

DARBA UZDEVUMS

PĒTNIECĪBAS PROJEKTAM "IT SISTĒMU LIETOTĀJU RADĪTA DIGITĀLĀ SATURA ANALĪZE, NOVĒRTĒŠANA UN KLASIFICĒŠANA, IZMANTOJOT AI/ML PIEEJU"

1. Ievads

SIA ABC software sadarbībā ar SIA Software Plus īsteno pētniecības projektu par tematu, kas saistīts ar digitāla satura novērtēšanu.

Pētniecības projekta nosaukums: "IT sistēmu lietotāja radīta digitālā satura analīze, novērtēšana un klasificēšana, izmantojot AI/ML pieeju."

Pētniecības projekts tiek veikts projektā "Informācijas un komunikācijas tehnoloģiju kompetences centrs digitalizācijai", projekta ID Nr. 2.2.1.3.i.0/1/24/A/CFLA/006, (turpmāk – Projekts), ko īsteno saskaņā ar 2024. gada 9. janvāra Ministru kabineta noteikumiem Nr. 34 "Latvijas Atveseļošanas un noturības mehānisma plāna 2.2.reformu un investīciju virziena "Uzņēmumu digitālā transformācija un inovācijas" 2.2.1.3.i. investīcijas "Atbalsts jaunu produktu un pakalpojumu ieviešanai uzņēmējdarbībā" īstenošanas noteikumi".

2. Pētniecības projekta mērķis

Pētniecības projekta mērķis ir izstrādāt metodi, kas spēj precīzi analizēt un novērtēt IT sistēmu lietotāju darbības un to veidoto digitālo saturu, pamatojoties uz digitālās mijiedarbības datiem.

3. Pētniecības projekta uzdevumi

- Literatūras apskats, pieejamo risinājumu un metožu izpēte. Nepieciešams apkopot un novērtēt pieejamos pētījumus par digitālā satura analīzi, lietotāju iesaisti un digitālo aktivitāšu uzraudzību un novērtēšanu.
- Lietotāju mijiedarbības datu vākšana, pirmapstrāde un analīze.
- Uz mašīnmācīšanos balstīta atbilstības novērtēšanas algoritma izstrāde. Esošu mašīnmācīšanās algoritmu pielāgošana, automātiskai mijiedarbības datu analīzei un novērtēšanai, atbilstības pakāpes iegūšanai.
- Modeļa apmācība un validācija: pielietot izstrādātos algoritmus modeļa izveidei ar apkopotajiem datiem un veikt to validāciju, lai nodrošinātu augstu precizitāti un uzticamību.
- Metodes integrācija un testēšana:
 - izstrādāto pieeju integrēt alfa versijas prototipā un to validēt mākslīgā vidē;
 - izstrādāt prototipa beta versiju un to validēt reālai videi pietuvinātā vidē.

Pētījuma mērķu sasniegšanai tiek izsludināts iepirkums un tā ietvaros iepirkta zinātniskās institūcijas ekspertu (turpmāk tekstā - pētnieki) pakalpojumi.

4. Pētnieku darba uzdevumi

4.1. Galvenās aktivitātes:

Pētniecības projekta uzsākšanas metodes izpētes gatavība ir atbilstoši TRL4. Pētniekiem, sadarbojoties ar SIA ABC software un SIA Software Plus ekspertiem, būs jāveic šādi darbi:

- veikt Rūpniecisko pētījumu (RP), kura ietvaros metodi ir plānots attīstīt un validēt laboratorijas vidē, sasniedzot TRL5 līmeni;
- veikt Eksperimentālo izstrādi (EI), kuras ietvaros metodes realizāciju ir paredzēts sākotnēji pārnest mākslīgā vidē sasniedzot TRL 6 līmeni un pēc tam izstrādāt jaunās metodes augstas gatavības prototipu, sasniedzot TRL 8 līmeni.

4.2. Detalizēts veicamo darbu un atbildību sadalījums

Aktivitāte Nr.1. pētījumam “IT sistēmu lietotāja radīta digitālā satura analīze, novērtēšana un klasificēšana, izmantojot AI/ML pieeju”

(Starpposma rezultāta plānotais termiņš – 31.10.2025.)

Rūpnieciskā pētījuma laikā plānots:

Darbība	Izpilda Zinātniskās institūcijas pētnieki	Izpilda SIA ABC software un SIA Software Plus eksperti
Analizēt saistīto zinātnisko literatūru par algoritmiem / metodēm / pieejām, kas tiek izmantotas pētījuma uzdevuma risināšanā	X	
Sagatavot un apstrādāt turpmākai analīzei datu kopu, kas saturēs sintētiskos datus, kas imitēs IT sistēmu lietotāju radītu digitālo saturu. Sintētiskie dati saturēs datus par dažādiem tematiem, lai būtu iespēja trenēt algoritmus noteiktu tematu atpazīšanai. Digitālā satura datu formāti būs pietuvināti populārāko biznesa lietojumu radītā satura datu formātam, piemēram, e-pasts, MS Office lietojumprogrammas, tērzēšanas programmas u.tml.		X
Konstruēt digitālā satura analīzes modeļus un veikt laboratorijas eksperimentus, optimizēt analīzes algoritmus. Eksperimentos noteikt pamata radītājus, ar kādiem visefektīvāk var veikt dažāda satura analīzi un tā novērtēšanu par atbilstību noteiktām tēmām, piemēram, finanses un nodokļi, jurisprudences, inženierprojektēšana, farmācija, privātus rakstus ar saturu.	X	

Izveidot tehnoloģiskos komponentus datu kopas glabāšanai, apstrādei un rezultātu saglabāšanai:

