

# Keelemudelite ja keeletehnoloogia vahendite kasutus e-kirjandite hindamisel: 2024. ja 2025. aasta katseeksamite võrdlus

**Merilin Aruvee, Kais Allkivi**, Andres Karjus,  
Katarin Leppik, Silvia Maine, Krister Kruusmaa

# Projekti eesmärgid ja tulemused

1. Lõime emakeele e-eksami kirjandite tekstikorpuse (kättesaadav [ELLE-s](#))
2. **Uurisime automaathindamise võimalusi ja täpsust kirjandite hindamisel**
  - suurte keelemudelite abil (nii sisu kui ka keelelised aspektid)
  - statistilise keeletötluse abil (lausestus, sõnastus, korrektsus)
3. **Analüüsisime masin- ja inimhindamise erinevusi**
4. Arendasime hindamismaatriksit
5. Tegime 5 Videoõpsi videot toetamaks eksamiülesannet ja katmaks kriteeriume
6. Salvestasime katsetuste koodi [Githubis](#)

# Kaks võrdlevat lähenemist

## 1. Statistiline keeletöötlus (vormilised tunnused)

- Keelekasutust iseloomustavad kvantitatiivsed tunnused
- Leitud spetsiaalsete tööriistadega (nt Stanza, korrektorid)
- Seosed teksti tunnuste ja inimhindajate skooride vahel → ennustused

## 2. Tehisar, suured keelemudelid (kõik tunnused)

- Automaathindamine kontekstiõppe vormis, ei vaja treenimist
- Sisend: hindamisskeem (kategooriate kaupa) + üks kirjand
- Katse: OpenAI GPT ja Google'i Gemini mudelid
- Väljakutse: millega võrrelda? Inimhindajate vahel ka suur variatsioon

## 9. klassi 2024. a kriteeriumid

pealkiri ja sissejuhatus	lausemoodustus
argumentatsioon	sõnavalik
alusteksti kasutamine	kirjavahemärgistus
lõpetus	õigekiri ja vormimoodustus
	liigendus ja vormistus

## 9. klassi 2025. a kriteeriumid

pealkiri	arusaadavus
kontekst	lausestuse sidusus
põhiväide	eksimused
alaväited	mitmekesisus
näited	kirjakeelsus
järeldused	sõnastusraskused
alusteksti kasutamine	kirjavahemärgistus
probleemile vastamine	õigekiri ja vormimoodustus
lõikude järeldused	liigendus
	trükkimine

