju ideju na temelju koje će Visla generirati videozapis. Nakon toga, korisnici mogu uređivati generirani videozapis prema vlastitim željama. Visla nudi i mogućnost uređivanja vlastitih videozapisa rezanjem ili spajanjem video isječaka, preuređivanjem scena, promjenom pozadinske glazbe i dr. Osim generiranja i uređivanja videa, Visla nudi i mogućnost dodavanja glasa videozapisima, pretvorbu audio sadržaja u video, opsežnu videoteku i preporuke za snimanje (Rivers, 2024). Moguće je također i izravno snimanje s vlastite kamere na platformu. Isto tako, Visla ima mogućnost izrezivanja dopunskih riječi, odnosno poštapalica kao što su „hm“ i „ah“ (Rebelo, 2023). Visla se može koristiti besplatno uz 10 GB prostora za pohranu i objavljivanje videozapisa u trajanju do 50 minuta. Detaljniji cjenovni planovi prikazani su u tablici 5. Na slici 18 možemo vidjeti kako izgleda početni ekran u alatu Visla.

Tablica 5: Cjenovni planovi za Vislu

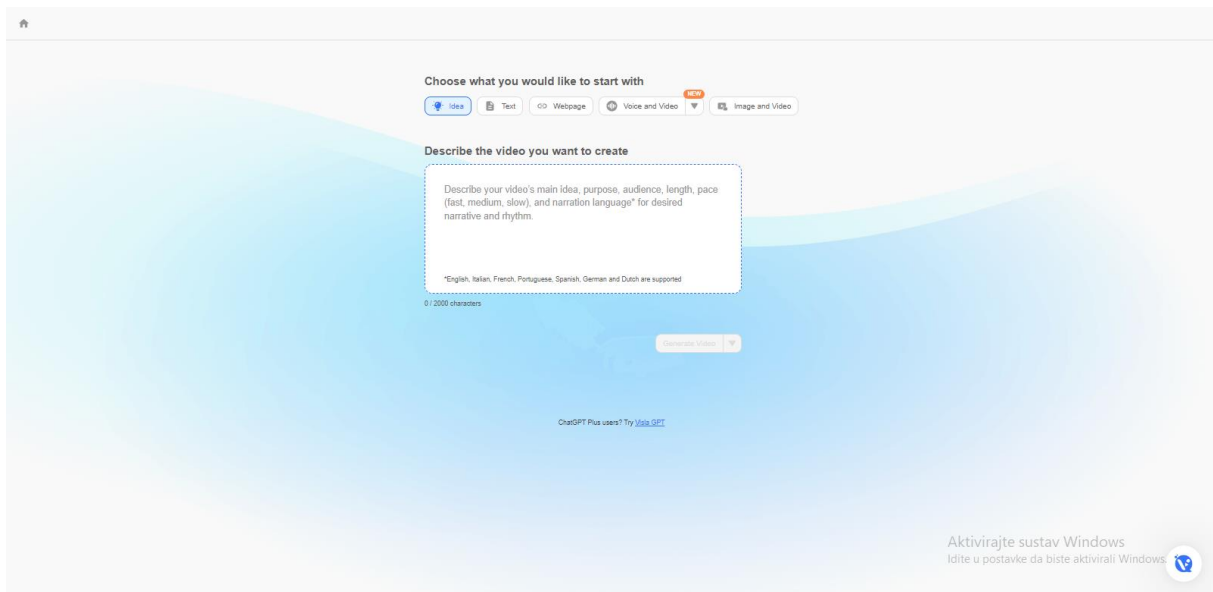
Free	Pro	Business	Enterprise
\$0	\$9	\$39	Prilagođena cijena
1000 mjesečno dodijeljenih kredita	Mjesečno dodijeljeni krediti počinju od 5000	Mjesečno dodijeljeni krediti počinju od 20 000	Prilagođeno dodjeljivanje mjesečnih kredita
Bez pristupa sadržajima iz Premium stock biblioteke	Pristup 2 milijuna sadržaja iz Premium stock biblioteke	Pristup više od 16 milijuna sadržaja iz Premium stock biblioteke	Pristup više od 16 milijuna sadržaja iz Premium stock biblioteke
Maksimalna rezolucija: 1080p	Maksimalna rezolucija: 1080p	Maksimalna rezolucija: 4k	Maksimalna rezolucija: 4k
Učitavanje medija: 3 GB/sat	Učitavanje medija: 3 GB/sat	Učitavanje medija: 20 GB/3 sata	Učitavanje medija: 20 GB/3 sata

(Prema: Visla, bez dat.)

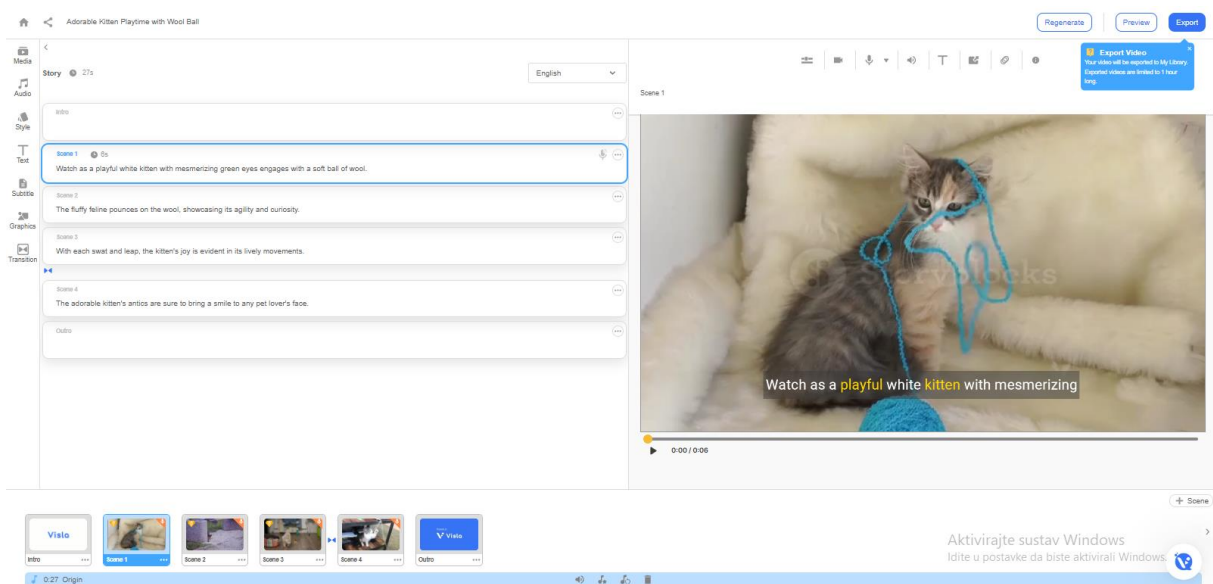


Slika 18: Izgled početnog ekrana alata Visla (Izvor: autorski primjer)

Rivers (2024) navodi određene prednosti i nedostatke alata Visla. Jedna od prednosti Visle je stvaranje profesionalnog videozapisa samo na temelju ideje ili scenarija. Visla također ima mogućnost generiranja skripte, dodavanja naracije i automatskog uklanjanja neželjenog sadržaja. Prednost Visle je i to što je jednostavna za korištenje pružajući intuitivno korisničko sučelje. Visla također nudi manje ograničenja u sklopu besplatnog plana u odnosu na konkurente pa tako korisnici mjesečno mogu objavljivati do čak 50 minuta video sadržaja. Visla ima i određene nedostatke. Jedan od nedostataka je taj što Visla nema mogućnost generiranja slika pomoću umjetne inteligencije, pa će korisnici kojima je to potrebno morati koristiti drugi generator slika i onda prenijeti svoje slike na Vislu. Još jedan od nedostataka je i dugo trajanje izvoza, preuzimanja i generiranja videozapisa. Kao posljednji nedostatak navedene su i moguće pogreške umjetne inteligencije, iako ovaj problem nije vezan izravno za Vislu, s obzirom na to da se taj problem može javiti i kod drugih AI video generatora. Na slici 19 prikazano je sučelje alata Visla za unos informacija o videu koji korisnik želi generirati, a na slici 20 sučelje za uređivanje videa u alatu Visla.



Slika 19: Sučelje alata Visla za unos informacija o videu (Izvor: autorski primjer)



Slika 20: Sučelje alata Visla za uređivanje generiranog videa (Izvor: autorski primjer)

## 5.4. Runway

Runway je alat umjetne inteligencije koji koristi strojno učenje za proces generiranja i uređivanja videozapisa i slika. Jedna od značajki Runwaya je generiranje videozapisa na temelju tekstualnih upita, a osim teksta, kao ulaz je moguće koristiti i slike. Runway također

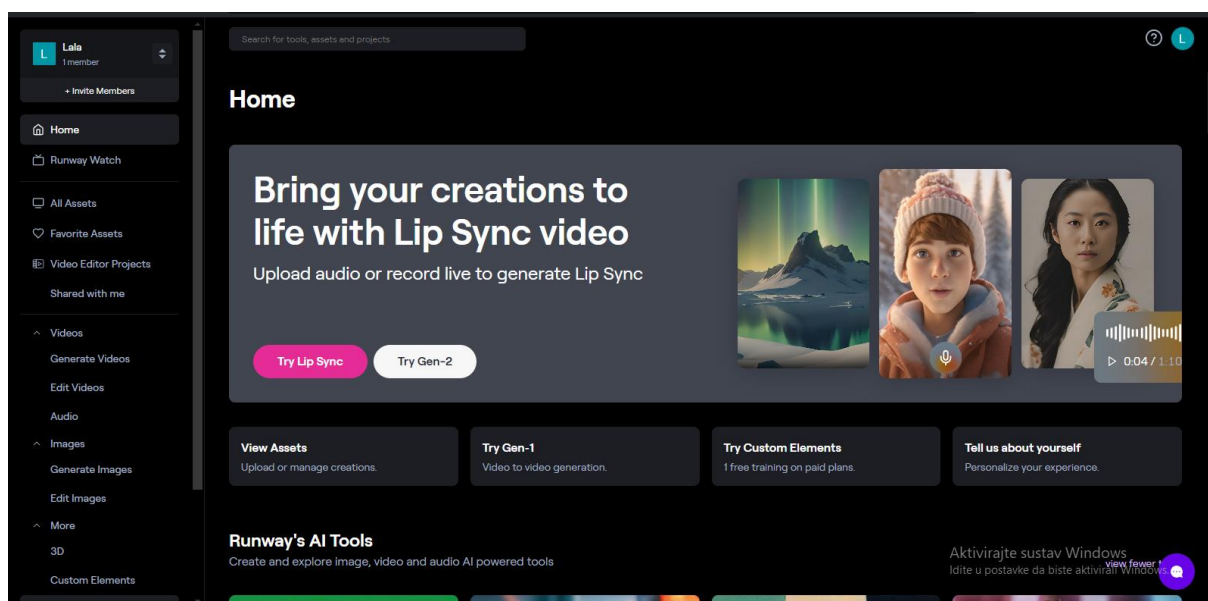


omogućuje stvaranje slika visoke rezolucije korištenjem modela latentne difuzije. Također je omogućeno uređivanje videozapisa dodavanjem specijalnih efekata, prijelaza, animacija i dr. Isto tako, Runway nudi velik broj alata unutar aplikacije koji se mogu koristiti za uklanjanje pozadine, usporavanje videa, pomicanje slika itd. Osim videa, Runway također pruža mogućnost generiranja slika iz tekstualnih upita. Još neke značajke ovog alata su i praćenje kretanja objekata u videozapisima, prijenos stila jedne slike na drugu sliku i procjena dubine slike što se može koristiti za stvaranje 3D efekata (Vikram, 2024). Runway nudi besplatni plan s određenim ograničenjima. Na primjer, u sklopu besplatnog plana moguće je generirati video iz videa u maksimalnom trajanju od četiri sekunde, dok video generiran iz teksta može trajati do 16 sekundi u sklopu proširenja videa. Detaljniji cjenovni planovi prikazani su u tablici 6. Nedavno je Runway predstavio svoj novi AI model Gen-3 Alpha koji generira video isječke iz tekstualnih unosa i fotografija, a u odnosu na prethodni model Gen-2 donosi poboljšanja u brzini generiranja i vjernosti (Wiggers, 2024). Na slici 21 možemo vidjeti kako izgleda početni ekran alata Runway.

Tablica 6: Cjenovni planovi za Runway

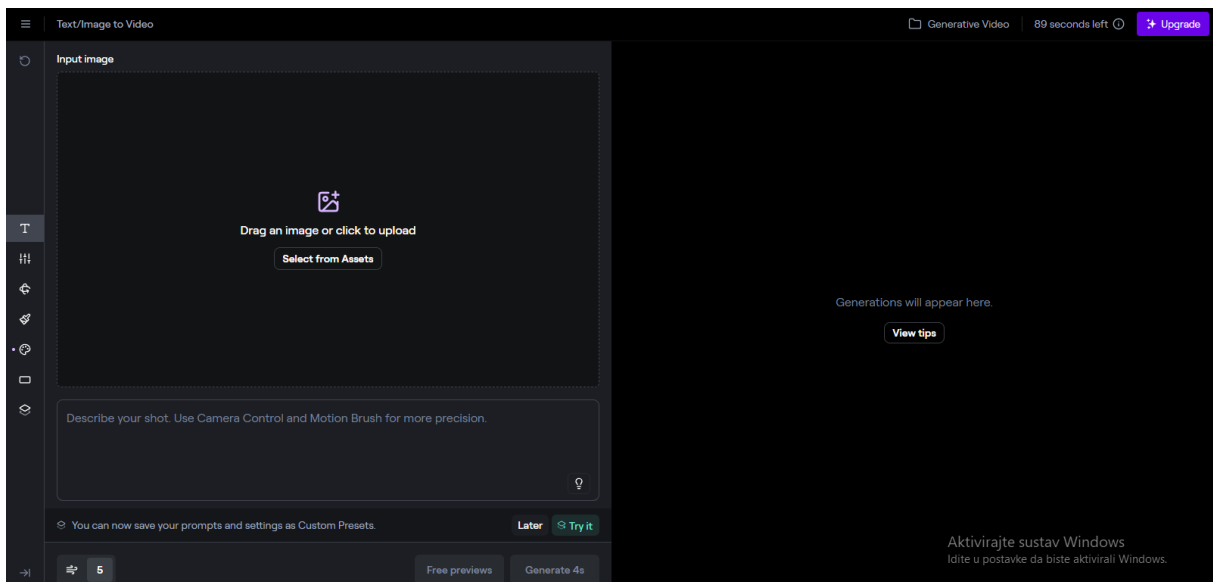
Basic	Standard	Pro	Unlimited	Enterprise
\$0	\$12	\$28	\$76	Prilagođena cijena
Ne mogu se kupiti dodatni krediti	Obnavljanje kredita na 625 svaki mjesec	Obnavljanje kredita na 2250 svaki mjesec	Sve značajke Pro plana	Sve značajke Pro plana
Ne može se povećati rezolucija ni ukloniti vodeni žigovi u Gen-1 i Gen-2	Gen-3 Alpha (tekst i slika u video) do 10 sekundi	Gen-3 Alpha (tekst i slika u video) do 10 sekundi	Neograničeno generiranje Gen-1, Gen-2 i Gen-3 Alpha u Explore modu	Jedinstveno prijavljivanje
Gen-1 (video u video) do 4 sekunde	Povećanje rezolucije u Gen-1 i Gen-2	Povećanje rezolucije u Gen-1 i Gen-2	Obnavljanje kredita na 2250 svaki mjesec s mogućnošću kupnje dodatnih kredita po potrebi	Prilagođeni iznosi kredita
Gen-2 (tekst i slika u video) do 16 sekundi putem opcije „Extend Video“	Gen-1 (video u video) do 15 sekundi	Gen-1 (video u video) do 15 sekundi		Prilagođeni prostor za pohranu
3 video projekta	Gen-2 (tekst i slika u video) do 16 sekundi putem opcije „Extend Video“	Gen-2 (tekst i slika u video) do 16 sekundi putem opcije „Extend Video“		Prilagodbe modela
5 GB resursa	Neograničeni broj video projekata	Neograničeni broj video projekata		Napredna sigurnost i usklađenost
Izvozi u 720p rezoluciji	100 GB resursa	500 GB resursa		Program stalnog uspjeha
Ograničene opcije izvoza slika	Izvozi u 4K rezoluciji	Obuka prilagođenih AI generatora		Prioritetna podrška

(Prema: Runway, bez dat.)



