

Keaty Siivelt
Justiits- ja digiministeerium

Teie: 22.07.2026 nr 1-17/4722-27

Meie: Kuupäev digitaallkirjas nr 20-50/482-2

**Vastused tehisarulahenduste toetusmeetme
taotluse täiendavatele küsimustele**

Lugupeetud Keaty Siivelt

Edastame vastused tehisarulahenduste toetusmeetme taotluse täiendavatele küsimustele.

1) Keelemudelite kavandatud kasutamine projektis.

Projektis ei kavandata baaskeelemudeli nullist treenimist. Kasutusele võetakse olemasolevad eel-treenitud keelemudelid ning neid kohandatakse FSDN-i ja turuinfo kasutusjuhtudeks süsteemijuhiste, viipemallide, RAG-teadmusbaasi, MCP-tööriistade ning valideeritud näidis- ja testülesannete abil. Piiratud peenhäälestamist teostatakse ainult juhul, kui projekti alguses tehtav võrdlustest näitab, et RAG-i, viipade ja tööriistadega ei saavutata vajalikku kvaliteeti. Peenhäälestamine ei ole projekti vältimatu ega ette lubatud tulemus.

Keelemudeleid kasutatakse nelja konkreetse funktsionaalsuse elluviimiseks:

Andmekoguja juhendaja. Keelemudel vastab FSDN-i metoodikat, Euroopa Komisjoni valideerimisreegleid ja METK-i juhendeid puudutavatele küsimustele. Vastuse kontekst leitakse RAG-teadmusbaasist ning kasutajale näidatakse allikaviited. Kui usaldusväärset allikat ei leita, peab lahendus vastamisest loobuma või suunama küsimuse spetsialistile.

Valideerimistulemuste selgitamine. FSDN-i süsteemi deterministlikud kontrollid ja Euroopa Komisjoni valideerimisreeglid jäävad autoriteetseks. Keelemudel ei otsusta, kas andmed on õiged, vaid selgitab arusaadavas keeles tuvastatud vea võimalikku põhjust, seost teiste näitajatega ja parandusvõimalust. Vajaduse korral kasutatakse eraldi anomaaliatuvastuse või statistilise analüüsi mudelit, mille tulemus antakse keelemudelile selgitamiseks.

Individuaalse tagasiside koostamine. Ettevõtte näitajad, võrdlusrühm, ajalooline areng ja prognoos arvutatakse eelnevalt valideeritud reeglite, SQL-päringute ja statistiliste või masinõppemudelite abil. Keelemudeli roll on muuta kontrollitud arvutustulemused arusaadavaks narratiiviks, mitte arvutada või prognoosida tulemusi iseseisvalt. See aitab automatiseerida 580 valimisse kuuluva majapidamise tagasiside koostamist.

Kontrollitud mittestandardised analüüsipäringud. Keelemudel aitab kasutaja küsimuse tõlgendada ja valida sobiva eeldefineeritud analüüsitööriista. Lahendus ei anna mudelile vaba SQL-juurdepääsu. Päringud tehakse ainult lubatud väljade, rühmituste ja konfidentsiaalsusreeglite piires ning väljastatakse piisavalt agregeeritud tulemus.

MCP roll on standardiseerida keelemudeli ja FSDN-i teenuste vaheline tööriistakiht. MCP-serverid vahendavad näiteks juhendite otsingut, valideerimistulemuste pärimist, võrdlusrühma näitajate arvutamist, prognoosimudeli käivitamist ning kinnitatud agregaatpäringuid. Iga tööriista sisend, väljund, kasutajaõigus, lubatud andmeväljad ja konfidentsiaalsusreeglid on eelnevalt määratud.

Mudelite vahetatavuse tagamiseks kasutatakse MCP-tööriistakihi kõrval eraldi mudeliadapterit või mudelivärvavat, mis võimaldab kasutada sama rakendusloogikaga lokaalset avatud mudelit või lubatud juhul pilvepõhist mudelit. Kui selleks on vajadus, siis konfidentsiaalseid üksikandmeid töödeldakse lokaalses, ReMITK-i, UT HPC või mõne teise IT partneri kontrollitavas keskkonnas. Pilvemudeleid kasutatakse suure tõenäolsusega üksnes avalike, sünteetiliste või pöördumatult anonüümitud andmetega pärast vastava riski- ja andmekaitsehinnangu läbimist.

See arhitektuur seob tegevused otseselt projekti eesmärkidega: vähendada turuinfo ja FSDN andmekogujate nõustamise ja tagasiside koostamise ajakulu, parandada kogutavate andmete kvaliteeti, võimaldada

kontrollitud mittestandardseid analüüse ning muuta FSDN-i majandus-, keskkonna- ja sotsiaalnäitajad põllumajandustootjale arusaadavaks.