## 12. klassi 2024. a kriteeriumid

SISU

STIIL

ÕIGEKEELSUS

ÜLESEHITUS

## 12. klassi 2025. a kriteeriumid

pealkiri ja sissejuhatus

argumentatsioon

alusteksti kasutamine

lõpetus

lausemoodustus

sõnavalik

kirjavahemärgistus

õigekiri ja vormimoodustus

liigendus ja vormistus

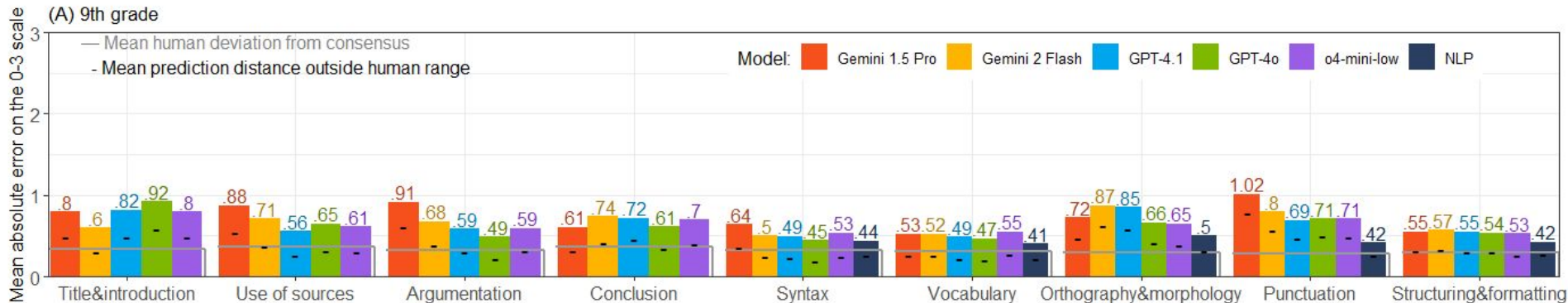


# *Tulemused I*

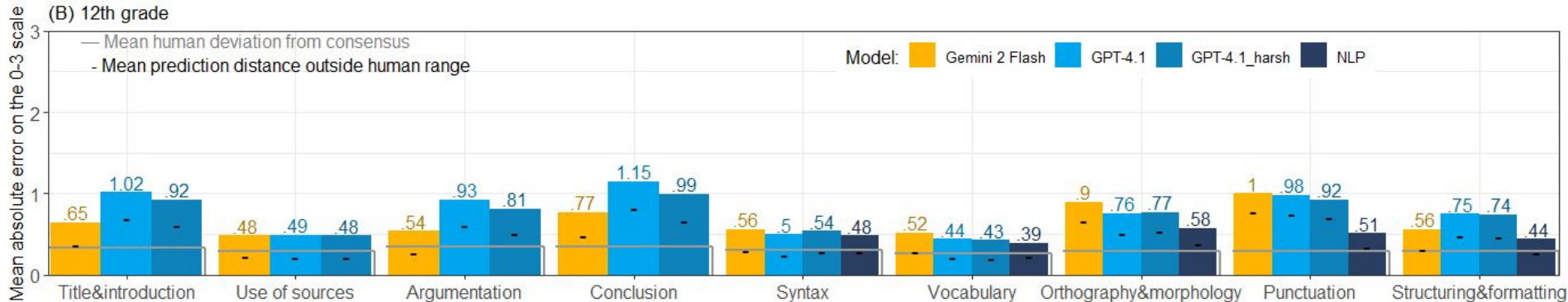
Automaathindamise täpsus

# 2024. aasta, 9. ja 12. klass

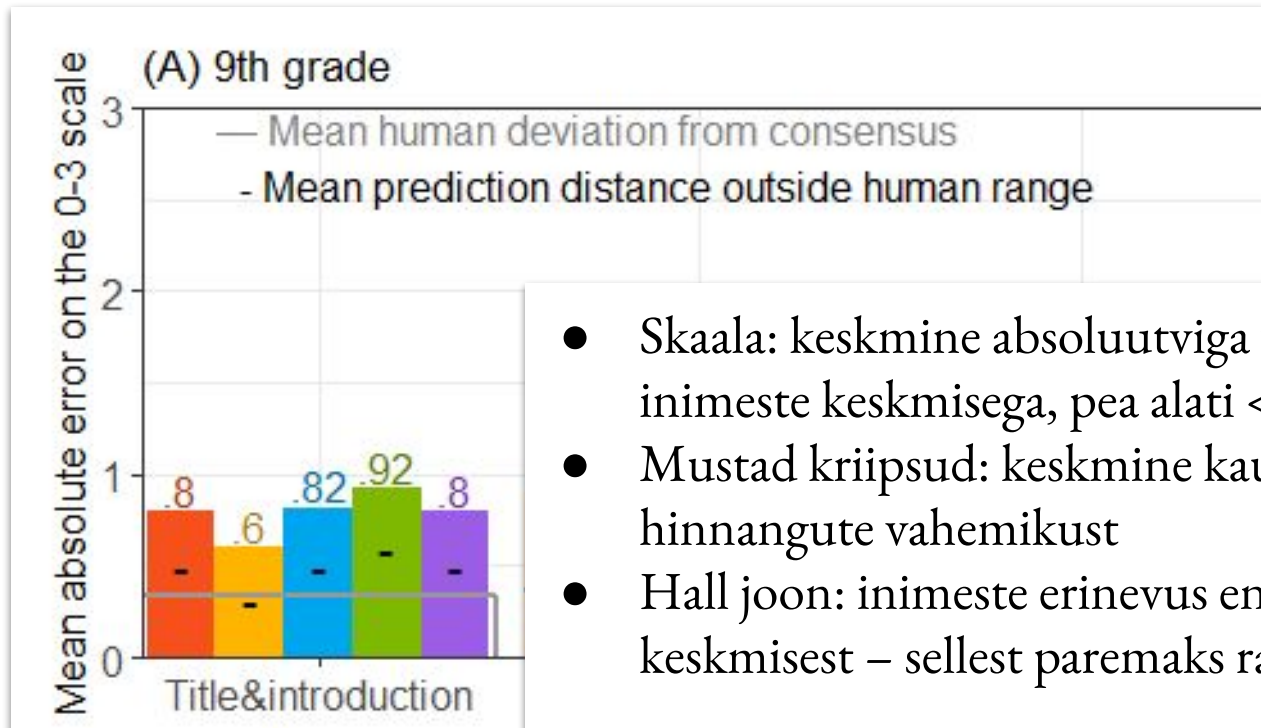
(A) 9th grade



(B) 12th grade



# Kuidas jooniseid lugeda



- Skaala: keskmine absoluutviga võrreldes inimeste keskmisega, pea alati  $< 1$
