

Videoiden ASR-tekstitys ja tekstityksen työnkulku

DigiCampus/Esteettömyys ja saavutettavuus -osahanke

Mikael Viitasaari/JAMK, Mikko Uimonen/HAMK



Sisältö

- Automaattiseen puheentunnistukseen (ASR) perustuva videotekstitys
 - Toimintaperiaate
 - Tutkimusaineisto ja sen käsittely
 - Havainnot ja johtopäätökset
- Tekstityksen työnkulku (Mikko Uimonen)

ASR –tekstityksen perusperiaate

1. Foneettinen taso: Äänteiden ja sanojen tunnistaminen
2. Semanttinen taso: Merkitysten tunnistaminen
 1. Taustalla kielimalli
 2. Sanatason tunnistuksen korjailu oletetun/tunnistetun sisällön merkityksen pohjalta
 3. Lauserakenteen määrittäminen (osassa järjestelmiä)

Tavoite

- Arviointityön tavoitteet
 - Selvittää aineistolähtöisesti millaisia tyypillisiä virheitä ASR –tekstityksissä ilmenee
 - Miten erilaiset videotallenteen laatuominaisuudet vaikuttavat ASR –tekstityksen laatuun
 - Arvioida, kuinka käyttökelpoista tekstitystä automaattisella puheentunnistuksella syntyy

Miten arvioitiin

- Keräsimme kolmesta korkeakoulusta yhteensä n. 25 h:n videoaineiston
 - Erilaisia puhujia, erilaisia tallennustilanteita, eri sisältöteemoja/oppiaineita, erilaisia laitteita käytössä jne.
- Aineisto ladattiin kolmeen videojärjestelmään + kaikkiin videoihin puheentunnistukseen pohjautuva tekstitys käyttöön

Arvioitujen järjestelmien tuottamien ASR-tekstitysten ominaispiirteitä ja –ominaisvirheitä

- Kaksi arvioinnissa ollutta ASR-toiminnallisuutta pyrki tuottamaan lauserakenteen ml. isot alkukirjaimet ja pisteet
- Yksi järjestelmä tuotti kauttaaltaan pienillä kirjaimilla tuotetun, ei lauserakennetta tavoittelevan tekstityksen.



Arvioinnin haaste

- Tavoitteena oli verrata ASR –tekstityksiä manuaalisesti tuotettuun sanatarkasti litteroituun tekstiin
 - Puhekielisyys, murteen käyttö
 - Toistot
- HAASTE:Arvioinnissa mukana olevien järjestelmien automaattitekstitys eteni toisistaan poikkeavalla rytmityksellä
- RATKAISU:Tekstitykset jäsenneltiin uudelleen yhtäläisellä rytmityksellä



Lähtötilanne: Eri tavoin rytmittyvät tekstitykset

5
00:00:25,000 --> 00:00:30,000
lähinnä vesihuoltoa. Tämä 15 opintopisteen paketti

6
00:00:30,000 --> 00:00:36,000
alkaa kyllä sillä että tutustutaan vesihuoltolakiin ja tutustutaan

7
00:00:36,000 --> 00:00:42,000
vesivaroihin noin kaiken kaikkiaan vesistöjen moninaiskäyttöasioihin

5
00:00:18.500 --> 00:00:23.840
viidentoista opintopisteen moduuli vesitekniikka tässä tapauksessa

6
00:00:23.840 --> 00:00:28.250
tarkoittaa lähinnä vesihuoltoa tämä

7
00:00:28.250 --> 00:00:32.900
viidentoista opintopisteen paketti alkaa kyllä sillä että tutustutaan

6
0:00:18,200 --> 0:00:21,240
15 opintopisteen.

7
0:00:22,280 --> 0:00:27,340
Tekniikka tässä tapauksessa tarkoittaa lähinnä vesihuoltoa ...

8
0:00:27,340 --> 0:00:31,636
tämä 15 opintopisteen paketti alkaa kyllä sillä että ...