Slika 21: Izgled početnog ekrana alata Runway (Izvor: autorski primjer)

Kao i ostali alati, i Runway ima određene prednosti i nedostatke. Jedna od prednosti Runwaya je sučelje prilagođeno korisniku pa ga vrlo jednostavno mogu koristiti i početnici. Prednost je i pristup unaprijed obučanim modelima što omogućuje korisnicima obavljanje raznih zadataka, kao što je generiranje slika, prijenos stilova itd. Runway također nudi povratne informacije u stvarnom vremenu i kompatibilan je s više platformi pa je dostupan široj publici. Jedna od prednosti je i redovito ažuriranje i poboljšanje što omogućava korisnicima da pristupe najnovijim značajkama umjetne inteligencije. Jedan od nedostataka Runwaya je cijena jer u sklopu besplatnog plana korisnicima je ograničen pristup naprednim značajkama i modelima. Jedan od problema može biti i zahtijevanje velikog broja resursa za pojedine modele umjetne inteligencije što može usporiti računalo. Nedostatak je i mogućnost ograničene prilagodbe pa bi nekim korisnicima moglo biti teško prilagoditi i podesiti model svojim zahtjevima (Shiksha, 2024).



Slika 22: Sučelje alata Runway za generiranje videa iz teksta ili slike (Izvor: autorski primjer)

## 5.5. Sora

Prema (Bharat, 2024) Sora je OpenAI-jev alat za generiranje videa iz slike ili teksta. Sora može generirati videozapise koji traju do jedne minute na temelju tekstualnog unosa. Kako bi generirala video, Sora koristi transformatorsku arhitekturu baziranu na difuziji. Radi na način da se video podaci raščlanjuju na okvire pri čemu se svaki okvir rastavlja na grupe piksela, a Sora bilježi vremenske informacije piksela. Sora koristi varijacijski autokoder (eng. Variational Autoencoder - VAE) i sažima neobrađeni video u latentni prikaz koji pohranjuje prostorne i

vremenske informacije. Sora koristi prostorne-vremenske zakrpe koje omogućuju rad s videozapisima različitih razlučivosti, duljina i omjera širine i visine. Sora transformira sve oblike vizualnih podataka u jedinstveni prikaz. Videozapisi se komprimiraju u latentne prostore i razgrađuju u prostorno-vremenske zakrpe.

Cotton (2024) navodi određena ograničenja Sore. Jedno od ograničenja je da Sora nema implicitno razumijevanje fizike pa se nekad ne pridržava fizičkih pravila stvarnog svijeta. Model ne razumije uzrok i posljedicu, a prostorni položaj objekata može se neprirodno pomaknuti. Osim toga, u pitanje se dovodi i pouzdanost Sore. Nije poznato koliko je videozapisa bilo potrebno generirati kako bi se dobio jedan upotrebljiv videozapis, ali ako je taj broj velik, to bi mogla biti velika prepreka. Odgovor na ovo pitanje ćemo dobiti tek kada alat postane široko dostupan. Na slici 23 prikazana je snimka zaslona iz videa koji je generirala Sora koji prikazuje ženu kako hoda ulicom Tokija.



Slika 23: Žena hoda ulicom Tokija u videu generiranom pomoću Sore

Jasaitis (2024) opisuje prednosti i nedostatke Sore. Jedna od prednosti je napredna tehnologija koja omogućuje stvaranje uvjerljivih videozapisa kinematografske kvalitete. Još jedna od prednosti je i ta što OpenAI koristi sigurnosne mjere i suradnju sa stručnjacima u području dezinformacija, govora mržnje i pristranosti kako bi proveli opsežna istraživanja Sore što uključuje dodavanje vodenih žigova svim videozapisima generiranim umjetnom inteligencijom. Osim toga, prednost su i personalizirani videozapisi koji značajnu povećavaju angažman i utjecaj. Sora ima i određenih nedostataka koje je potrebno navesti. Jedan od nedostataka je taj što postoji rizik zlouporabe što može dovesti do etičkih problema. Problem

je i taj što su stručnjaci u područjima kao što su video produkcija i grafički dizajn u opasnosti od gubitka posla jer Sora može preuzeti njihove poslove. Na kraju, pretjerano korištenje umjetne inteligencije može ugušiti ljudsku kreativnost i kritičko razmišljanje pa treba uspostaviti ravnotežu između ljudskog doprinosa i doprinosa umjetne inteligencije.

## 6. Izrada video sadržaja u alatu Visla

Izrada video sadržaja pomoću umjetne inteligencije bit će demonstrirana u alatu Visla. Ovaj alat odabran je zbog toga što je vrlo jednostavan za korištenje i nudi veći raspon mogućnosti generiranja video sadržaja u sklopu besplatnog plana u usporedbi s ostalim alatima. Visla nudi mogućnost generiranja video sadržaja na sljedeće načine: na temelju ideje, na temelju tekstualnog unosa, korištenjem URL-a web stranice, na temelju audiozapisa i na temelju slike. U ovom poglavlju bit će prikazani svi ovi načini generiranja video sadržaja, a zatim će biti analizirana njihova uspješnost.

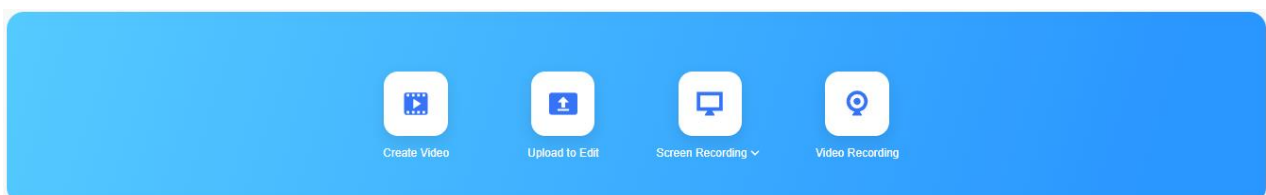
### 6.1. Korisničko sučelje alata Visla

Nakon što se prijavimo sa svojim korisničkim računom za Vislu ili svojim Gmail računom, otvara se početni ekran. Na početnom ekranu imamo četiri glavne opcije koje također možemo vidjeti i na slici 24:

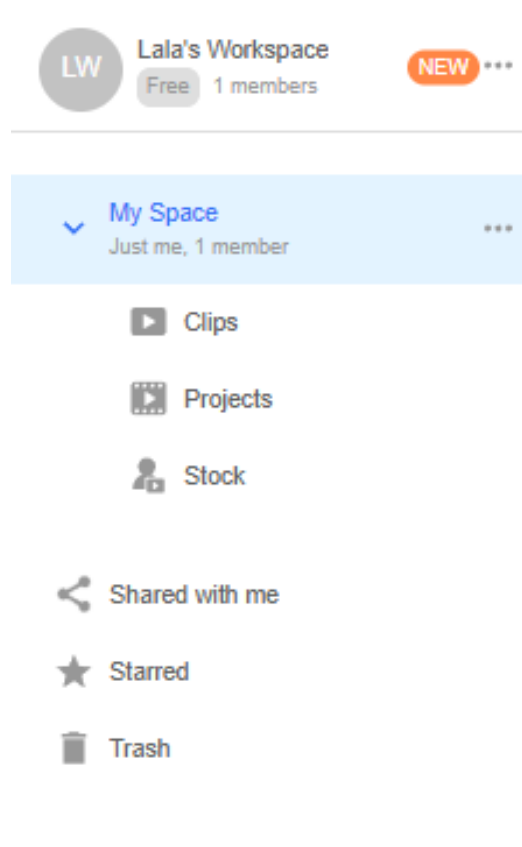
- Kreiranje videa (*eng. Create video*)
- Učitavanje videa za uređivanje (*eng. Upload to Edit*)
- Snimanje ekrana (*eng. Screen Recording*)
- Snimanje videa (*eng. Video Recording*)

Ispod ovih opcija nalaze se videozapisi koje je korisnik nedavno kreirao ili uređivao. S lijeve strane nalazi se bočna traka sa sljedećim opcijama prikazana na slici 25:

- Isječci (*eng. Clips*)
- Projekti (*eng. Projects*)
- Stock sadržaji (*eng. Stock*)
- Podijeljeno sa mnom (*eng. Shared with me*)
- Označeno zvjezdicom (*eng. Starred*)
- Koš za smeće (*eng. Trash*)

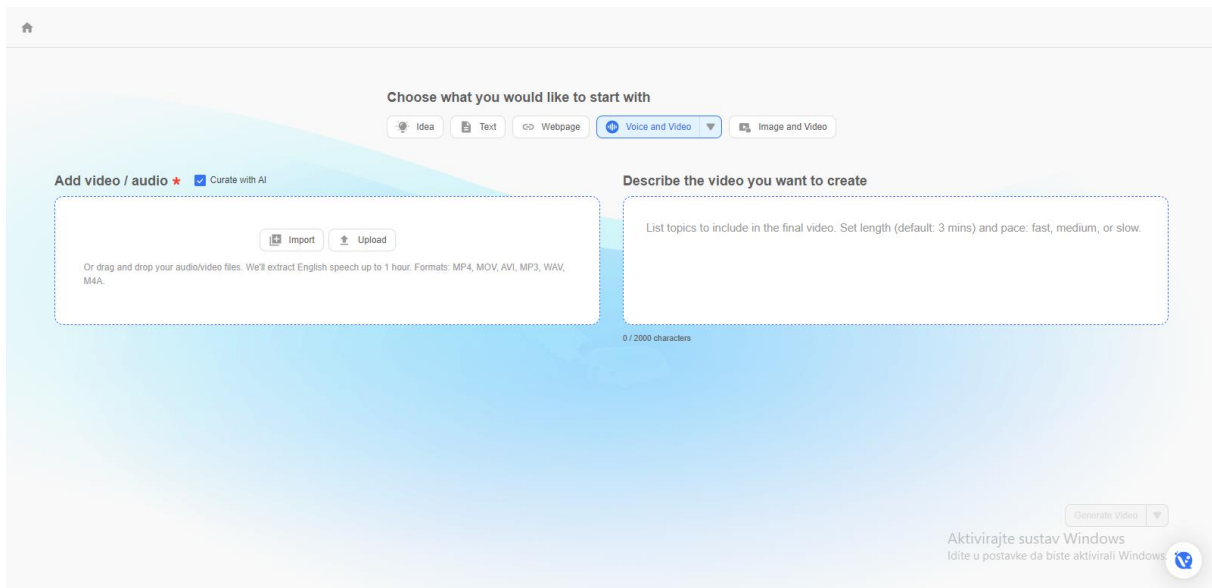


Slika 24: Četiri glavne opcije na početnom ekranu Visle (Izvor: autorski primjer)



Slika 25: Bočna traka početnog ekrana Visle (Izvor: autorski primjer)

Opcija koja će biti potrebna za generiranje video sadržaja je opcija *Create video* pa kada kliknemo tu opciju, otvara nam se sučelje s tabovima koji nam omogućuju da odaberemo vrstu unosa na temelju kojeg će se video generirati. Za unos teksta dostupni su jednostavni tekstualni okviri, a postoje i gumbi za učitavanje sadržaja s računala. Na slici 26 vidimo sučelje za odabir i unos odgovarajuće vrste upita za generiranje video sadržaja.



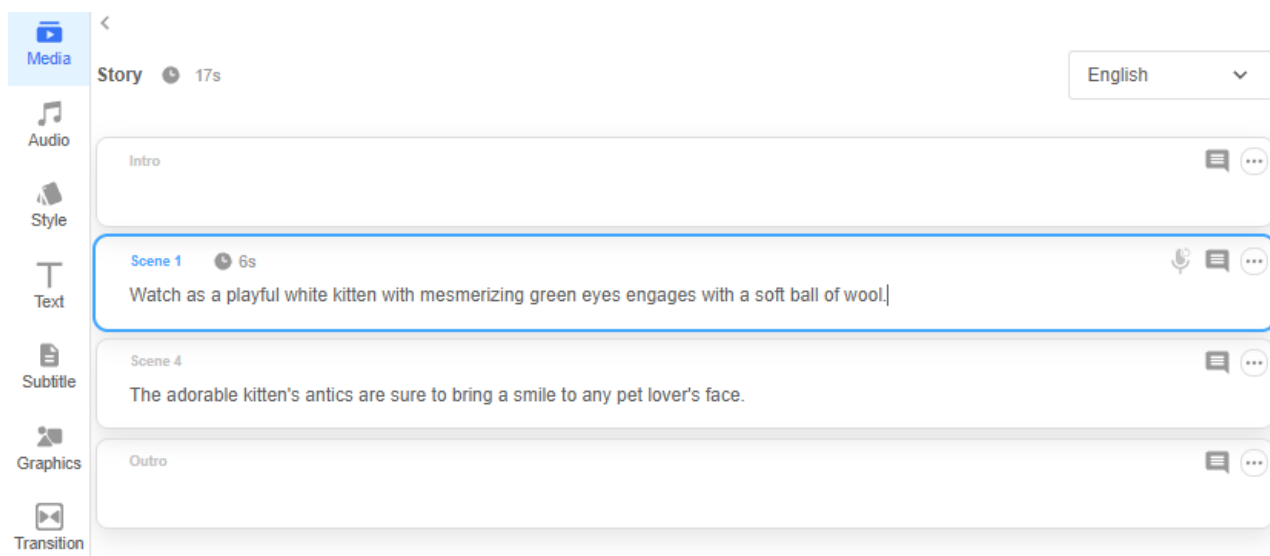
Slika 26: Odabir i unos željenog upita za generiranje video sadržaja u Visli (Izvor: autorski primjer)

Sučelje za uređivanje videa dijeli video na scene tako da korisnik može uređivati svaku scenu zasebno što je vidljivo na slici 27. S lijeve strane nalazi se bočna traka s mogućnostima uređivanja videozapisa. Jedna od mogućnosti je dodavanja stock videozapisa od kojih su neki besplatni, a za neke je potrebno odabrati neki od planova koji nisu besplatni kako bi bilo moguće njihovo korištenje. Opcija *Audio* nudi mogućnost generiranja glasovne naracije pomoću umjetne inteligencije, kao i učitavanje zvučne datoteke s računala. Također je dostupna i opcija upravljanja s pozadinskom glazbom. Opcija *Text* nudi mogućnost odabira stila teksta, pozicije teksta, kao i veličine teksta, a iste mogućnosti su dostupne i za uređivanje titlova. Alat još nudi i opciju dodavanja različitih grafičkih oblika, kao i dodavanje tranzicija. Svaka scena se još može naknadno promijeniti promjenom teksta u tekstualnom okviru za pojedinu scenu što možemo vidjeti na slici 28.