- Mustad kriipsud: keskmine kaugus inimeste hinnangute vahemikust
- Hall joon: inimeste erinevus enda keskmisest – sellest paremaks raske saada



# Üldpilt ja kallutatatus

Keskmine viga 0–27  
skaalal ~3 punkti

(A) 9th grade				
Model	Bias	<i>p</i>	MAE	In range
(Human average:	16.08)			
Gemini 1.5 Pro	-3.66	***	4.24	46%
Gemini 2 Flash	-1.11	***	2.95	61%
GPT-4.1	+1.71	***	2.94	64%
GPT-4o	+0.46	*	2.66	67%
o4-mini-low	0.00		2.68	68%

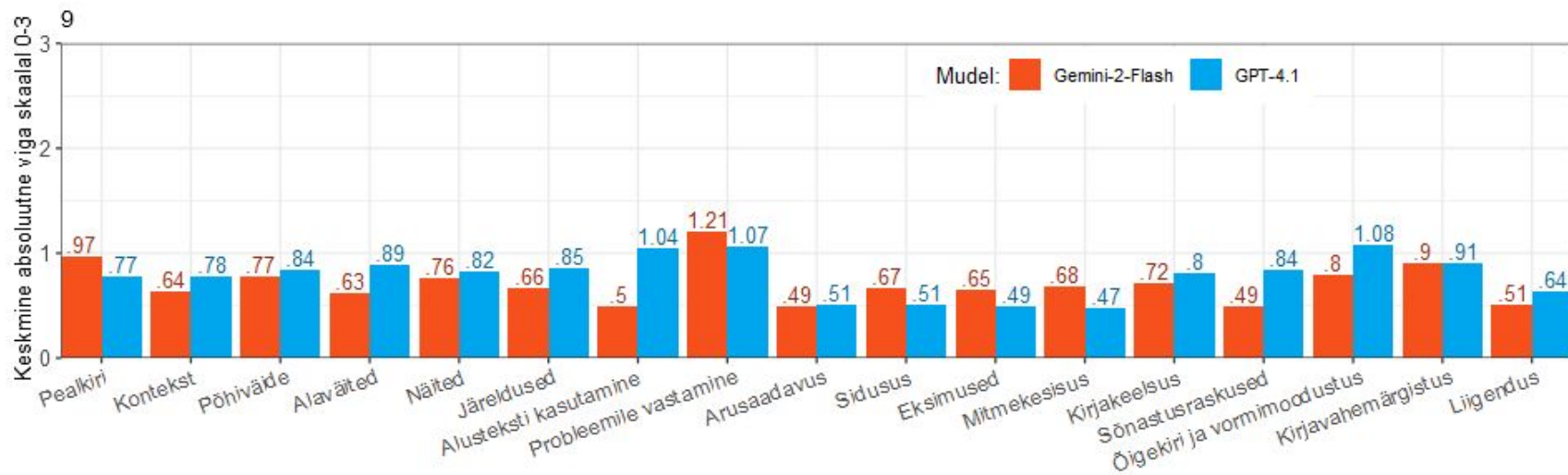
Pea kõigi mudelite lõpphinne jääb enamasti (>60%) kahe inimhindaja vahele