”Synkronoidut” tekstisisällöt

Vesitekniikka. Ja tässä tapauksessa tarkoittaa lähinnä vesihuoltoa. Tämä 15 opintopisteen paketti alkaa kyllä sillä, että tutustutaan vesihuoltolakiin, tutustutaan vesivaroihin noin kaiken kaikkiaan ja vesistöjen moninaiskäyttöasioihin ja tuota...

Tekniikka tässä tapauksessa tarkoittaa lähinnä vesihuoltoa tämä 15 opintopisteen paketti alkaa kyllä sillä että tutustutaan vesihuoltolaki. Ostetaan vesivaroihin on kaiken kaikkiaan ja vesistöjen MoniNaiset käyttö asioihin ja. Tuota

tekniikka tässä tapauksessa tarkoittaa lähinnä vesihuoltoa. Tämä 15 opintopisteen paketti alkaa kyllä sillä että tutustutaan vesihuoltolakiin ja tutustutaan vesivaroihin noin kaiken kaikkiaan vesistöjen moninaiskäyttöasioihin ja

vesitekniikka tässä tapauksessa tarkoittaa lähinnä vesihuoltoa tämä viidentoista opintopisteen paketti alkaa kyllä sillä että tutustutaan vesihuoltolakiin tutustutaan vesivaroihin on kaiken kaikkiaan ja vesistöjen moninaiskäyttö asioihin

- Puheentunnistuksen tuottamia tekstityksiä ei muokattu, ainoastaan rytmitettiin etenemään yhtäläisesti
- Lopullista tekstitystä ei tuotettu, tavoitteena oli verrata kolmen erilaisen puheentunnistuksen tuottamaa tekstiä manuaalisesti tuotettuun tekstiin

Aineiston analysointi

- Aineistolähtöisyys: tunnistettiin aineistosta toistuvia virheitä ja luokiteltiin niiden mukaan

	KOMMENTTI	SANAN VIRHETUNNISTUS							toisto tai täytesana poistettu		puhe
		SANAN VIRHETUNNISTUS	YKSITTÄISIÄ EPÄSANOJA	MUU KUIN TÄYTESANA	SANA JÄÄ POIS	SISÄLTÖÄ JÄÄ PALJON POIS	YLIMÄÄRÄISIÄ SANOJA	toisto tai täytesana poistettu			
seminaarissa käsitellyistä asioista.		x			x						
lyistä asioista.		.			x						
sa käsitellyistä asioista		x			.						

Toistuvia virheitä

- Yksittäisen sanan virhetunnistus
 - Korvautuminen merkityksellisellä, mutta väärällä sanalla
 - Korvautuminen ”epäsanalla”
 - Sana jää kokonaan pois.
 - Paljon sisältöä jää pois
- Isoja kirjaimia väärissä paikoissa
- Pieni kirjain aloittaa lauseen
- Merkitys muuttuu
- Lauserakenne virheellinen
- Torsolauseet
- HUOM!
 - Puheentunnistus tekee myös hyödyllistä tiivistystä: puhekielen täytesanat pois
 - Puhekieli → Kirjakieli



Sanan virhetulkinta ja/tai koko merkityksen muuttuminen

elikkä Soul...Soulissa asuu noin 10 miljoonaa ihmistä, joten nähtävää ja koettavaa taatusti riitti.

Eliikkäs Oulussa asuu noin 10 miljoonaa. Nähtävää ja koettavaa taatusti re.

eli Soulissa asui noin 10 miljoonaa ihmistä joten nähtävää ja koettavaa. Taatusti riitti.

elikkä se oli se asuu noin kymmenen miljoonaa ihmistä en nähtävää ja koettavaa taatusti riitti

Apilapelto ei kestä kolmatta säilörehukorjuuta. Helppoutena on jo ensimmäisenä vuonna pellon hyvä kasvusto.