Slika 27: Podjela videa na scene (Izvor: autorski primjer)



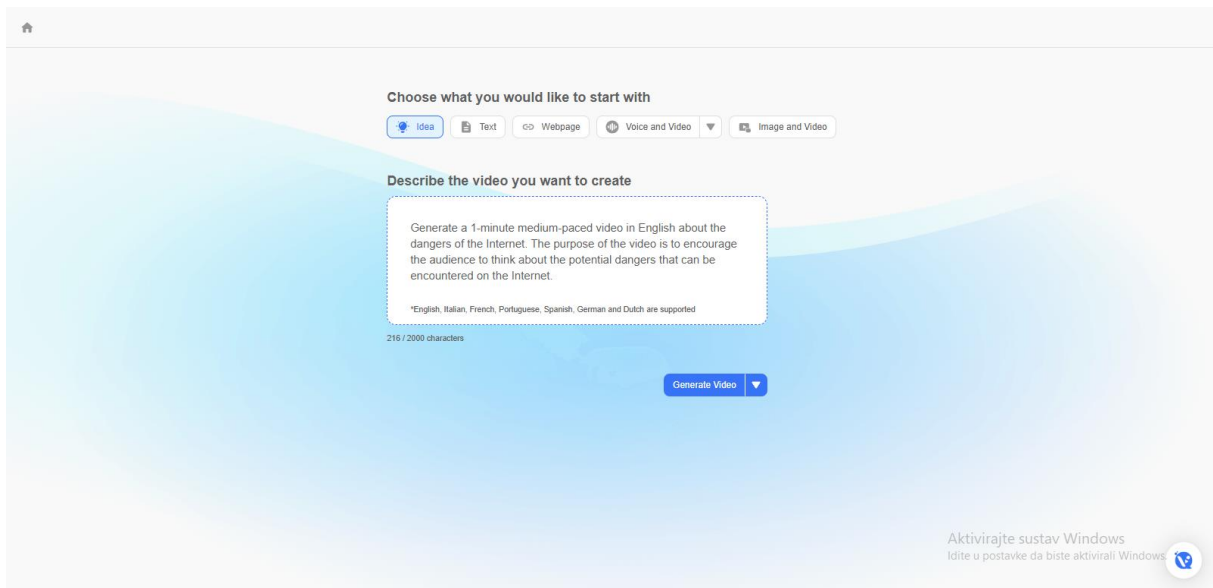


Slika 28: Tekstualni okviri za svaku scenu (Izvor: autorski primjer)

## 6.2. Izrada video sadržaja na temelju ideje

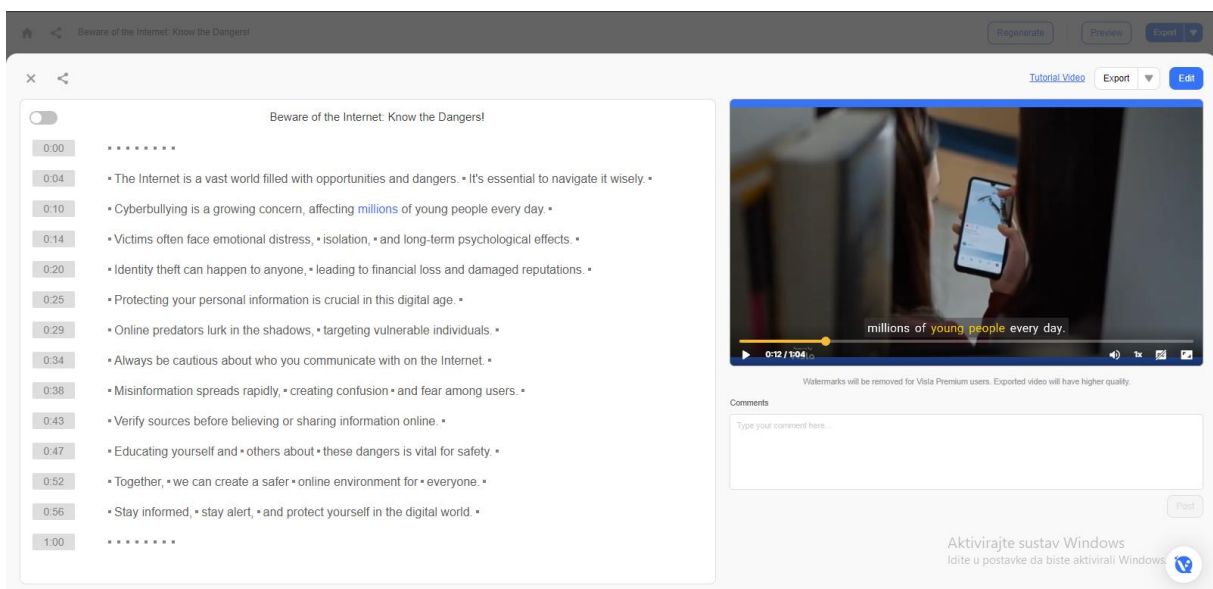
Prvi primjer izrade video sadržaja u alatu Visla će kao unos koristiti ideju i na temelju nje će se generirati videozapis. Kako bismo kreirali video, potrebno je na početnom ekranu odabrati opciju *Create Video*. Nakon toga, otvara se ekran s odabirom vrste upita. Za vrstu upita potrebno je odabrati ideju (opcija *Idea*) i unijeti upit u tekstualni okvir. S obzirom na to da alat ne podržava hrvatski jezik, upit će biti zadan na engleskom jeziku.

Tekst koji će biti unesen kao upit je sljedeći: *Generate a 1-minute medium-paced video in English about the dangers of the Internet. The purpose of the video is to encourage the audience to think about the potential dangers that can be encountered on the Internet.* Ovim upitom je Visli zadano da generira video srednjeg tempa dugačak jednu minutu na engleskom jeziku o opasnostima na internetu. Svrha videa je potaknuti publiku na razmišljanje o potencijalnim opasnostima s kojima se može susresti na internetu. Unos ovog upita možemo vidjeti na slici 29.



Slika 29: Unos upita za izradu video sadržaja na temelju ideje (Izvor: autorski primjer)

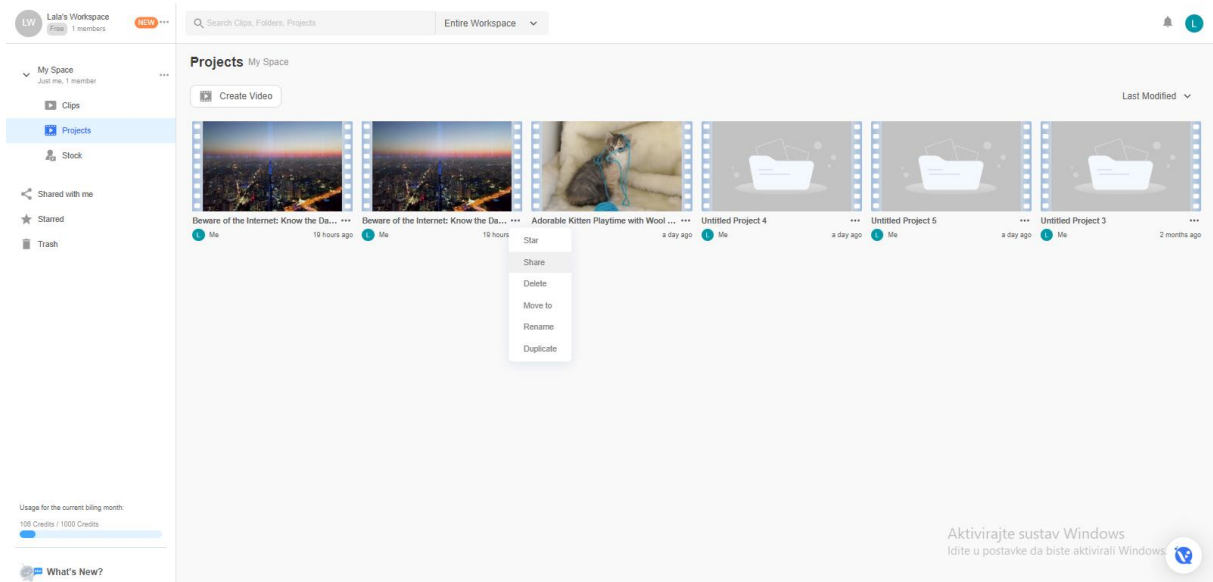
Nakon unosa teksta, potrebno je pritisnuti gumb *Generate Video* za generiranje videa. Nakon određenog vremena generirat će se video i otvorit će se prozor s generiranim videom i scenama od kojih se video sastoji što možemo vidjeti i na slici 30.



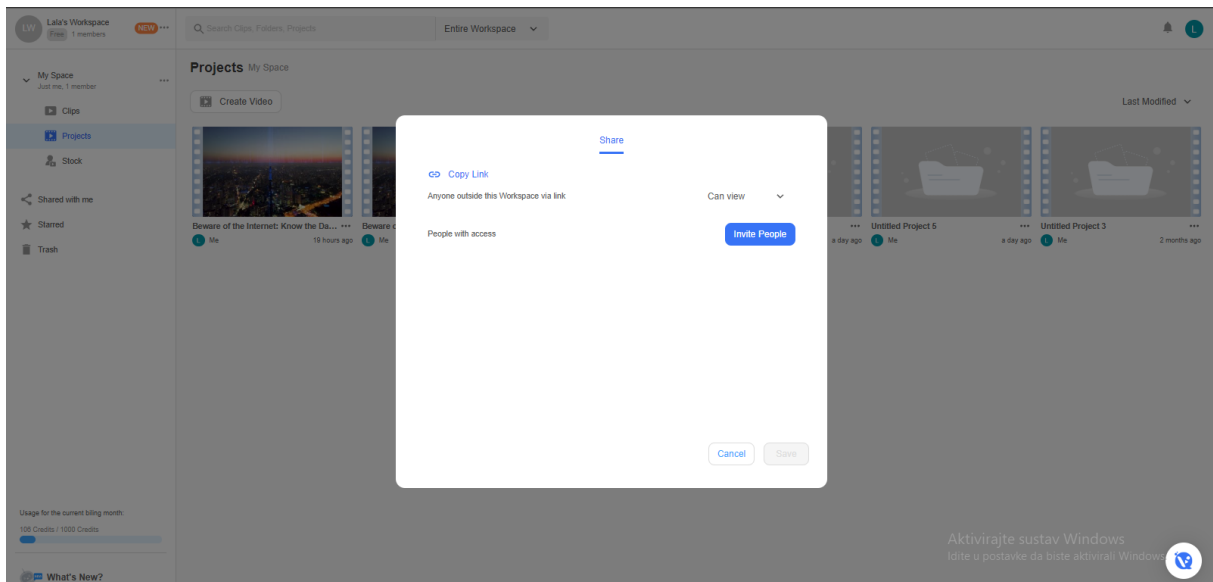
Slika 30: Prikaz generiranog videa u Visli na temelju ideje (Izvor: autorski primjer)

Generirani video može se uređivati pritiskom na gumb *Edit*, a može se i podijeliti tako da ga može pregledati svatko s vezom. Kako bismo podijelili video, potrebno je na početnom

zaslonu odbrati opciju *Projects*, pronaći video koji želimo podijeliti, kliknuti na tri točkice u donjem desnom kutu i odabrati opciju *Share* kao što vidimo i na slici 31. Nakon toga se otvara prozor gdje možemo jednostavno kopirati link i podijeliti ga sa svima za koje želimo da pregledaju video što je prikazano na slici 32.



Slika 31: Opcija *Share* za dijeljenje videa (Izvor: autorski primjer)



Slika 32: Kopiranje poveznice na video (Izvor: autorski primjer)

Na ovoj poveznici nalazi se video koji je generiran na temelju ideje: <https://app.visla.us/project/1254196944843100160>.

### 6.3. Izrada video sadržaja na temelju teksta

Drugi primjer izrade video sadržaja u alatu Visla će kao unos koristiti tekst za generiranje videa. Ova opcija omogućava unos narativnog ili vizualnog scenarija, a postoji i poseban tekstualni okvir za opis videa koji želimo generirati, odnosno za postavljanje tempa videa. Kako bismo kreirali takav video, potrebno je kod odabira vrste upita odabrati *Text* i unijeti vizualni ili narativni scenarij u tekstualni okvir.