Osadel mudelitel väike (1–3 punkti) positiivne või negatiivne kalle (o4 arutlevat tüüpi mudelil mitte)

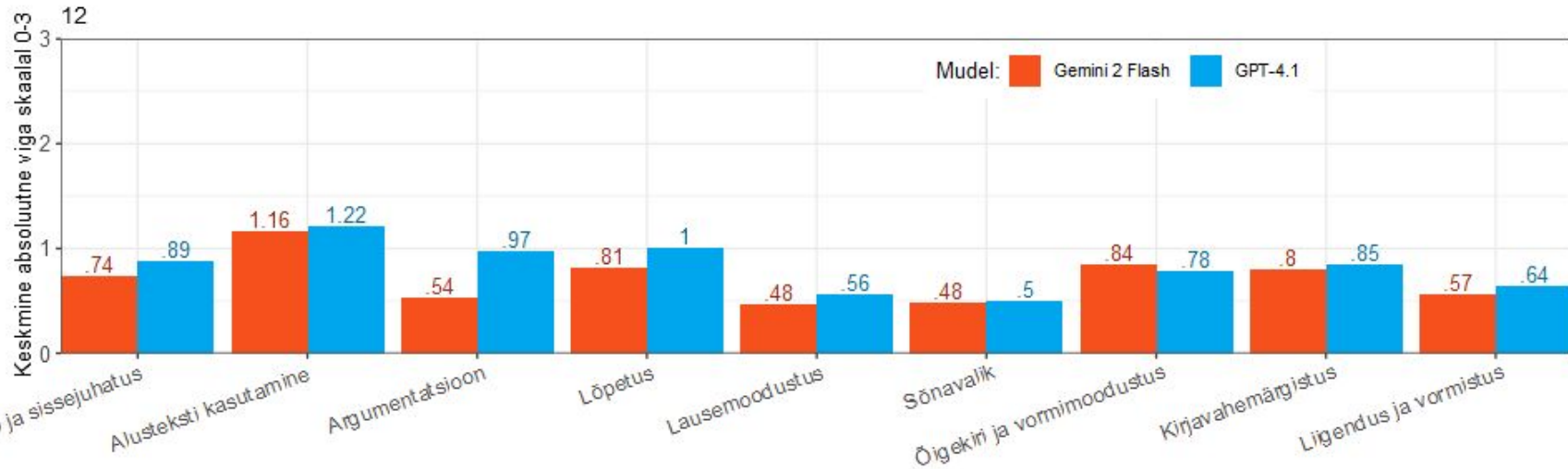


# 2025:

## 9. klass



## 12. klass



TALLIN Digiinstituut

# Mida teada saime?

- Keelemudelid saavad üsna hästi hakkama nii sisuliste kui vormiliste kategooriatega (ja vahepeal on välja tulnud veel paremad mudelid)
- Selleks eraldi treenitud statistilised mudelid on natuke paremad vormilistes kriteeriumites, nagu kirjavahemärgid, õigekiri ja vormimoodustus, lausemoodustus, sõnavalik ja liigendus
- Nii tehisaru kui ka inimesed vajavad häid täpseid juhendeid
- Hindamise automatiseerimine on täiesti võimalik kas inimese toena eksamihindamisel või õppetöö käigus, andes kirjutamist harjutavale õpilasele isiklikku tagasisidet