Selline fookus on FSDN-i eesmärkidega hästi kooskõlas. Euroopa Komisjoni järgi laienes FSDN 2025. aastal majandusandmetelt keskkonna- ja sotsiaalnäitajatele ning riikidelt oodatakse ettevõtetele võrdlevat tagasisidet sarnaste majapidamiste kestlikkusnäitajate alusel. Samuti peab üksikandmete avaldamine olema välistatud.

2) Täiendatud eelarve kirjeldus.

6. Projekti eelarve ja ajakava	
<i>Lisa iga etapi ning tegevuse juurde tegevuse kirjeldus, algus ja lõpp ning eelarve.</i>	
I etapp	Ettevalmistus, hanked ja väliseksperdi kaasamine Aeg: 01.10.2026 – 31.12.2026 (3 kuud) Esimese etapi hinnanguline eelarve: 70 000 eurot+15% omaosalus
sh	Välise AI-eksperdi hanke läbiviimine ja eksperdi kaasamine meeskonda; esmase meetodilise ülesehituse ja modelleerimisloogika kaardistamine. Kaasame välise AI eksperdi leidmaks parim võimalik lahendus FSDN ja turuinfo andmete töötlemiseks ja raportite (võrdlev tagasiside andmeesitajatele, faktilised, kokkuvõtted) koostamiseks keelemudeli abil. Koostame eksperdi abil auditi, et täpsustada kontseptsiooni detailid. Täpsustame tegevuskava ja kaardistame modelleerimisloogika, et AI lahenduste abil täita kontaktasutusele pandud ülesandeid (kogutud andmete kvaliteedi tagamine (sh juhendaja andmekogujatele (tugineb meetodikal ja erinevate koostatud juhendmaterjalidel, mille alusel vastab tekkivatele küsimustele ning suunab andmekogujaid), tagasiside andmine andmeesitajatele). Eksperdi eeldatav töömaht kuni 100 töötundi. Hinnanguline eelarve: 10 000 eurot+15% omaosalus
sh	Tarkvara ja litsentside hankimine ja vajadusel riistvara (serverite) tehniliste tingimuste koostamine ja hankimine Eksperdiga koostöös valitud lahendusele vastavalt hangime tarkvara ja litsentsid arvestades kaasnevaid hoolduskulusid pikas perspektiivis. Kaalume erinevaid variante (rent, ost). Vaatame üle olemasolevate serverite tehnilised tingimused, vajadusel hangime lisaserveri, mis vastab valitud lahenduse rakendamiseks ja paigutamiseks vajalikele nõuetele. Vajalik riistvara on ka graafikakaardid. Hinnanguline eelarve: 55 000 eurot+15% omaosalus
sh	Andmekaitsealase mõjuhinna (DPIA) algatamine; anonüümimismetoodika ja MCP turvanõuete esmane disain. FSDN ka turuinfo üksikandmed on konfidentsiaalsed, mistõttu anonüümimise andmed enne keelemudelile töötlemiseks andmist. FSDN ja turuinfo andmetele on määratud kaitsetarve. Teeme andmekaitsealase mõjuhinna lähtudes eksperdiga koostöös välja valitud võimalikust lahendusest. Hinnanguline töötundide maht (tooteomanik ja mõju hindaja) 75 tundi. Hinnanguline eelarve: 5 000 eurot+15% omaosalus
II etapp	Taristu käivitamine, andmete ettevalmistamine ja RAG-lahenduste testimine Valmistame ette keelemudelite ja RAG-lahenduste katsetamiseks vajalikud andmed ning seadistame infoturbe- ja andmekaitsealastele vastava käituskeskonna REMITK või sobiva välise teenusepakkuja taristus. Valime mudelite käitamiseks inferentsi- ja orkestreerimislahenduse, mis võimaldab hallata mudelite koormust, versioone ja kättesaadavust.