Scenarij za kreiranje ovog videa generirao je ChatGPT i nalazi se na sljedećoj poveznici: <https://chatgpt.com/share/277f9994-3819-4264-b261-38dab4a01313>. Na slici 33 možemo vidjeti postavljanje upita za generiranje videa na temelju teksta. Tema ovog videa će također biti opasnosti na internetu, ali ovaj puta se očekuje drugačiji video zbog drugačije postavljene upita. U tekstualni okvir uneseni je sljedeći upit, odnosno scenarij:

*[INTRO - Upbeat music playing, bright visuals of people using the internet]*

*Narrator (Voiceover):*

*The internet is a world of endless possibilities—connecting us, entertaining us, and making our lives easier. But beneath its surface, there are dangers we all need to be aware of.*

*[CUT TO - Darker visuals, eerie music fades in]*

*Narrator (Voiceover):*

*Cyberbullying—one of the internet's biggest threats. It can happen to anyone, anywhere. Words can hurt, and online harassment can leave lasting scars.*

*[CUT TO - Screen showing a person typing on their computer, looking worried]*

*Narrator (Voiceover):*

*Identity theft is another major risk. Sharing too much personal information online can lead to criminals stealing your identity, leaving you to deal with the aftermath.*

*[CUT TO - Glitch effect, red warning signs flashing]*

*Narrator (Voiceover):*

*Then there's malware—viruses, ransomware, spyware. Clicking on the wrong link can infect your device and compromise your data in an instant.*

*[CUT TO - Visual of someone scrolling through social media, surrounded by distorted images]*

*Narrator (Voiceover):*

*And don't forget about fake news. Misinformation spreads like wildfire, distorting our views and affecting real-world decisions.*

*[CUT TO - Friendly visual, upbeat music returns]*

*Narrator (Voiceover):*

*But you can stay safe! Be mindful of what you share, think before you click, and always question what you see online. The internet can be a wonderful place, but only if you navigate it wisely.*

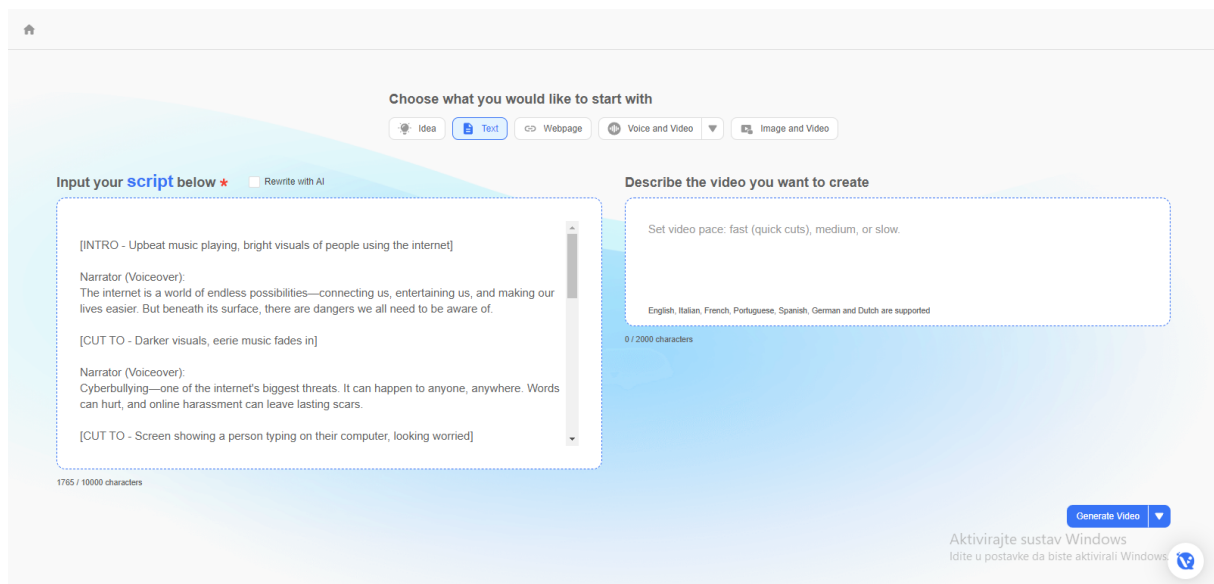
*[OUTRO - Bright visuals of people happily using the internet again]*

*Narrator (Voiceover):*

*Stay smart, stay safe, and enjoy the best of what the internet has to offer.*

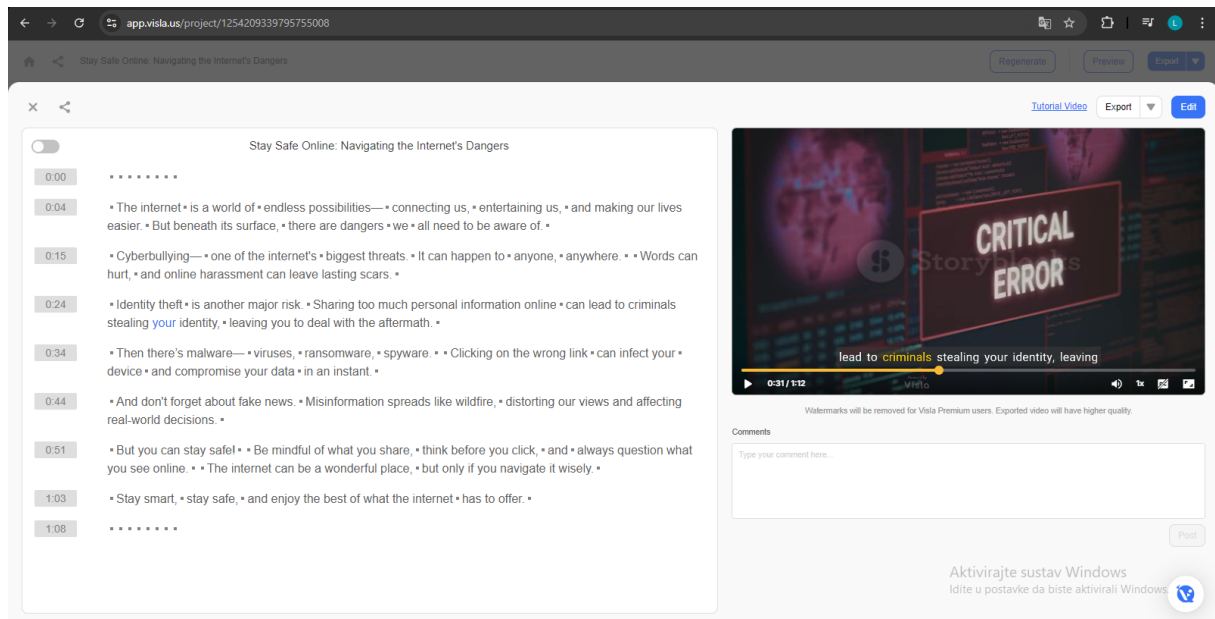
*[END SCREEN - Text: "Be safe online."]*

*[LOGO and website of the organization, soft fade out]*



Slika 33: Unos upita za izradu video sadržaja na temelju teksta (Izvor: autorski primjer)

Nakon unosa teksta, potrebno je ponovno pritisnuti gumb Generate Video za generiranje videa. Nakon što se video generira, možemo ga pregledati zajedno sa scenama od kojih se video sastoji kao što je prikazano na slici 34. Ovaj video nalazi se na sljedećoj poveznici: <https://app.visla.us/project/1254209339795755008>.

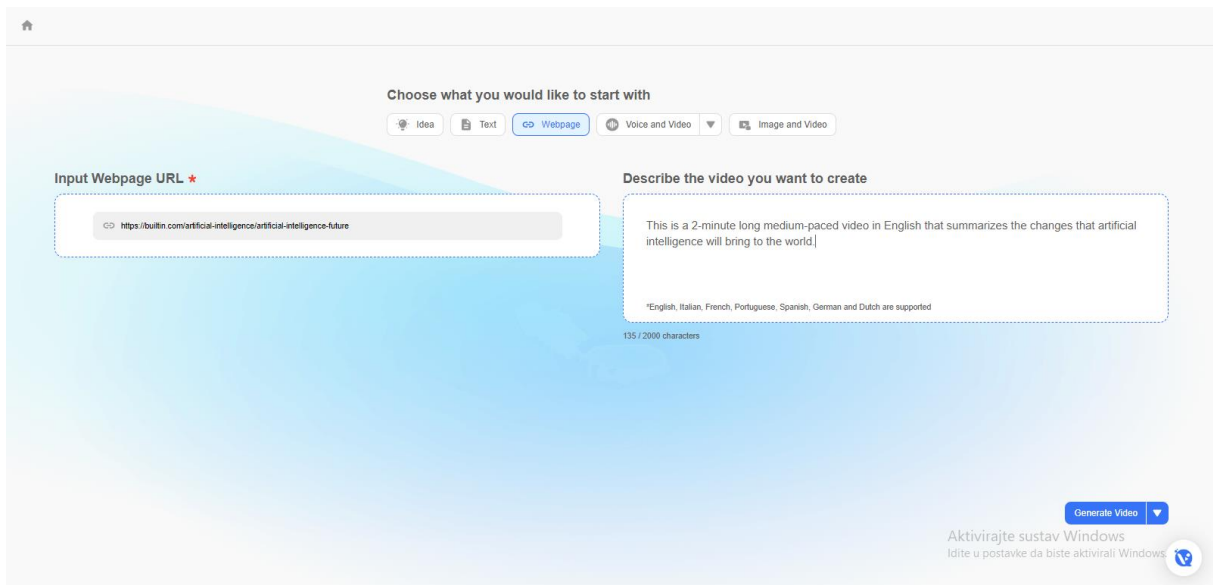


Slika 34: Prikaz generiranog videa u Visli na temelju teksta (Izvor: autorski primjer)

## 6.4. Izrada video sadržaja na temelju URL-a web stranice

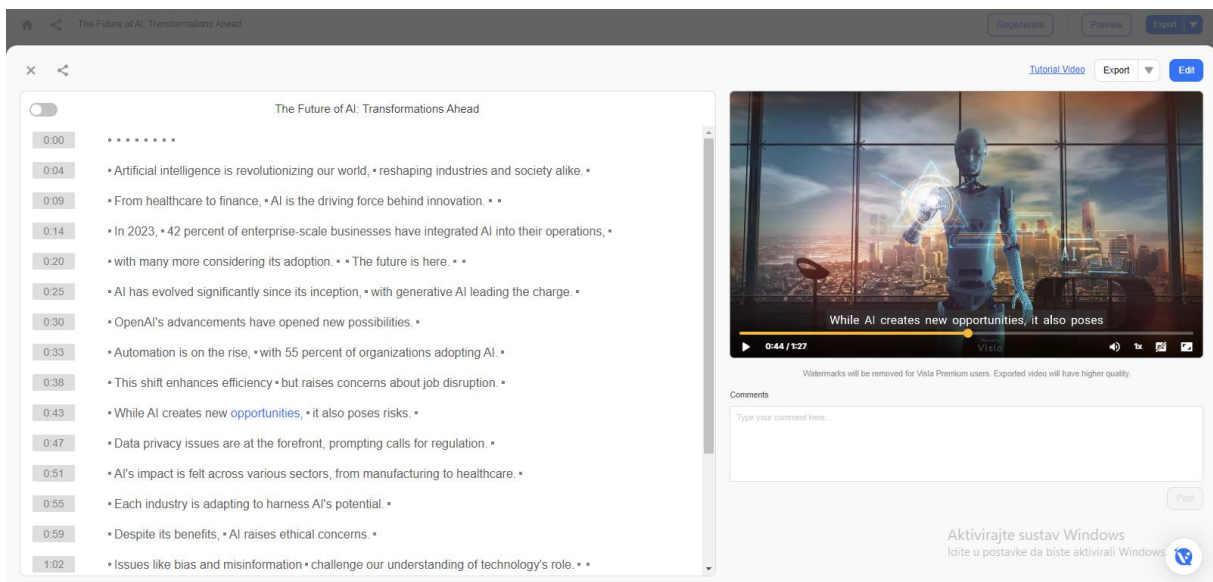
Sljedeći video koji će se kreirati bit će generiran na temelju URL-a web stranice. Ponovno je potrebno na početnom ekranu odabrati opciju *Create video* za kreiranje videa, a zatim je potrebno kod odabira vrste unosa kliknuti na opciju *Webpage* kako bi se generirao video na temelju URL-a web stranice. Thomas (2024) je autor članka koji govori o budućnosti umjetne inteligencije i taj će se članak koristiti za generiranje videa. Osim URL-a web stranice, dostupan je i tekstualni okvir za opis videa gdje korisnik može unijeti dodatne podatke o videu, kao što su duljina, svrha, tempo, publika itd.

U polje za unos linka treba unijeti poveznicu: <https://builtin.com/artificial-intelligence/artificial-intelligence-future>, a u tekstualni okvir za unos videa uneseni je sljedeći tekst: *This is a 2-minute long medium-paced video in English that summarizes the changes that artificial intelligence will bring to the world..* Ovime je zadano da se generira video u trajanju od dvije minute srednjeg tempa na engleskom jeziku koji sažeto navodi promjene koje će u svijetu donijeti umjetna inteligencija. Na slici 35 možemo vidjeti unos upita za generiranje videa na temelju URL-a web stranice.



Slika 35: Unos upita za izradu video sadržaja na temelju URL-a web stranice (Izvor: autorski primjer)

Na slici 36 možemo vidjeti prikaz generiranog videa na temelju URL-a web stranice, kao i popis scena od kojih se video sastoji. Ovaj video nalazi se na sljedećoj poveznici: <https://app.visla.us/project/1254335511930114048>.



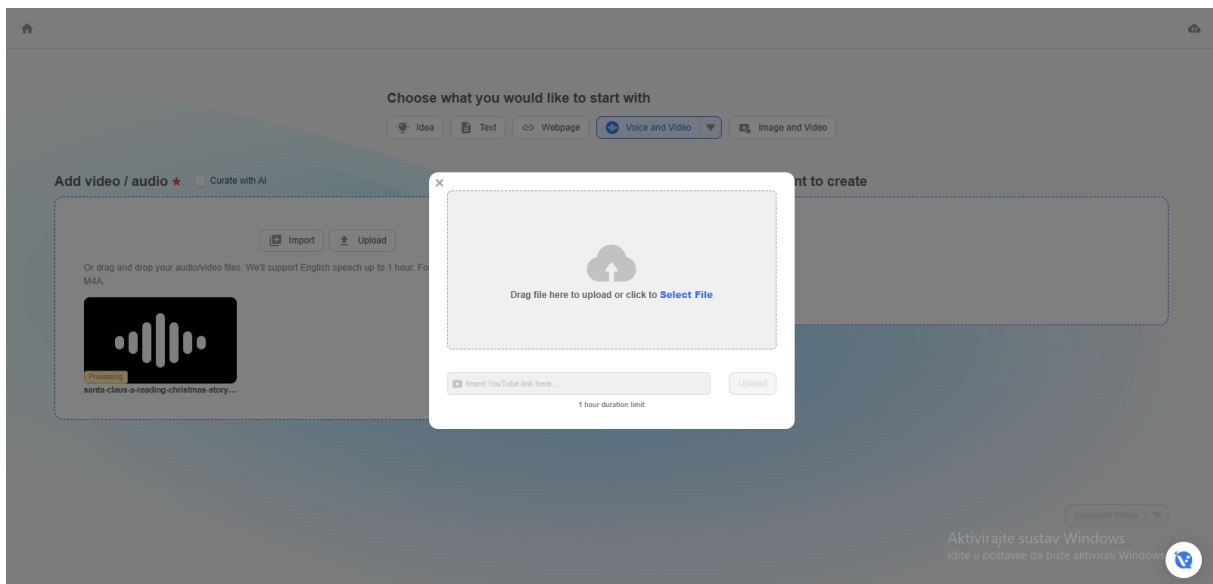
Slika 36: Prikaz generiranog videa u Visli na temelju URL-a web stranice (Izvor: autorski primjer)



## 6.5. Izrada video sadržaja na temelju audiozapisa

Visla također nudi mogućnost generiranja video sadržaja na temelju audiozapisa. Kako bismo generirali video na temelju audiozapisa, potrebno je kod odabira vrste unosa odabrati opciju *Voice and Video*. Audiozapise je moguće učitati s računala ili umetnuti poveznicu na audiozapis s YouTubea. Moguće je učitati audiozapise u trajanju do jednog sata. Ovdje također postoji dodatni tekstualni okvir za unos dodatnih informacija o videu.