# *Tulemused II*

Kirjandite keeletunnused



TALLINNA ÜLIKOOL  
Digitehnoloogia  
instituut



TALLINNA ÜLIKOOL  
Humanitaarteaduste  
instituut

# Keeletunnused

- **Morfosüntaktilised:** sõnaliigid ja nende alamliigid, käänd- ja tegusõnade vormikasutus
- **Leksikaalsed:** sõnavara mitmekesisus (sh sõnaliigiti), harvade sõnade osakaal, nimisõnade abstraktsus, sisusõnade osakaal
- **Pindtunnused:** sõnade ja lausete pikkus, keerukusindeksid, lõikude pikkus
- **Veatunnused:** eri tüüpi automaatsete paranduste sagedus eesti keele grammatikakorrektori väljundi põhjal

# 2024. ja 2025. aasta 9. klassi kirjandite võrdlus

- **Meetod:** korrelatsioonanalüüs hindamisaspektide ja keeletunnuste vahel
- **Küsimus:** kui võrd seostuvad sarnaste hindamisaspektide keskmised skoorid kahel aastal samade keeletunnustega?
- **Võimaldas vaadata:**
  - kas skooore mõjutavad stabiilselt ühed ja samad tunnused
  - kas hindajad eristavad uusi detailsemaid aspekte

# Õigekiri ja vormistus

2024		2025	
Kirjavahe- märgistus	<ul style="list-style-type: none"><li>• Puuduvad kirjavahemärgid -0,75</li><li>• Puuduva kirjavahe- märgi paranduste osakaal -0,66</li><li>• Liigsed kirjavahemärgid -0,13</li></ul>	Kirjavahe- märgistus	<ul style="list-style-type: none"><li>• Puuduvad kirjavahemärgid -0,70</li><li>• Puuduva kirjavahe- märgi paranduste osakaal -0,66</li><li>• Liigsed kirjavahemärgid -0,13</li></ul>
Õigekiri ja vormimoodustus	<ul style="list-style-type: none"><li>• Grammatika- parandusi sõna kohta -0,69</li><li>• Sõnaasendused -0,57</li><li>• Kobarparandused -0,56</li></ul>	Õigekiri ja vormimoodustus	<ul style="list-style-type: none"><li>• Grammatika- parandusi sõna kohta -0,65</li><li>• Sõnaasendused -0,62</li><li>• Kobarparandused -0,37</li></ul>
Liigendus ja vormistus	<ul style="list-style-type: none"><li>• Lõikude sõnade arvu st.hälve -0,35</li><li>• Lõikude sõnade arvu vahe -0,34</li><li>• Sõnaasendused -0,29</li><li>• Lõikude lausete arvu st.hälve - 0,22</li></ul>	Liigendus	<ul style="list-style-type: none"><li>• Lõikude sõnade arvu st.hälve -0,33</li><li>• Lõikude sõnade arvu vahe -0,31</li><li>• Lõikude lausete arvu st.hälve -0,21</li></ul>
		Trükkimine	<ul style="list-style-type: none"><li>• Sõnaasendused -0,37</li></ul>