Apilapelto IKEAsta kolmatta täällä. Helppo uutena on jo ensimmäisenä vuonna pellon hyvä.

apilapelto ei kestä kolmattaynnä rehunkorjuuta. helppohoitona on jo ensimmäisenä vuonna pennun hyvä käskystä.

apilapelto ei kestä kolmatta säilörehun korjuuta helppoutena on jo ensimmäisenä vuonna pellon kasvusta

Mulla on kysymys sinulle ja se on että millaisia ajatuksia nää kuvat herättää sinussa johtamisesta ja johtajista

Mulla on kysymys. Sinulle ja se on että millaisia ajatuksia nää kuvat herättää. Sinussa johtamisesta. Johtaja.

Mulla on kysymys niin ole se on että millaisia ajatuksia nämä kuvat herättää Finnairissa johtamisesta ja johtajista

mulla on kysymys sinulle isonettä millasia ajatuksia näin kuvat herättää sinussa johtamisesta ja johtajista



Muu kuin täytesana jää pois

Tällä videolla käydään läpi johtamisen ja johtajuuden ja organisaation peruskäsitteitä

Tällä videolla käydään läpi johtamisen ja johtajuuden organisaation.

tällä videolla käydään läpi johtamisen ja johtajuuden ja organisaation peruskäsitteitä.

tällä videolla käydään läpi johtonsa ja johtajuuden organisaation peruskäsitteitä



Isot kirjaimet, puuttuu tai liikaa

Tämä tässä käsitellään myös siis kaivot lähteet ja tuota porakaivoasiat myös se veden hankintapuoli mutta on koulutus selkeästi keskittynyt tähän ajankohtaiseen kiinteistökohtaisen jäteveden käsittelyjärjestelmien suunnitteluun

Tässä käsitellään myös siis kaivot lähteet ja ja Tuota on Porakaivo. Myös set vedenhankinta puoli mutta koulutus selkeästi keskittynyt tähän ajankohtaiseen kiinteistökohtaisen jäteveden käsittely järjestelmien suunnittelu.

jossa käsitellään myös kaivot lähteet ja porakaivot myöshankintapuoli mutta on koulutus selkeästi keskittynyt tähän ajankohtaiseen kiinteistökohtaisiin jäteveden käsittelyjärjestelmiin suunnitteluun

tässä käsitellään myös siis kaivot lähteet hihaa porakaivoasiat myös se noiden hankintapuun mutta koulutus selkeästi keskittynyt tähän ajankohtaseen kiinteistökohtaisen jäteveden käsittelyjärjestelmien suunnitteluun

Taustametelin vaikutus...

(Vaihtuu uusi miespuhuja, navetan taustameteliä, hakkausta). No itse asiassa siihen on aika montaki ajatusta elikkä tuota ensimmäinen tietysti on että haetaan kustannussäästöjä, lannotteiden käytössä,

Montakin ajattelin sitten tota ensimmäinen tietysti että laita ... ja sitten käänny Sitten

vitsaili aika monta ja kaikkilla. Ensimmäinen tietokone on aika vaikea ja lampaiden käytössä

löylyt toinen palvelu olja blondi luulet rahaa lunta ollu kunules seuroilla v**** sinne taas vaikeaa puolassa esimerkiksi sitten saada se on jäsenmäärä kiva ajaa esimerkiksi vaikka olo siellä vai joki oon oon joki pitää olla esimerkiksi mies niilo lol mustuaisen tarvitaan historiaa mielentilaan näin mutta pyrin on jannen tiedottaa vai sinne



Johtopäätökset

- ASR –tekstitys ei aineiston perusteella ole valmis kelvollinen tekstitys
- Laadukkaassa kuvaustilanteessa ja selkeällä, suunnitellulla puheella voidaan tuottaa video, jonka ASR-tekstitystä voidaan käyttää hyvänä pohjana valmiin tekstityksen laatimiseksi
- Hyvin usein manuaalisesti tuotettu tekstitys on nopeampi tapa päästä hyvään lopputulokseen tekstityksen osalta
- Kehittyvien tekoälyratkaisuiden myötä ASR-tekstityksen laatu tulee kehittymään merkittävästi