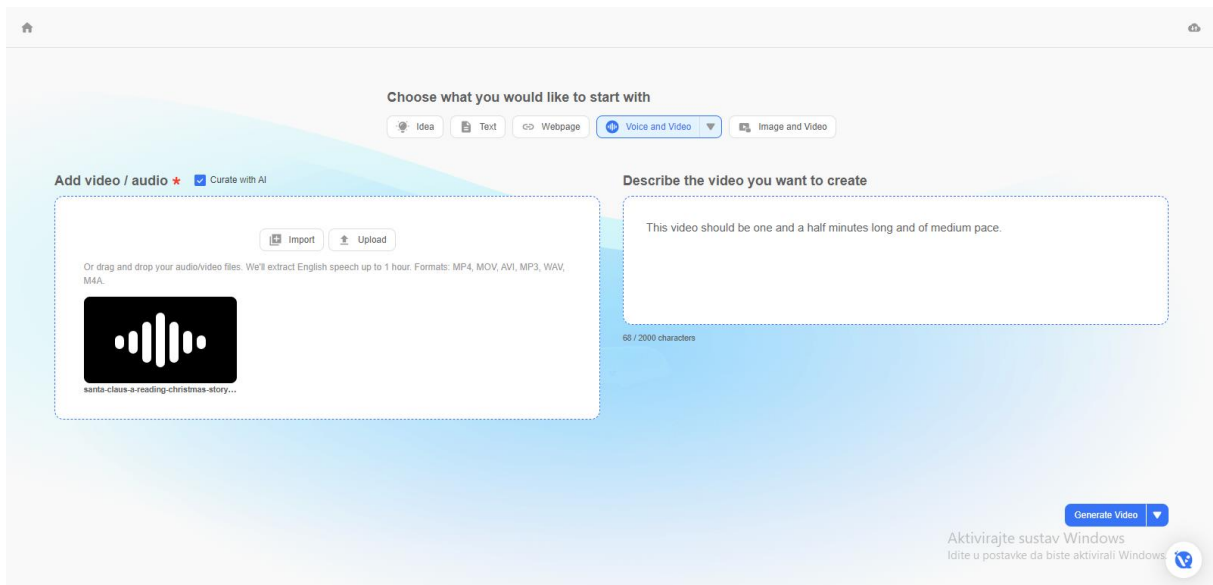
U ovom primjeru učitat ćemo audiozapis koji je prethodno preuzet s Pixabay sa sljedeće poveznice: <https://pixabay.com/sound-effects/santa-claus-a-reading-christmas-story-17777/>. Ovo je audiozapis s glasom Djeda Mraza koji čita božićnu priču. Kako bismo učitali audiozapis s računala, potrebno je kliknuti gumb *Upload* pri čemu se otvara prozor gdje možemo dovući željeni audiozapis što možemo vidjeti i na slici 37.



Slika 37: Učitavanje audiozapisa za generiranje video sadržaja (Izvor: autorski primjer)

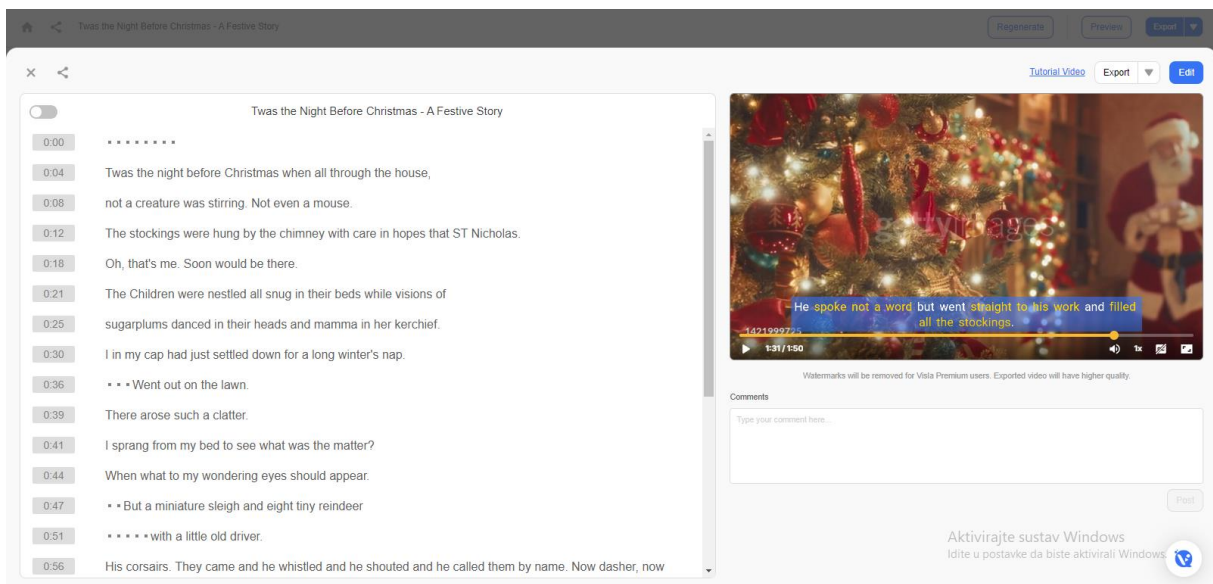
Još dodatno možemo postaviti da video traje jednu i pol minutu i da je srednjeg tempa. Na slici 38 možemo vidjeti unos upita za izradu video sadržaja na temelju audiozapisa.





Slika 38: Unos upita za izradu video sadržaja na temelju audiozapisa (Izvor: autorski primjer)

Na slici 39 možemo vidjeti video generiran na temelju audiozapisa i scene od kojih se video sastoji. Ovaj video nalazi se na sljedećoj poveznici: <https://app.visla.us/project/1271169163330600960>.

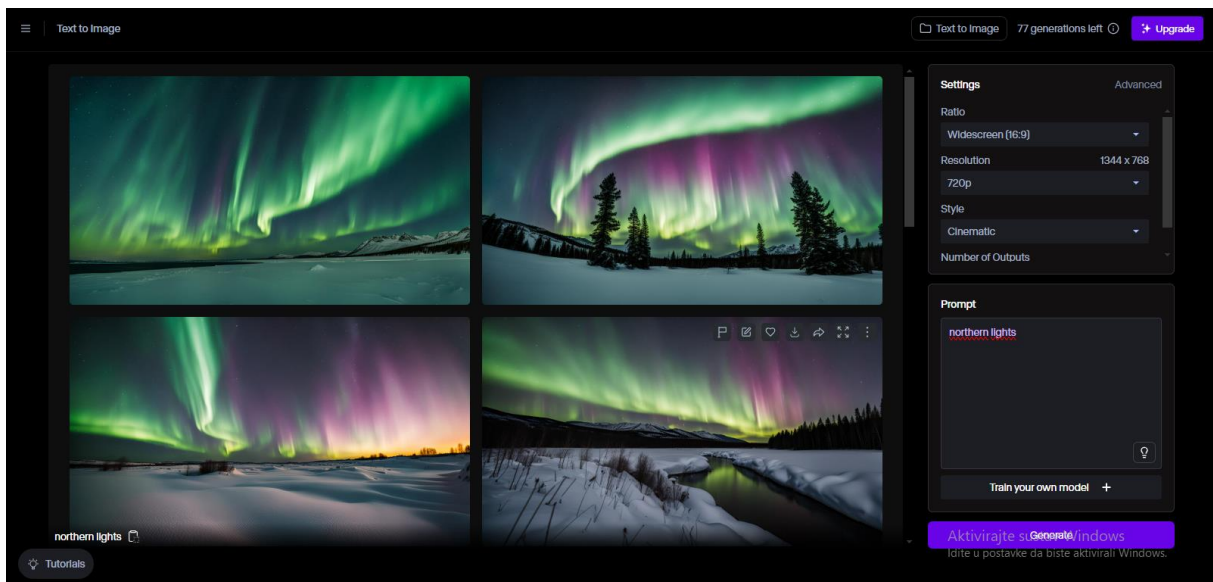


Slika 39: Prikaz generiranog videa u Visli na temelju audiozapisa (Izvor: autorski primjer)

## 6.6. Izrada video sadržaja na temelju slike

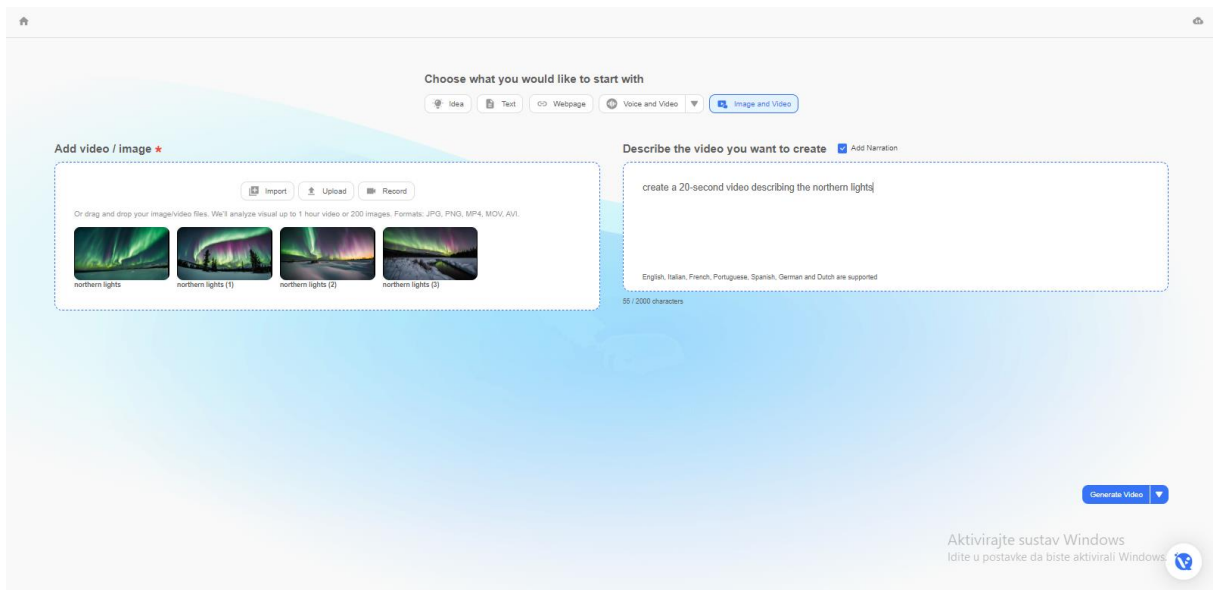
Posljednji primjer izrade video sadržaja u alatu Visla kao unos koristit će sliku. Općenito je moguće učitati najviše dvjesto slika ili čak i video u trajanju do najviše jednog sata. Kako bismo generirali video na temelju slike, potrebno je kod odabira vrste unosa odabrati opciju *Image and Video*. Video je također moguće i snimiti klikom na gumb *Record*.

Za ovaj primjer koristit ćemo slike koje su generirane pomoću umjetne inteligencije u alatu Runway. Slike su generirane na temelju tekstualnog upita, a uneseni tekst je bio *polarna svjetlost (eng. northern lights)*. Alat Runway je generirao četiri slike koje možemo vidjeti na slici 40. Ove slike bit će korištene za generiranje videa u Visli.



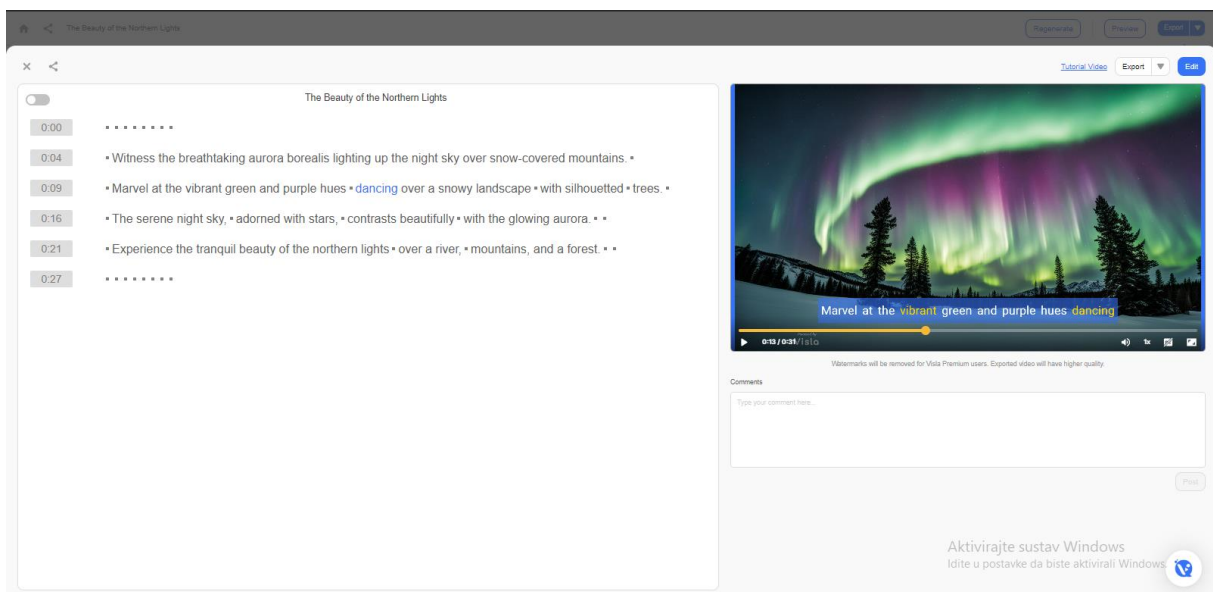
Slika 40: Slike generirane u alatu Runway (Izvor: autorski primjer)

Klikom na gumb *Upload* možemo učitati željene slike. Osim toga, ovdje također imamo dodatan tekstualni okvir za unos opisa videa pa tako možemo odrediti duljinu generiranog videa. Na slici 41 možemo vidjeti unos upita za generiranje videa na temelju slika polarne svjetlosti.