6. Projekti eelarve ja ajakava	
<i>Lisa iga etapi ning tegevuse juurde tegevuse kirjeldus, algus ja lõpp ning eelarve.</i>	
	Kaardistame FSDN-i juhendid, eeskirjad, Euroopa Komisjoni reeglistiku ja varasemad tagasisidearuanded. Määrame dokumentide laadimise ja uuendamise töövood, tükeldamise, metaandmed, vektorindeksi, versioonihalduse, allikaviited ning vastuste korrektsuse hindamise testid. Loome juhendmaterjalidel põhineva RAG-prototüübi ning katsetame personaliseeritud tagasiside koostamist piiratud testandmestikuga. Struktureeritud FSDN-i andmeid kasutatakse kontrollitud andmepäringute kaudu, RAG-i aga dokumentidest asjakohase teabe leidmiseks. Rakendame testandmetele andmete minimeerimise, agregeerimise ja anonüümimise meetodeid ning hindame andmete kasulikkuse säilimist. Pilvepõhiste mudelitega testitakse ainult nõuetekohaselt anonüümitud andmeid. Valime ja testime inferentsimootorit, kvantimist, kontekstiakna suurust, paralleeltöötlust ja KV-cache'i seadeid ning mõõdame latentsust, läbilaskevõimet ja GPU-mälu kasutust. Aeg: 01.01.2027 – 31.03.2027 (3 kuud) Teise etapi hinnanguline eelarve: 40 000 eurot+15% omaosalus
sh	Keelemudelite taristu seadistamine Seadistame keelemudelite katsetamiseks ja käitamiseks vajaliku taristu REMITK või nõuetele vastava välise teenusepakkuja keskkonnas. Valime mudelite haldamiseks orkestreerimis- ja inferentsilahenduse ning seadistame juurdepääsukontrolli ja logimise. Testime kvantimist, kontekstiakna, paralleeltöötlust ja KV-cache'i ning mõõdame latentsust, läbilaskevõimet ja GPU-mälu kasutust. Hinnanguline eelarve: 10 000 eurot+15% omaosalus
sh	Juhendmaterjalide RAG-lahenduse ettevalmistamine Kaardistame ja valmistame ette FSDN-i juhendid, eeskirjad ning Euroopa Komisjoni reeglistiku. Määrame dokumentide laadimise ja uuendamise töövood, tükeldamise põhimõtted, metaandmed, vektorindeksi ja allikaviidete nõuded. Loome esmase RAG-lahenduse juhendmaterjalidest vastuste leidmiseks ning testime vastuste korrektsust, allikatele vastavust ja usaldusväärsust. Hinnanguline eelarve: 15 000 eurot+15% omaosalus
sh	FSDN-i testandmete anonüümimine ja mudelite testimine Valmistame ette piiratud FSDN-i testandmestiku ning rakendame andmete minimeerimise, agregeerimise ja anonüümimise meetodeid. Hindame taasidentifitseerimise riski ja andmete statistilise kasulikkuse säilimist. Anonüümitud andmetega testime kohalikke ja pilvepõhiseid kommertsimudeleid ning personaliseeritud tagasiside koostamist struktureeritud näitajate ja varasemate PDF-aruanete põhjal. Võrdleme mudelite vastuste adekvaatsust, faktilist korrektsust ja allikatele vastavust. Hinnanguline eelarve: 15 000 eurot+15% omaosalus
III etapp	Mudel-agnostilise integratsiooni- ja MCP-kihi arendamine Arendame integratsioonikihi, mis ühendab keelemudelid FSDN-i API ja kontrollitud andmetööriistadega. MCP-d kasutatakse standardiseeritud liidesena, mille kaudu saavad

6. Projekti eelarve ja ajakava	
<i>Lisa iga etapi ning tegevuse juurde tegevuse kirjeldus, algus ja lõpp ning eelarve.</i>	
	keelemudelid kasutada ainult eelnevalt määratud õiguste ja piirangutega tööriistu. Mudelite vahetatavus tagatakse eraldiseisva mudeliadapteri kaudu. Arendame tööriistad valideerimisvigade tuvastamiseks ja selgitamiseks, individuaalse tagasiside koostamiseks, võrdlusanalüüsiks ning piiratud analüüsipäringute tegemiseks. Keelemudelile ei anta vahetut juurdepääsu PostgreSQL-i andmebaasile, vaid andmeid vahendatakse FSDN-i API või kontrollitud päringute kaudu. Rakendame autentimise ja õiguste kontrolli, andmete minimeerimise ja agregeerimise reeglid, sisendite ja väljundite valideerimise, logimise ning auditeerimise. Liidestame valitud lokaalsed ja pilvepõhised mudelid ning testime mudelite vahetatavust, vastuste korrektsust, jõudlust ja andmepiirangute järgimist. Väliseksperdi kaasabil viime lõpule lahenduse metoodilise ülesehituse, andmekaitse- ja turvareeglid, testimise põhimõtted ning tehnilise dokumentatsiooni. Väliseksperdi kavandatud töömaht on kuni 400 tundi. Aeg: 01.04.2027 – 30.06.2027 (3 kuud) Kolmanda etapi hinnanguline eelarve: 50 000 eurot+15% omaosalus
sh	MCP andmevahetuskahi arendamine Arendame mudel-agnostilise MCP-kihi, mis ühendab keelemudelid FSDN-i API ja kontrollitud andmetööriistadega. Looame komponendarhitektuuri ning tööriistad valideerimisvigade selgitamiseks, individuaalse tagasiside koostamiseks, võrdlusanalüüsiks ja piiratud analüüsipäringuteks. Rakendame õiguste kontrolli, logimise, sisendi ja väljundi valideerimise, testid ning tehnilise dokumentatsiooni. Hinnanguline eelarve: 25 000 eurot+15% omaosalus
sh	Metoodika ja andmepiirangute seadistamine Viime väliseksperdi juhtimisel lõpule AI-lahenduse metoodilise ja tehnilise ülesehituse. Määrame tööriistade kasutusõigused, lubatud andmeväljad, agregeerimise ja andmete minimeerimise reeglid ning väljundite kontrollimise põhimõtted. Testime piirangute toimimist ja dokumenteerime lahenduse turbe-, andmekaitse- ja auditeerimisnõuded. Hinnanguline eelarve: 15 000 eurot+15% omaosalus
sh	Keelemudelite liidestamine ja vahetatavuse testimine Liidestame MCP-kihiga valitud lokaalsed ja pilvepõhised keelemudelid, kasutades pilvemudelite puhul ainult nõuetekohaselt anonüümitud andmeid. Looame mudeliadapteri, mis võimaldab mudeleid vahetada ilma FSDN-i tööriistu ümber arendamata. Testime mudelite funktsionaalset sobivust, vastuste korrektsust, jõudlust ja andmepiirangute järgimist. Hinnanguline eelarve: 10 000 eurot+15% omaosalus
IV etapp	Mudelite valideerimine, testperiood ja turvakontroll Aeg: 01.07.2027 – 30.09.2027 (3 kuud) Neljanda etapi hinnanguline eelarve: 30 000 eurot+15% omaosalus

