

Päivitykset

22.10.2024	README luotu; Projektin kuvaus, Jatkokehitys ja Erityismaininnat lisätty
25.10.2024	Asennus- ja käyttöohjeet lisätty
18.11.2024	Lisätty ominaisuudet Order, Match ja Artiklojen tarkkuudella luokittelu README päivittäminen

README – Direktiivikone V1.0

Direktiivikone on alustavasti tehty Oikeusministeriön sisäiseen käyttöön. Ohjelman tarkoitus on merkittävästi vähentää tekstin vertailuun ja kopioimiseen muutoin kuluva aikaa. Direktiivikone automatisoi voimassa olevan direktiivin vertailun, tähän tulleen muutosehdotuksen kanssa, hyväksikäyttäen ehdotuksessa liitteenä olevaa vertailutaulukkoa.

Projektin kuvaus

Ominaisuudet

1. Lukee, luokittelee ja tallentaa tekstisisällön annetuilta sivuilta
2. Lukee vertailutaulukon uuden ehdotuksen liitteistä
3. Yhdistää kerätyn tekstidatan vertailutaulukon mukaisesti yhdeksi *Pandas Dataframeksi*
4. Vertailee, onko ehdotuksessa tullut sisällöllistä muutosta voimassa olevaan
5. Tallentaa luodun taulukon Excel-tiedostona

Direktiivikone lukee [EUR-Lex](#) sivulta, käyttäjän antamilta sivuilta tekstidatan HTML-muodosta. Kone lukee sivuja osa kerrallaan, ja luokittelee datan sen mukaan, kuuluuko kyseinen osa artiklaan, vai johonkin sen alaosiin. Kerätyt tekstidatat tallennetaan omiin tietorakenteisiin vertailua varten.

Direktiivikone formatoi uuden ehdotuksen liitteenä olevan vertailutaulukon muotoon, joka on hyödyllinen tekstisisällön vertailun kannalta.

Direktiivikone käsittelee datan siten, että syötetyt tiedot pystytään yhdistävän ID:n (esim. 1(2)(a)) avulla yhdistämään yhdeksi tietorakenteeksi, josta ilmenee artiklojen välinen suhde, tekstisisältö, sekä onko tekstisisällössä muutoksia.

Lopuksi tulos tallennetaan Excel-tiedostona jatkokäyttöä varten.

Asennusohjeet

Alustava suunnitelma ohjelman jakamiseksi rajatummalle käyttäjäkunnalle on jakaa ohjelmaan liittyvät tiedostot Direktiivikone Teams-kanavalla. Ohjelman käyttäminen ohjeiden laatimishetkellä edellyttää tiedostojen siirtämistä käyttäjän työkoneelle.

Kehitysvaiheessa todettiin myös, että ohjelman ajon voisi suorittaa esim. **Aleksi Kirjonen** tai **Saara Sahinoja**.

Käyttöohjeet

Suorittaakseen ohjelman omalla työasemallaan käyttäjän tulee:

1. Asentaa Python työasemalle Software Centeristä
2. Asentaa seuraavat kirjastot, joita ohjelma käyttää. Linkit lisätty, mikäli kiinnostaa katsoa kirjastojen ominaisuuksia tarkemmin.
 - a. Pandas (<https://pypi.org/project/pandas/>)
 - b. Numpy (<https://pypi.org/project/numpy/>)
 - c. Openpyxl (<https://pypi.org/project/openpyxl/>)
 - d. BeautifulSoup (<https://pypi.org/project/beautifulsoup4/>)
 - e. Requests (<https://pypi.org/project/requests/>)
3. Siirtää Teams-kanavalla olevat tiedostot:
 - a. Main.py
 - b. TextExtract.py
 - c. TableExtract.py
 - d. Combine.py

Omalle työasemalle yhteen kansioon

Ohjelman käynnistäminen

1. Avaa komentokehote kansiossa, johon olet edellä mainitut, .py-päätteiset tiedostot tallentanut. Tähän yksinkertainen tapa on ensin avata

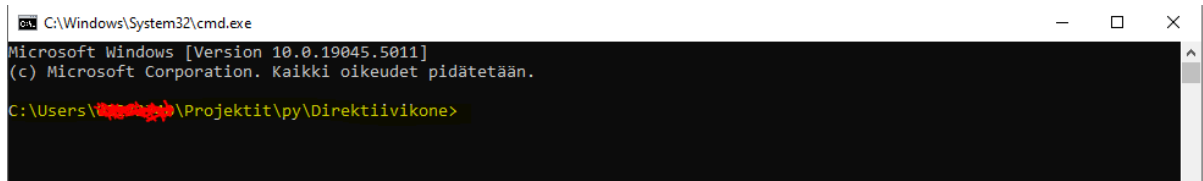
Resurssienhallinta.



2. Resurssienhallinnassa siirry resurssienhallinnassa kansioon, jossa Direktiivikone sijaitsee.
3. Kirjoita ikkunan yläreunassa olevan tiedostopolun tilalle **'cmd'** ja paina Enter. Komentokehotteen pitäisi aueta kyseisessä kansiossa ohjelman käynnistämistä varten.

Tiedät, että olet toiminut oikein kun, tiedosto polku lukee komentokehotteessa,

kuten korostettuna alla



```
C:\Windows\System32\cmd.exe
Microsoft Windows [Version 10.0.19045.5011]
(c) Microsoft Corporation. Kaikki oikeudet pidätetään.
C:\Users\Toni Tähkisen\Projektit\py\Direktiivikone>
```

4. Käynnistä ohjelma kirjoittamalla komentokehotteeseen:

python main.py

5. Seuraa komentokehotteeseen tulevia ohjeita.

Jatkokehitys

Roomalaisin numeroin merkittyjen alakohtien lukemisen lisääminen

Tiedon lukemisen kehittäminen sen pohjalta, miten lähtömateriaalien kehittäminen etenee.

Ongelmatilanteet

Ohjelman k

Erityismaininnat

Suunnitelma Direktiivikoneen toteutuksesta on tehty **Julius Korhosen** tekemän työn pohjalta.

Tavoitteet ohjelmalle on asetettu **Toni Tähtisen** ohjeiden mukaan.

Direktiivikoneen toimintaa on arvioitu yhteistyössä **Saara Sahinojan, Aleksi Kirjosen** ja **Annica Helsingiuksen** kanssa.

Suunnittelusta, ohjelmoinnista ja testauksesta vastuussa on **Heikki Martikainen**.