Slika 41: Unos upita za izradu video sadržaja na temelju slika (Izvor: autorski primjer)

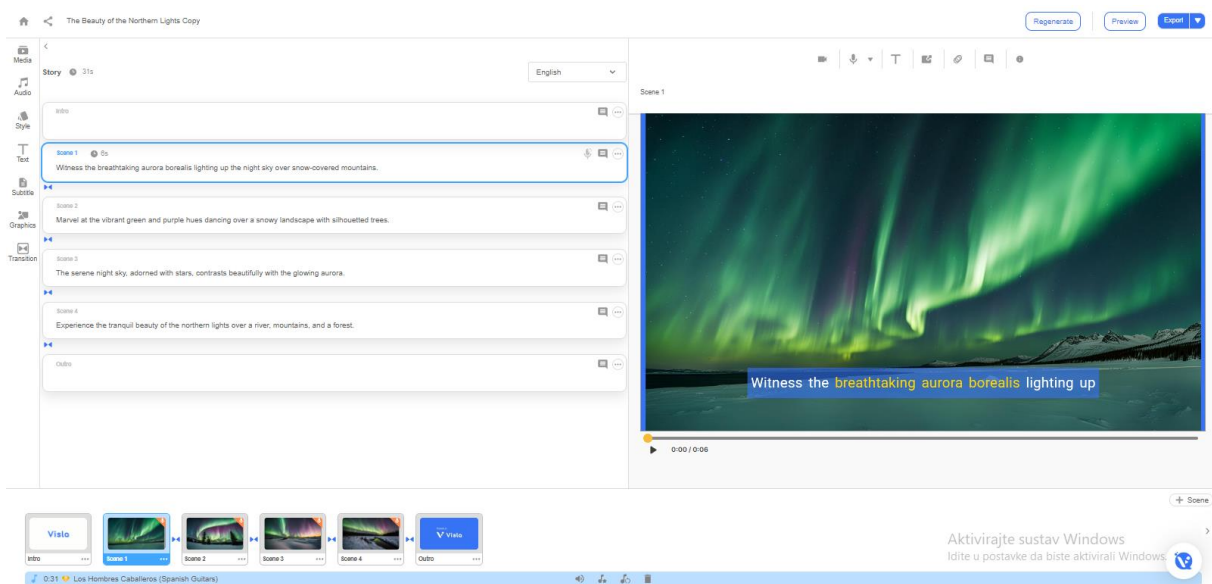
Na slici 42 prikazan je generirani video o polarnoj svjetlosti. Ovaj video nalazi se na sljedećoj poveznici: <https://app.visla.us/project/1272064553521315840>.



Slika 42: Prikaz generiranog videa u Visli na temelju slika (Izvor: autorski primjer)

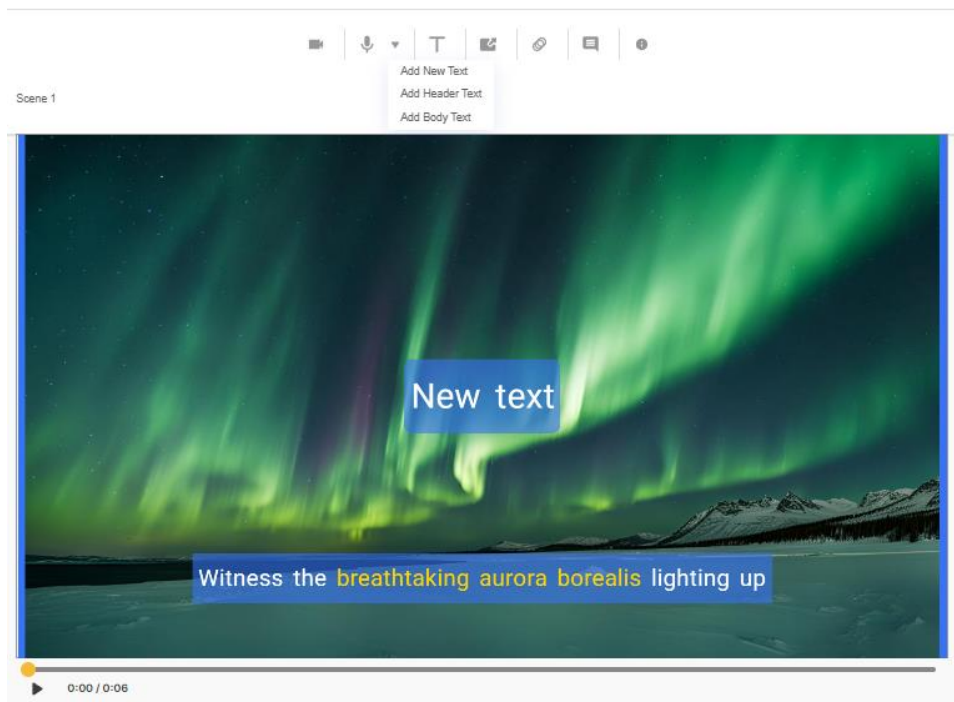
## 6.7. Uređivanje video sadržaja

Visla osim kreiranja video sadržaja, također nudi mogućnost njihovog uređivanja. Kako bismo mogli uređivati željeni videozapis, potrebno je na početnom ekranu dvaput kliknuti na video koji želimo generirati i nakon toga će nam se otvoriti sučelje za uređivanje. U ovom primjer, uređivat ćemo video koji smo generirali na temelju slika polarne svjetlosti. Na slici 43 možemo vidjeti sučelje za uređivanje videa o polarnoj svjetlosti.

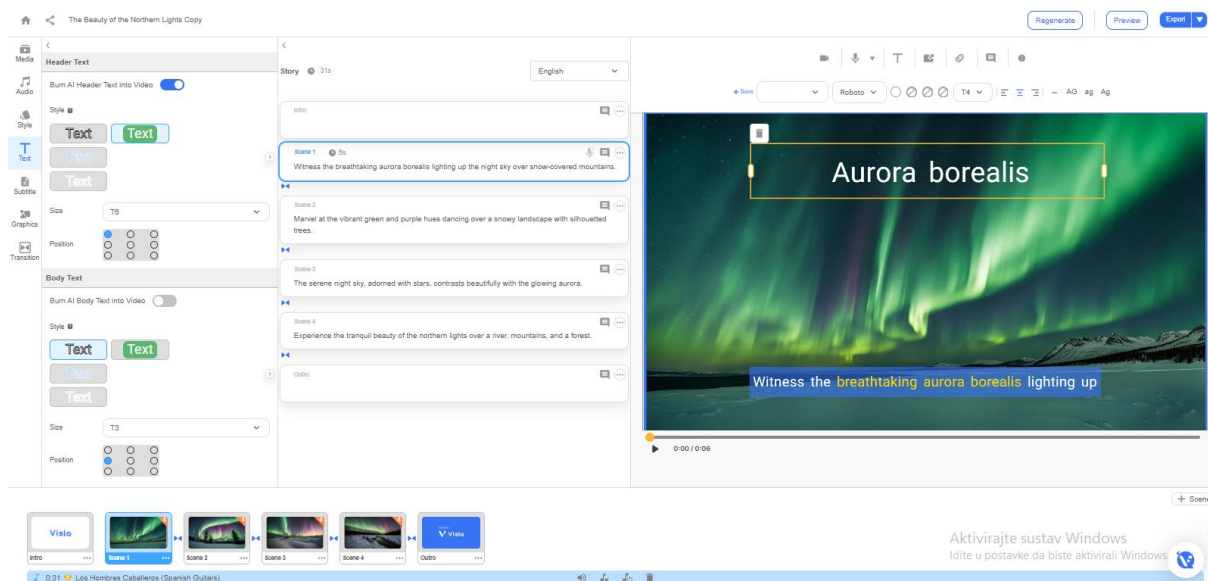


Slika 43: Sučelje za uređivanje videa o polarnoj svjetlosti (Izvor: autorski primjer)

Za početak, na početku videa možemo dodati tekst „Aurora borealis“. Kako bismo to učinili, potrebno je iznad videa odabrati ikonu teksta i kliknuti opciju *Add New Text* kao što vidimo na slici 44. Kada kliknemo na tekstualni okvir s dodanim tekstom, pojavljuju se dodatne opcije za promjenu stila teksta, fonta, veličine i boje teksta. Na slici 45 možemo vidjeti ove opcije, kao i uređeni tekst.

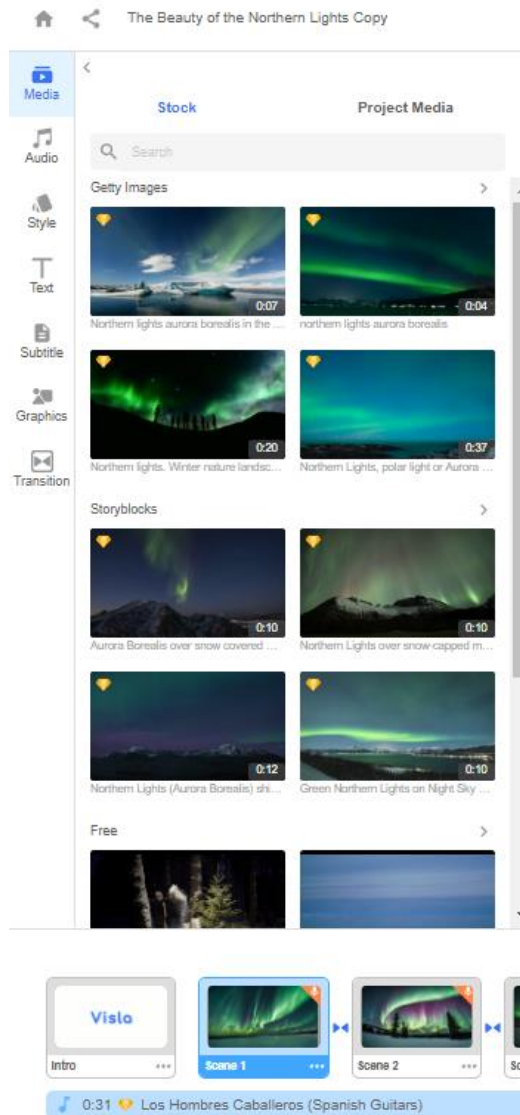


Slika 44: Dodavanje teksta (Izvor: autorski primjer)



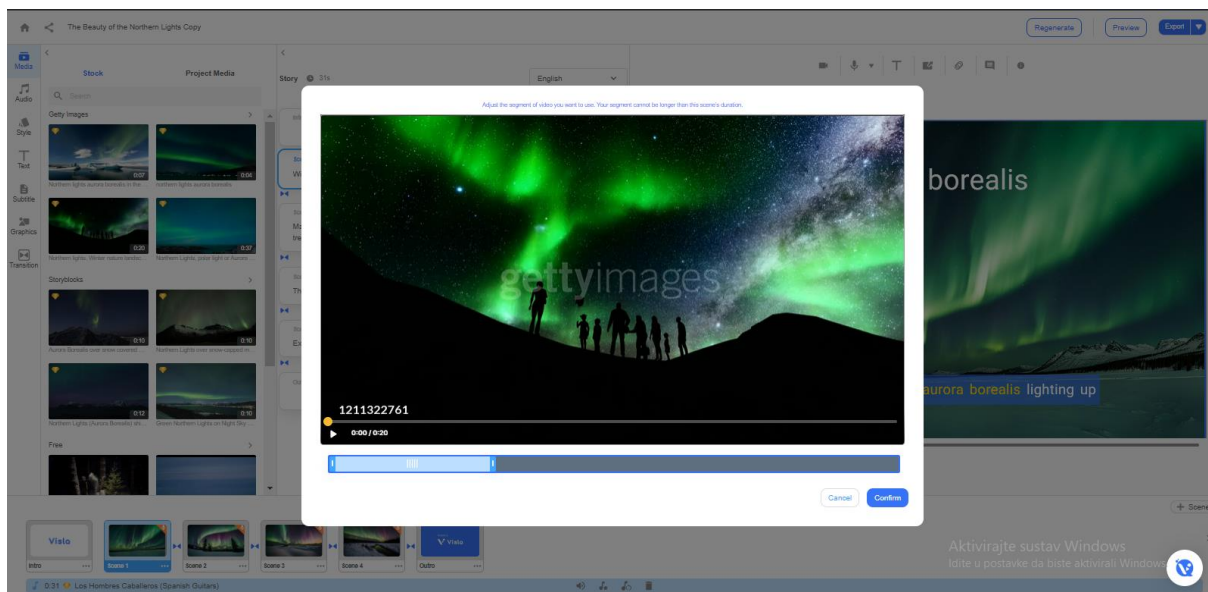
Slika 45: Oblikovanje teksta (Izvor: autorski primjer)

Visla također nudi mogućnost dodavanja stock videozapisa. Neki videozapisi u sklopu besplatnog plana će imati vodeni žig, dok su oni bez vodenog žiga potpuno besplatni. Kako bismo dodali određeni isječak videozapisa, potrebno je odabrati opciju *Media*. Visla će odmah ponuditi isječke videozapisa koji se podudaraju s temom videa koji uređujemo. Sve što je potrebno napraviti je kliknuti na željeni videozapis, odabrati segment tog videozapisa koji ne traje duže od scene na koju dodajemo taj isječak, i kliknuti na gumb *Confirm*.