6. Projekti eelarve ja ajakava	
<i>Lisa iga etapi ning tegevuse juurde tegevuse kirjeldus, algus ja lõpp ning eelarve.</i>	
sh	Pilootkasutus reaalsete andmekogujatega (<i>human-in-the-loop</i>). Välise eksperdi poolt masinõppemudelite ja tagasiside (<i>Data Storytelling</i>) genereerimise loogika valideerimine. Andmekogujate (FSDN ja turuinfo) tööaeg lahenduse kasutamisel ja tagasiside andmisel (150 tundi). Väliseksperdi tööaeg lahenduse eesmärgile vastavuse valideerimisel ja kogujate tagasiside arvestamisel (kuni 150 tundi), Hinnanguline eelarve: 10 000 eurot+15% omaosalus
sh	Valitud lahenduste võrdlemine. Lahenduse kiiruse, stabiilsuse ja tulemuste usaldusväärsuse võrdlev analüüs lokaalsete ja pilvepõhiste mudelite vahel läbi MCP kihi. Testime ja võrdleme lahenduse kiirust ja stabiilsust ning usaldusväärsust pilvepõhiste mudelite erilahenduse puhul. Hinnanguline eelarve: 10 000 eurot+15% omaosalus
sh	Turva- ja lekketestid. Süsteemi turva- ja lekketestid (kontrollimaks, et MCP reeglid ja anonüümimisloogika välistavad isikuandmete/konfidentsiaalsete üksikandmete lekke pilveteenustesse). Testime loodud lahenduse turvalisust ja anonüümimisloogika toimivust. Tööaeg ca 300 tundi. Hinnanguline eelarve: 10 000 eurot+15% omaosalus
V etapp	MVP Live ja lähtekoodi avalikustamine. Läbipaistvuse, AI Acti vastavuse ja lõppdokumentatsiooni koostamine Eelarve on kujunenud hinnanguliselt 5% ulatuses kogu planeeritud eelarve mahust. Sisuliselt METK töötajate tööaeg seoses lähtekoodi lisamisega riiklikusse koodivaramusse, vormi esitamisega kratid.ee registrisse ja lõpparuande koostamisega. Aeg: 01.10.2027 – 31.10.2027 (1 kuu) Viienda etapi hinnanguline eelarve: 10 000 eurot+15% omaosalus
sh	Lahenduse lähtekoodi puhastamine ja avalikustamine riiklikus koodivaramus/GitHubis (ülekanavuse tagamiseks).
sh	Tehisaru Kratid.ee registrisse „Algoritmi kasutuse ülevaate“ vormi esitamine ja läbipaistvuse tagamine.
sh	Projekti lõpparuandlus ja paikvaatluse korraldamine toetuse andjale.

3) Projekti tulemuste kavandatud laiem taaskasutatavus.

Projekti tulemused võimaldavad teistel asutustel olemasolevaid ja arendatud komponente sama arhitektuuri-, turbe- ja integratsioonimustri alusel kiiremini ning väiksema analüüsi- ja katsetamiskuluga kasutusele võtta.

Lugupidamisega

/allkirjastatud digitaalselt/

Katre Kirt

Maamajanduse valdkonna juhataja

Koopia: Monika Karu 5885 1179, monika.karu@justdigi.ee