

Automaattisen sisällönkuvailun työryhmän kokous

31.3.2022 klo 9.15-11

Paikalla: Mirja Anttila, Outi Hintikka, Juho Inkinen, Mona Lehtinen, Ilona Lindfors, Kristiina Näyhö, Jaakko Tuohiniemi, Tarja Turunen ja Leena Vainikainen

Ohje puoliautomaattiseen sisällönkuvailuun

Sovittiin, että Mona alkaa työstää dokumenttia kappaleesta ”Puoliautomaattinen ja automaattinen sisällönkuvailu” ja laittaa ryhmälle sähköpostitse tiedon kun tekstiä on sen verran, että sitä kannattaa kommentoida.

Ohjeen julkaisupaikalle listattiin vaihtoehtoiksi [Finto AI:n aswi-sivu](#), Annif.org-sivusto, Finto AI -lomakkeen sivu, [kuvailuohjeiden sivusto](#) (mahdollisesti osana Sisällönkuvailuopasta tai erikseen) ja [Sisällönkuvailun asiantuntijaryhmän aswi-sivu](#). Keskusteltiin varsinaisen ohjedokumentin julkaisusta Doriassa, mutta päädyttiin siihen, että päivitettävä verkkosivu voisi olla toimivampi. Ohjeeseen voidaan linkittää useammasta paikasta. Jos ohje julkaistaisiin sekä Doriassa, että erikseen verkkosivulla, on riskinä se, että dokumentit alkavat eriytyä sisällöltään. Päätetään asiasta kun ohje on valmis.

Keskusteltiin ohjeen kohderyhmästä – ammattikuvailijat vai opiskelijat ja muut ei-ammattilaiset. Tarja huomautti, että ei-ammattilaiskäyttäjille ala termistö on vierasta ja tekstiä mahdollisesti muokattava sopivaksi. Päätettiin kuitenkin, että pyritään tekemään yksi dokumentti, joka soveltuu laajasti käytettäväksi. Tarvittaessa sen voisi luetuttaa esim. opiskelijoilla etukäteen ja lisätä linkkejä sivuille, joilla aihetta on avattu laajemmin.

Finto AI:n käyttöliittymän uudistus

Juho ja Mona esittelivät eri vaihtoehtoja lomakkeen uudistuksen muotoilusta. Ryhmä kannatti vaihtoehtoa 4, jossa tekstin kopiointi ja tiedoston raahaus ovat samassa laatikossa ja tiedoston lataus erikseen. Vaihtoehtoja voi vielä kommentoida osoitteessa https://docs.google.com/presentation/d/1jRcwDf9Q-2BI9DaBAcziyAX2xxN0UXrVocbU0yuzmrE/edit#slide=id.g11eaac10cf5_1_84.

Huomattiin, että on tarvetta myös toiminnallisuudelle, jossa teksti voidaan syöttää antamalla lomakkeelle URI (tai raahaamalla linkki).

Kielivalinnoista kannatettiin vaihtoehtoa, jossa olisi erikseen valittavana tekstin kieli, käytettävä sanasto tai luokitus ja ehdotusten kieli. Kannatettiin myös tekstin kielen automaattista päättelyä. Ehdotusten kieli voisi olla oletuksena sama, kuin käyttöliittymän kieli. Pidettiin parhaana, että myös raahattaessa tai ladattaessa tiedosto, ehdotukset saa näkyviin vasta klikkaamalla ”anna aihe-ehdotukset” -valintaa.

Keskusteltiin tapauksista, joissa aineisto on kopiosuojattu. Tällä hetkellä näissä tilanteissa Kuvailtava teksti -laatikko jää tyhjäksi, mutta myös erillinen virheilmoitus voisi olla hyödyllinen. Sovittiin että kerätään ja testataan ongelmatapauksia, esim. tilanteita, joissa manuaalisesti on mahdollista kopioida vain yksi sivu tekstiä kerrallaan (toimiiko kuitenkin ladattaessa/raahattaessa oikein).

Testiversio: <http://ai.dev.finto.fi.ext.ocp-kk-test-0.k8s.it.helsinki.fi/?locale=fi>.

Ryhmän seuraava kokous pidetään toukokuussa.