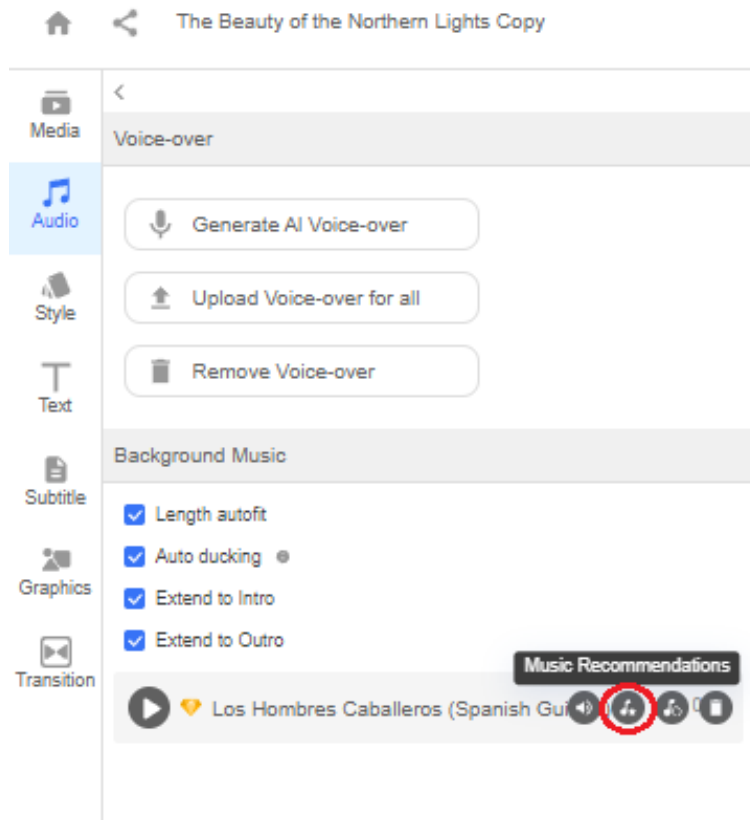
Slika 46: Dodavanje stock videozapisa putem opcije *Media* (Izvor: autorski primjer)



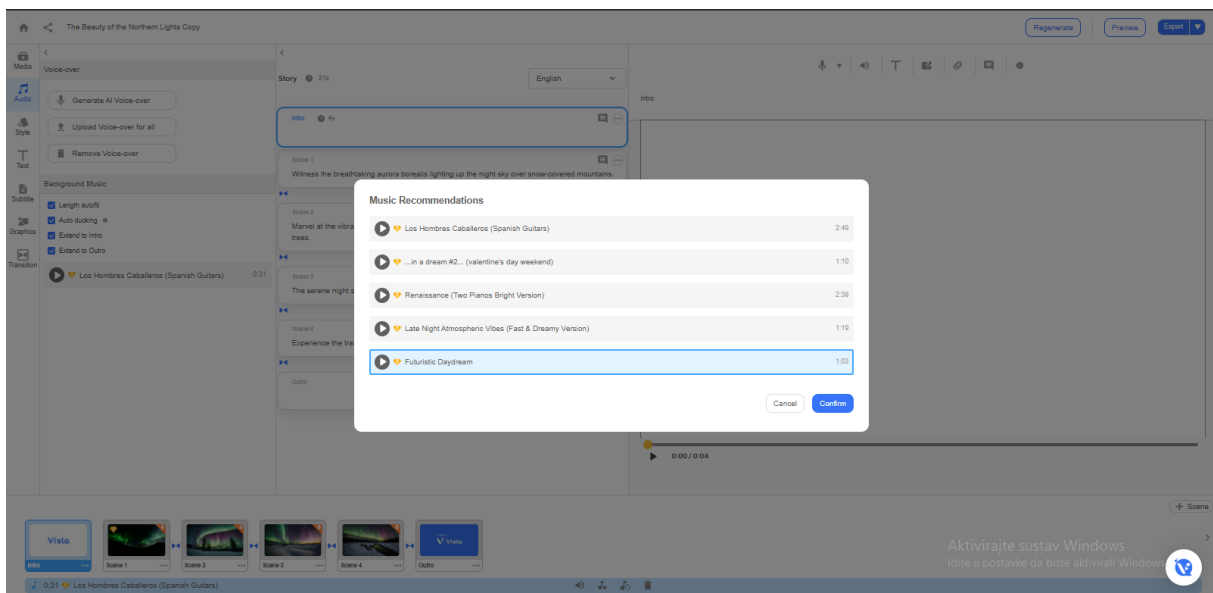


Slika 47: Odabir željenog isječka stock videozapisa (Izvor: autorski primjer)

Opcija *Audio* omogućuje korisniku generiranje i uklanjanje glasovne naracije, ali i učitavanje glasovne naracije s računala. Također je moguće promijeniti i zadanu pozadinsku glazbu. Kako bismo promijenili pozadinsku glazbu, potrebno je pod opcijom *Audio* pokraj pozadinske glazbe odabrati opciju *Music Recommendations*, odabrati željenu pozadinsku glazbu i potvrditi odabir klikom na gumb *Confirm*.



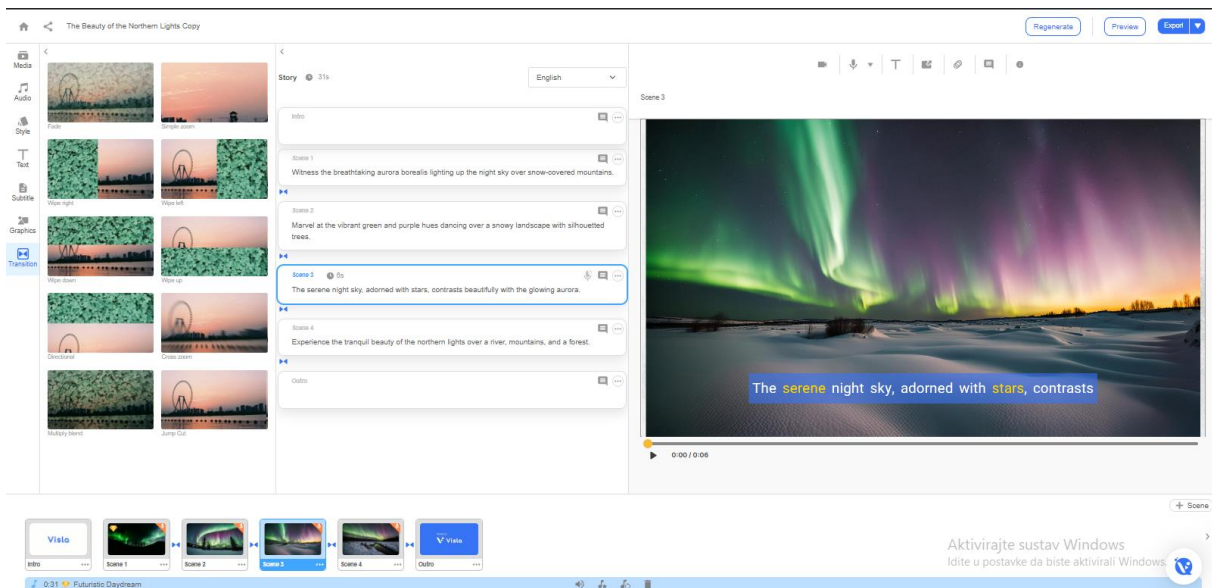
Slika 48: Promjena pozadinske glazbe (Izvor: autorski primjer)



Slika 49: Odabir pozadinske glazbe (Izvor: autorski primjer)



U Visli je moguće dodati i tranzicije između scena. Potrebno je samo na bočnoj traci odabrati opciju *Transition* i odabrati scenu na koju želimo dodati tranziciju na sljedeću scenu. Ovo su samo neki od primjera uređivanja videozapisa u Visli. Postoji još niz opcija za uređivanje videa, kao što su npr. promjena boje i veličine titlova, dodavanje grafičkih oblika i dodavanje stila. Ovaj uređeni video dostupan je na sljedećoj poveznici: <https://app.visla.us/project/1272334219690971136>.



Slika 50: Dodavanje tranzicije između scena (Izvor: autorski primjer)

## 6.8. Analiza učinkovitosti

Kako bismo mogli analizirati učinkovitost alata Visla, potrebno je uzeti u obzir nekoliko čimbenika. Analiza će se odnositi na besplatni plan jer su svi videozapisi predstavljeni u ovom radu generirani u sklopu besplatnog plana. U sklopu besplatnog plana maksimalna rezolucija generiranih videozapisa je 1080p, dok za *Business* i *Enterprise* planove maksimalna rezolucija iznosi 4k. Visla koristi stock videozapise od kojih sastavlja jedan cjeloviti video. Prilikom pregledavanja videozapisa u Visli, možemo primijetiti da je slika pomalo zamućena, što se najbolje vidi u uvodnom i završnom dijelu s logotipom Visle. Generirani videozapisi dobro su usklađeni s temom koju je na početku zadao korisnik i postoji logičan slijed između scena. Korisničko sučelje Visle je intuitivno i jednostavno za korištenje s jasnim nazivima opcija koje nedvosmisleno upućuju korisnika na određenu radnju. Ovaj alat prikladan je za početnike i mogu ga koristiti i korisnici koji nemaju prethodnog iskustva s kreiranjem i uređivanjem videa. Jedna od vrlo korisnih značajki je i mogućnost uređivanja videozapisa pa korisnik može

dodavati tekst, mijenjati pozadinsku glazbu i glasovnu naraciju, dodavati tranzicije i vlastite medijske datoteke itd. Za generiranje videozapisa u trajanju od 1-2 minute Visli je bilo potrebno manje od minute vremena pri čemu je najduže trajalo generiranje videozapisa na temelju URL-a web stranice. Visla također nudi mogućnost prilagodbe omjera stranica za različite platforme društvenih medija pa tako korisnik može optimizirati videozapis za YouTube, TikTok, Instagram, Facebook itd. Ako gledamo cijenu i broj dostupnih značajki u sklopu besplatnog plana, Visla korisnicima nudi puno veću slobodu u generiranju i uređivanju videozapisa od svojih konkurenata. Svaki korisnik ima na raspolaganju tisuću kredita koji se korisniku iznova dodjeljuju svaki mjesec i koje može koristiti za generiranje videozapisa. Broj utrošenih kredita ovisi o vrsti upita i trajanju videozapisa, ali općenito za jednominutni video kreiran na temelju ideje, Visla će korisniku uzeti oko 100 kredita. To daje korisniku mogućnost kreiranja gotovo 10 kratkih videozapisa mjesečno što je vrlo solidno ako govorimo o besplatnom planu. Ipak, ako koristimo besplatni plan, nemoguće je izbjeći vodene i zvučne žigove koji se inače maknu kod planova koji se plaćaju, ali to je razumljivo s obzirom na to da videozapise kreiramo besplatno. Na temelju svega navedenog, možemo zaključiti da je Visla učinkovit alat koji možemo koristiti za generiranje kratkih videozapisa s logičnim slijedom priče i prikladnim videoisječcima koji odgovaraju zadanoj temi. Iako ponegdje neki videoisječak nije u potpunosti u skladu s trenutnom pričom, cijeli videozapis će na kraju gledateljima biti razumljiv i uspješno će prenijeti poruku što je i svrha alata umjetne inteligencije za izradu video sadržaja.

Primjena alata umjetne inteligencije može značajno unaprijediti procese stvaranja video sadržaja. Prije svega, korištenje alata umjetne inteligencije smanjuje vrijeme potrebno za proizvodnju video sadržaja. Dovoljno je imati samo ideju ili scenarij prema kojemu će umjetna inteligencija kreirati video sadržaj. Osim toga, upotreba umjetne inteligencije smanjuje i troškove za proizvodnju videozapisa jer nije potrebno angažirati velik broj stručnjaka koji sudjeluju u stvaranju video sadržaja. Smanjenjem vremena i troškova za kreiranje video sadržaja dolazi i do povećanja produktivnosti. Umjetna inteligencija također se može koristiti i za automatizaciju jednostavnih zadataka kao što je dodavanje titlova ili stvaranje opisa za pojedini videozapis. Stručnjaci se tako mogu više usredotočiti na kreativni dio procesa proizvodnje video sadržaja. S obzirom na to da su alati umjetne inteligencije za izradu video sadržaja dostupni sve većem broju ljudi, to pomaže kreativnim pojedincima koji inače ne bi imali dovoljno resursa i financijskih sredstava za stvaranje video sadržaja. Alati umjetne inteligencije također se mogu koristiti za poboljšanje kvalitete videozapisa jer imaju mogućnost poboljšati boje, dodavati vizualne elemente, prilagoditi kontrast ili napraviti druga poboljšanja. Osim toga, umjetna inteligencija može se koristiti i za stvaranje personaliziranog sadržaja stvaranjem videozapisa koji će odgovarati ciljanoj publici.

## 7. Zaključak

Kroz ovaj završni rad predstavljeni su alati umjetne inteligencije za izradu video sadržaja, navedeni su primjeri njihove primjene i na konkretnim primjerima demonstrirana je izrada video sadržaja u odabranom alatu. Svaki od dostupnih alata nudi različite načine generiranja video sadržaja, kao i različite cjenovne planove. Neki alati kao što su InVideo AI i Sora nude mogućnost generiranja video sadržaja samo na temelju tekstualnog upita, dok neki alati kao što je Runway nude mogućnost generiranja videa sadržaja i na temelju slike.

S obzirom na to da je Visla alat koji nudi velik broj dostupnih značajki u sklopu besplatnog plana, taj je alat odabran za demonstraciju procesa izrade video sadržaja. Video sadržaji su u Visli generirani na nekoliko načina: na temelju ideje, na temelju teksta, na temelju audiozapisa, na temelju URL-a web stranice i na temelju slika. S obzirom na vrijeme potrebno za generiranje i kvalitetu video sadržaja, može se reći da je Visla vrlo učinkovit alat za stvaranje kratkih videozapisa za različite namjene.

Generiranje video sadržaja pomoću umjetne inteligencije ima mnoge prednosti, ali i nedostatke. Umjetna inteligencija smanjuje vrijeme i troškove potrebne za proizvodnju video sadržaja, poboljšava kvalitetu videozapisa i automatizira zadatke na koje bi ljudi inače utrošili puno vremena. S druge strane, neki pojedinci bi mogli zloupotrijebiti ovakav jednostavan način izrade video sadržaja stvaranjem lažnih videozapisa koji predstavljaju poznate osobe kako govore nešto što u stvarnosti nikad nisu rekli. Osim širenja *deepfake* tehnologije, postoji opasnost od gubitka posla za ljude koji rade kao video montažeri jer bi njihov posao mogla preuzeti umjetna inteligencija. Također se javlja i problem s autorskim pravima jer umjetna inteligencija koristi skup postojećih podataka za generiranje sadržaja, a ti su podaci u nečijem vlasništvu.

Alati umjetne inteligencije za izradu video sadržaja još uvijek se razvijaju. Očekuje se da će do kraja 2024. godine Sora biti dostupna široj javnosti, a već sada je stekla veliku popularnost. S razvojem generativne umjetne inteligencije, ovi alati će početi stvarati sve realističnije i uvjerljivije videozapise i s vremenom će biti sve teže uočiti da se radi o videozapisu koji je generirala umjetna inteligencija.

## Popis literature

Shi, Z. (2020). *Advanced Artificial Intelligence*. Singapur: World Scientific Publishing Co. Pte. Ltd.

Patel, D. M. (2023). *Artificial intelligence & Generative AI for Beginners: The Complete Guide*. Wrocław, Poljska: Amazon Fulfillment Poland Sp. z o.o.

Russell, S. J., i Norvig, P. (2020). *Artificial intelligence: A Modern Approach*. Upper Saddle River, New Jersey, USA: Pearson Education, Inc.