Datu glabātuve sākumdatu, starprezultātu un gala rezultātu glabāšanai		X
AI/ML algoritmi digitālā satura analīzei un klasifikācijai	X, Realizācijai varētu izmantot: • Tikai pozitīvus piemērus, proti, dati satur nelielus piemaisījumus ar negatīviem piemēriem (jānodrošina stingrs <i>labeling</i> šādiem datu eksemplāriem); • Citi uzņ. vai projektu artefakti (dokumenti, oficiāli ziņojumi utt); • Satura analīzei izmanto artefaktus, kurus var dabūt <i>plaintext</i>: ○ Recent files (docx, pdf, txt, xls) piemēram, 1. KB; ○ Nosūtīti/saņemti e-pasti, 1.KB; ○ Pārlūka apmeklētas lapas (vismaz publiskām lapām), 1.KB. Vai var nodrošināt arī to saturu, ko ievada lietotājs apmeklēta lpp? • <i>Teams</i> tērzēšanas saturs.	
Komponents vai matemātiskais algoritms, kas nodrošinās eksperimentos iegūto rezultātu salīdzināšanu	X, Realizācija varētu ietvert (izmantot): ○ metodi līdzīgu <i>Siluethe</i> klasterēšanai (metožu rezultātu salīdzināšanai);	

	○ Izmantojot, piemēram, <i>cross</i> validāciju vai līdzīgo pieeju; ○ Jūtīgums un specifiskums.	
Veikt metodes un tās atbalsta komponentu pārbaudi, paredzot šādus eksperimentu veidus laboratorijas vidē: • Eksperimenti ar tehnoloģiskiem komponentiem; • Eksperimenti ar tehnoloģisko komponentu savstarpējo integrāciju.	X	X
<u>Sasniegtie rezultāti:</u>		
Izstrādāta metode digitālā satura analīzei un novērtēšanai laboratorijas līmenī, kas darbojas ar sintētiskiem datiem	X	
Izveidoti tehnoloģiskie komponenti izstrādātas metodes atbalstam un veikta galveno komponentu integrācija, un to kopdarbība ir pārbaudīta laboratorijas vidē	X	X
Sagatavots eksperimentu apraksta dokuments	X	X

Aktivitāte Nr.2 pētījumam "IT sistēmu lietotāja radīta digitālā satura analīze, novērtēšana un klasificēšana, izmantojot AI/ML pieeju".

(Starpposma rezultāta plānotais termiņš – 28.02.2026.)

Aktivitātes ietvaros ir plānots:

Darbība	Izpilda Zinātniskās institūcijas pētnieki	Izpilda SIA ABC software un SIA Software Plus eksperti
<u>Veikt metodes tehnoloģisko komponentu pārnesei no laboratorijas vides uz mākslīgi radīto vidi</u>		
Laboratorijas vide tiks nodrošinātā ABC software datu centrā un/vai Azure makoņskaitļošanas platformā		X
Tiks nodrošināti ETL procesi apmācības datu krātuves aizpildīšanai		X
Metodes realizācija izvēlēta platformā	X	
Modeļa apmācība un uzturēšana (ar papildus apmācību)	X	
Jauno piemēru testēšana un kvalifikācija (piemēram, izmantojot kāds API)	X	
Vizuāla rezultātu prezentācija		X
Jānodrošina visu izveidoto procesu horizontālo mērogošanu (pēc iespējas)	X	X
Izstrādāt digitālā satura analīzes modeļa atbalsta prototipa primārās funkcionālās prasības		X
Izstrādāt algoritma funkcionāls apraksts	X	
Integrēt jaunās metodes funkcionalitāti prototipa alfa versijā		X
Validēt tehnoloģiju mākslīgi radītā vidē	X	X
Apkopot pētījuma starprezultātus un sagatavot divu zinātnisko publikāciju sākotnējās versijas	X	
<u>Sasniegtie rezultāti:</u>		
Izstrādāta jaunās metodes tehniskā specifikācija		X
Izstrādāta jaunās metodes prototipa alfa versija un veikta funkcionalitātes validācija mākslīgā vidē	X	X

Aktivitāte Nr.3 pētījumam **“IT sistēmu lietotāja radīta digitālā satura analīze, novērtēšana un klasificēšana, izmantojot AI/ML pieeju”** .

(Gala rezultāta plānotais termiņš – 30.06.2026.)

Aktivitātes ietvaros ir plānots:

Darbība	Izpilda Zinātniskās institūcijas pētnieki	Izpilda SIA ABC software un SIA Software Plus eksperti
Sagatavot reālai videi pietuvinātu vidi ar gandrīz reāliem datiem	X, par savu komponenti	X, par savu komponenti
Sagatavot un publicēt divus zinātniskos rakstus, kas indeksēti Web of Science, SCOPUS, ERIH (A vai B) ScienceDirect vai Elsevier datubāzēs; MKN 34 70.3.punkts	X	
Pilnveidot tehnisko specifikāciju, papildinot to ar uzlabojumiem, kuru nepieciešamība konstatēta pēc aktivitātē 2 prototipa alfa versijas testēšanas		X
Pilnveidot jaunās metodes prototipu un veikt tā darbības pārbaudi līdz tā galīgajai formai – prototipa beta versija	X, par savu komponenti	X, par savu komponenti
<u>Sasniegtie rezultāti:</u>		
Iesniegtas publicēšanai divas zinātniskās publikācijas, atbilstoši MKN34 70.3.punktam.	X	
Izveidota digitālā satura analīzes, novērtēšanas un klasificēšanas metodes prototipa beta versija tās galīgajā formā	X	X
Pēc nepieciešamības aktualizēta digitālā satura analīzes, novērtēšanas un klasificēšanas metodes tehniskā specifikācija		X