IBM (bez dat.) *What is machine learning (ML)?* Preuzeto 19.6.2024. s <https://www.ibm.com/topics/machine-learning>

Kanade, V. (2022). *What Is Machine Learning? Definition, Types, Applications, and Trends*. Preuzeto 19.6.2024. s <https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/>

de Albuquerque, V. H. C., Raj, P., Yadav, S. P. (2024). *Toward Artificial General Intelligence*. Berlin/Boston: Walter de Gruyter GmbH

Gillis, A. S. (2023). *deep learning*. Preuzeto 20.6.2024. s <https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network>

Macukow, B. (2016). *Neural Networks – State of Art, Brief History, Basic Models and Architecture*. Švicarska: Springer International Publishing

SAS (bez dat.) *Artificial Neural Networks What they are & why they matter*. Preuzeto 20.6.2024. s [https://www.sas.com/el\\_gr/insights/analytics/neural-networks.html#technical](https://www.sas.com/el_gr/insights/analytics/neural-networks.html#technical)

Crabtree, M. (2024). *What is Natural Language Processing (NLP)? A Comprehensive Guide for Beginners*. Preuzeto 20.6.2024. s <https://www.datacamp.com/blog/what-is-natural-language-processing>

DataJobs (bez dat.) *Natural Language Processing*. Preuzeto 20.6.2024. s [https://datajobs.com/data-science-repo/NLP-Background-\[SU\].pdf](https://datajobs.com/data-science-repo/NLP-Background-[SU].pdf)

Simplilearn (2024). *The Power of Computer Vision in AI: Unlocking the Future!* Preuzeto 20.6.2024. s <https://www.simplilearn.com/computer-vision-article>

Mihajlović, I. (2019). *Everything You Ever Wanted To Know About Computer Vision*. Preuzeto 20.6.2024. s <https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e>

Porter, A. (2023). *Unveiling 6 Types of Generative AI*. Preuzeto 20.6.2024. s <https://bigid.com/blog/unveiling-6-types-of-generative-ai/>

Feuerriegel, S., Hartmann, J., Janiesch, C., Zschech, P. (2024). *Generative AI. Business & Information Systems Engineering*, 66(1), 111-126. <https://doi.org/10.1007/s12599-023-00834-7>

- Nvidia (bez dat.). *What is Generative AI?* Preuzeto 20.6.2024. s <https://www.nvidia.com/en-us/glossary/generative-ai/>
- AI Video Generation (2024). Preuzeto 21.6.2024. s <https://deepgram.com/ai-glossary/ai-video-generation>
- VideoZen (2023). *The Evolution of AI in Video Generation: A Historical Perspective*. Preuzeto 26.7.2024. s <https://medium.com/@videozen/the-evolution-of-ai-in-video-generation-a-historical-perspective-8bc46d8bf242>
- Michelson, L. (2023). *How AI Video Generators Transform Concepts into Visual Masterpieces*. Preuzeto 21.6.2024. s <https://www.d-id.com/blog/ai-video-generators-transform-concepts-into-visual-masterpieces/>
- AIContentfy team (2024). *AI-generated content for video marketing and multimedia content*. Preuzeto 21.6.2024. s <https://aicontentfy.com/en/blog/ai-generated-content-for-video-marketing-and-multimedia-content>
- Talvola, E. (2023). *The Ethics of AI Video Generators: Pros and cons*. Preuzeto 25.7.2024. s <https://animoto.com/blog/video-tips/ai-video-generator-ethics>
- Elena (2023). *Write the Best Prompts to Generate AI Videos You Need*. Preuzeto 6.8.2024. s <https://www.flexclip.com/learn/text-to-video-prompt.html#part2>
- Saleem, A. (2024). *7 tested prompting techniques to use AI video generators – Evolution of Art industry*. Preuzeto 6.8.2024. s <https://datasciencedojo.com/blog/prompting-to-use-ai-video-generators/>
- Sharma, R. i Achar, C. (2024). *Generative AI for Learning: Investigating the Potential of Synthetic Learning Videos*. *International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)*, 13(6), doi: 10.15680/IJIRSET.2024.1306143
- Revolutionising Education: AI Avatar Videos for Engaging Learning* (2023). Preuzeto 21.6.2024. s <https://www.emotech.ai/blog/ai-avatar-videos-education>
- Odabashian, R. (2024). *SORA and Healthcare: How Text-to-Video AI Can Revolutionize Patient Care*. Preuzeto 21.6.2024. s [https://www.linkedin.com/pulse/sora-healthcare-how-text-to-video-ai-can-revolutionize-zx66e?trk=public\\_post\\_main-feed-card\\_reshare\\_feed-article-content](https://www.linkedin.com/pulse/sora-healthcare-how-text-to-video-ai-can-revolutionize-zx66e?trk=public_post_main-feed-card_reshare_feed-article-content)
- "Revolutionizing Learning & Training in Hospitals: The Pros and Cons of Using Deepfake Generators for AI-Powered Videos"* (bez dat.) Preuzeto 21.6.2024. s <https://www.trainday.io/industry/hospitalsandhealthcare/revolutionizing-learning-and-training-in-hospitals-the-pros-and-cons-of-using-deepfake-generators-for-ai-powered-videos+6519559b8fb1233bde0cdf5d/>
- Andriasyan, S. (bez dat.). *How Artificial Intelligence Is Used in the Film Industry*. Preuzeto 21.6.2024. s <https://smartclick.ai/articles/how-artificial-intelligence-is-used-in-the-film-industry/>
- Xagoraris, K. (2020). *Using Artificial Intelligence in the VFX and Film Industry*. Preuzeto 21.6.2024. s <https://www.katexagoraris.com/artificial-intelligence-and-film-ind>
- Santiago, E. (2024). *14 Best AI Video Generators to Use in 2024*. Preuzeto 22.6.2024. s <https://blog.hubspot.com/marketing/ai-video-generator>

McCallister, N. (2023). *InVideo Exposed | We Tested the New Web Based Video Creation Software*. Preuzeto 22.6.2024. s <https://entresource.com/invideo-review/>

InVideo (bez dat.). *Pricing*. Preuzeto 6.8.2024. s <https://invideo.io/studio-pricing>

Harry (2024). *InVideo AI Review: Pros, Cons, and User Experiences*. Preuzeto 22.6.2024. s <https://aiprotips.net/invideo-ai-review-pros-cons-and-user-experiences/>

Jalli, A. (2023). *11 Best AI Video Generators (June 2024)*. Preuzeto 22.6.2024. s <https://medium.com/@artturi-jalli/ai-video-generators-14312e5bcb62>

*Lumen5 review: Is this Video Maker Right for you? Features, Pro's & Con's* (bez dat.). Preuzeto 22.6.2024. s <https://yourentrepreneurresources.com/lumen5-review/>

Lumen5 (bez dat.). *Plans for your video content creation strategy*. Preuzeto 6.8.2024. s <https://lumen5.com/pricing/>

„Lumen5“ (bez dat.). Preuzeto 22.6.2024. s <https://findmyaitool.com/tool/lumen5>

Rivers, A. (2024). *Visla: Turn text into Videos Made Simple with AI [2024]*. Preuzeto 22.6.2024. s <https://softgist.com/visla-review#pros--cons-of-visla-ai>

Rebelo, M. (2023). *The best AI video generators*. Preuzeto 22.6.2024. s <https://zapier.com/blog/best-ai-video-generator/#visla>

Visla (bez dat.). *Choose the plan that fits your needs*. Preuzeto 6.8.2024. s <https://www.visla.us/pricing>

Vikram (2024). *Runway - Features, Pricing, Pros & Cons (June 2024)*. Preuzeto 22.6.2024. s <https://siteefy.com/ai-tools/runway/>

Wiggers, K. (2024). *Runway's new video-generating AI, Gen-3, offers improved controls*. Preuzeto 22.6.2024. s <https://techcrunch.com/2024/06/17/runways-new-video-generating-ai-gen-3-offers-improved-controls/>

Runway (bez dat.). *Choose the best plan for you*. Preuzeto 6.8.2024. s <https://runwayml.com/pricing>

Shiksha (2024). *Runway ML: Where Art Meets AI for Limitless Possibilities*. Preuzeto 22.6.2024. s <https://www.shiksha.com/online-courses/articles/runwayml-where-art-meets-ai-for-limitless-possibilities-blogId-153191#7>

Bharat (2024). *SORA: The New Text to Video Generation Model from OpenAI*. Preuzeto 5.8.2024. s <https://opencv.org/blog/sora-openai/>

Cotton, R. (2024). *What is Open AI's Sora? How it Works, Use Cases, Alternatives & More*. Preuzeto 5.8.2024. s <https://www.datacamp.com/blog/openai-announces-sora-text-to-video-generative-ai-is-about-to-go-mainstream>

Jasaitis, A. (2024). *OpenAI Sora Review - Features, Pros and Cons of Sora AI Text to Video Generator*. Preuzeto 5.8.2024. s <https://www.wps.com/blog/openai-sora-review-features-pros-and-cons-of-sora-ai-text-to-video-generator/>  
<https://chatgpt.com/share/277f9994-3819-4264-b261-38dab4a01313>

Thomas, M. (2024). *The Future of AI: How Artificial Intelligence Will Change the World*.  
Preuzeto 10.8.2024. s <https://bultin.com/artificial-intelligence/artificial-intelligence-future>

# Popis slika

Slika 1: Dijagram podataka piksela (Izvor: Mihajlović, 2019).....	7
Slika 2: Generativna suparnička mreža (Izvor: Feuerriegel, Hartmann, Janiesch i Zschech, 2024).....	9
Slika 3: Dječak s pjegama iz videa Toys „R“ Us (Izvor: YouTube).....	17
Slika 4: Dječak bez pjega iz videa Toys „R“ Us (Izvor: YouTube).....	17
Slika 5: Nepravilne naočale iz videa Toys „R“ Us (Izvor: YouTube).....	18
Slika 6: Primjer dobro generiranog video avatara u obrazovnom videu (Izvor: YouTube).....	19
Slika 7: Primjer loše generiranog video avatara u obrazovnom videu (Izvor: YouTube).....	20
Slika 8: Video avatar s tekstualnim opisom teme videa (Izvor: YouTube).....	22
Slika 9: Video avatar objašnjava pacijentu jedan od načina uzimanja lijekova (Izvor: YouTube).....	22
Slika 10: Prikaz zalaska sunca iz trailera za film The Outworld (Izvor: YouTube).....	24
Slika 11: Prikaz automobila na plaži iz trailera za film The Outworld (Izvor: YouTube).....	24
Slika 12: Izgled početnog ekrana alata InVideo AI (Izvor: autorski primjer).....	26
Slika 13: Sučelje za unos teksta u alatu InVideo AI (Izvor: autorski primjer).....	27
Slika 14: Sučelje za uređivanje generiranog videa u alatu InVideo AI (Izvor: autorski primjer)	28
Slika 15: Izgled početnog ekrana alata Lumen5 (Izvor: autorski primjer).....	29
Slika 16: Sučelje alata Lumen5 za unos informacija o videu (Izvor: autorski primjer).....	30
Slika 17: Sučelje alata Lumen5 za uređivanje generiranog videa (Izvor: autorski primjer)....	30
Slika 18: Izgled početnog ekrana alata Visla (Izvor: autorski primjer).....	32
Slika 19: Sučelje alata Visla za unos informacija o videu (Izvor: autorski primjer).....	33
Slika 20: Sučelje alata Visla za uređivanje generiranog videa (Izvor: autorski primjer).....	33
Slika 21: Izgled početnog ekrana alata Runway (Izvor: autorski primjer).....	35
Slika 22: Sučelje alata Runway za generiranje videa iz teksta ili slike (Izvor: autorski primjer)	36
Slika 23: Žena hoda ulicom Tokija u videu generiranom pomoću Sora.....	37
Slika 24: Četiri glavne opcije na početnom ekranu Visle (Izvor: autorski primjer).....	39
Slika 25: Bočna traka početnog ekrana Visle (Izvor: autorski primjer).....	40
Slika 26: Odabir i unos željenog upita za generiranje video sadržaja u Visli (Izvor: autorski primjer).....	41
Slika 27: Podjela videa na scene (Izvor: autorski primjer).....	41
Slika 28: Tekstualni okviri za svaku scenu (Izvor: autorski primjer).....	42
Slika 29: Unos upita za izradu video sadržaja na temelju ideje (Izvor: autorski primjer).....	43
Slika 30: Prikaz generiranog videa u Visli na temelju ideje (Izvor: autorski primjer).....	43
Slika 31: Opcija <i>Share</i> za dijeljenje videa (Izvor: autorski primjer).....	44



Slika 32: Kopiranje poveznice na video (Izvor: autorski primjer).....	44
Slika 33: Unos upita za izradu video sadržaja na temelju teksta (Izvor: autorski primjer).....	46
Slika 34: Prikaz generiranog videa u Visli na temelju teksta (Izvor: autorski primjer).....	47
Slika 35: Unos upita za izradu video sadržaja na temelju URL-a web stranice (Izvor: autorski primjer).....	48
Slika 36: Prikaz generiranog videa u Visli na temelju URL-a web stranice (Izvor: autorski primjer).....	48
Slika 37: Učitavanje audiozapisa za generiranje video sadržaja (Izvor: autorski primjer).....	49
Slika 38: Unos upita za izradu video sadržaja na temelju audiozapisa (Izvor: autorski primjer)	50
Slika 39: Prikaz generiranog videa u Visli na temelju audiozapisa (Izvor: autorski primjer) .	50
Slika 40: Slike generirane u alatu Runway (Izvor: autorski primjer) .....	51
Slika 41: Unos upita za izradu video sadržaja na temelju slika (Izvor: autorski primjer) .....	52
Slika 42: Prikaz generiranog videa u Visli na temelju slika (Izvor: autorski primjer).....	52
Slika 43: Sučelje za uređivanje videa o polarnoj svjetlosti (Izvor: autorski primjer).....	53
Slika 44: Dodavanje teksta (Izvor: autorski primjer) .....	54
Slika 45: Oblikovanje teksta (Izvor: autorski primjer).....	54
Slika 46: Dodavanje stock videozapisa putem opcije <i>Media</i> (Izvor: autorski primjer).....	55
Slika 47: Odabir željenog isječka stock videozapisa (Izvor: autorski primjer).....	56
Slika 48: Promjena pozadinske glazbe (Izvor: autorski primjer).....	57
Slika 49: Odabir pozadinske glazbe (Izvor: autorski primjer) .....	57
Slika 50: Dodavanje tranzicije između scena (Izvor: autorski primjer).....	58

# Popis tablica

Tablica 1: Definicije umjetne inteligencije u dvije dimenzije .....	2
Tablica 2: Alati umjetne inteligencije za izradu video sadržaja.....	25
Tablica 3: Cjenovni planovi za InVideo AI .....	26
Tablica 4: Cjenovni planovi za Lumen5 .....	29
Tablica 5: Cjenovni planovi za Vislu .....	31
Tablica 6: Cjenovni planovi za Runway .....	35

# Prilozi

Poveznice na generirane video sadržaje:

Video generiran na temelju ideje: <https://app.visla.us/project/1254196944843100160>

Video generiran na temelju teksta: <https://app.visla.us/project/1254209339795755008>

Video generiran na temelju URL-a web stranice:  
<https://app.visla.us/project/1254335511930114048>

Video generiran na temelju audiozapisa: <https://app.visla.us/project/1271169163330600960>

Video generiran na temelju slika: <https://app.visla.us/project/1272064553521315840>

Uređeni video generiran na temelju slika: <https://app.visla.us/project/1272334219690971136>