# Sõnavalik



2024		2025	
Sõnavalik	<ul style="list-style-type: none"><li>• Kobarparandused -0,42</li><li>• Sõnaasendused -0,39</li><li>• Keskmise sõnapikkus 0,36</li><li>• Guiraud' indeks 0,32</li><li>• MTLD indeks 0,30</li><li>• [---]</li><li>• Harvade sõnade osakaal 0,30</li><li>• Omadussõnade osakaal 0,28</li><li>• Leksikaalne tihedus 0,25</li></ul>	Mitmekesisus	<ul style="list-style-type: none"><li>• Guiraud' indeks 0,47</li><li>• MTLD indeks 0,45</li><li>• [---]</li><li>• Keskmise sõnapikkus 0,35</li><li>• Harvade sõnade osakaal 0,33</li><li>• Kobarparandused -0,32</li></ul>
		Kirjakeelsus	<ul style="list-style-type: none"><li>• Keskmise sõnapikkus 0,32</li><li>• Kobarparandused -0,30</li><li>• Sõnaasendused -0,28</li><li>• Maasi indeks -0,27</li><li>• Leksikaalne tihedus 0,26</li></ul>
		Sõnastus-raskused	<ul style="list-style-type: none"><li>• Sõnaasendused -0,45</li><li>• Kobarparandused -0,38</li><li>• MTLD indeks 0,29</li><li>• Keskmise sõnapikkus 0,29</li><li>• Maasi indeks -0,27</li></ul>

# Lausemoodustus



2024		2025	
Lause- moodustus	<ul style="list-style-type: none"><li>• Kobarparandused -0,46</li><li>• Keskmise sõnapikkus 0,40</li><li>• Harvade sõnade osakaal 0,34</li><li>• Omadussõnade osakaal 0,32</li><li>• Keskmise lausepikkus -0,29</li><li>• Omadussõna käändevormide arv 0,26</li><li>• Leksikaalne tihedus 0,26</li></ul>	Arusaadavus	<ul style="list-style-type: none"><li>• Kobarparandused -0,44</li><li>• Keskmise sõnapikkus 0,34</li><li>• Harvade sõnade osakaal 0,33</li><li>• Keskmise lausepikkus -0,27</li><li>• Omadussõnade osakaal 0,25</li><li>• Omadussõna käändevormide arv 0,25</li></ul>
		Lausete sidusus	<ul style="list-style-type: none"><li>• Kobarparandused -0,40</li><li>• Keskmise sõnapikkus 0,37</li><li>• Harvade sõnade osakaal 0,33</li><li>• Nimetava käände osakaal -0,27</li><li>• Keskmise lausepikkus -0,26</li><li>• Leksikaalne tihedus 0,26</li></ul>
		Eksimused	<ul style="list-style-type: none"><li>• Kobarparandused -0,42</li><li>• Keskmise sõnapikkus 0,28</li><li>• Harvade sõnade osakaal 0,26</li><li>• Keskmise lausepikkus -0,25</li><li>• Leksikaalne tihedus 0,21</li><li>• Liigsed sõnad -0,19</li></ul>

# 2024 ja 2025 võrdlus, 9. klass, korrelatsioonid üldskooriga

Veatunnuste mõju  
vähenes, keerukuse  
ja mitmekesisuse  
mõju suurenes.

2024	2025
<p>Korrektus:</p> <ul style="list-style-type: none"><li>• Grammatikaparandusi sõna kohta -0,73</li><li>• Grammatikaparanduste arv -0,65</li><li>• Grammatikaparandusi lause kohta -0,63</li><li>• Puuduvad kirjavahemärgid -0,55</li><li>• Kobarparandused -0,53</li><li>• Sõnaasendused -0,43</li><li>• Puuduvate kirjavahemärkide osakaal parandustest -0,37</li><li>• Sõnaasenduste osakaal parandustest 0,33</li></ul> <p>Keerukus/mitmekesisus:</p> <ul style="list-style-type: none"><li>• Keskmine sõnapikkus 0,48</li><li>• Harvade sõnade osakaal 0,38</li><li>• MTLD indeks 0,35</li><li>• Keskmine lausepikkus -0,35</li><li>• Omadussõnade osakaal 0,33</li><li>• Leksikaalne tihedus 0,32</li></ul>	<p>Korrektus:</p> <ul style="list-style-type: none"><li>• Grammatikaparandusi sõna kohta -0,63</li><li>• Grammatikaparandusi lause kohta -0,59</li><li>• Grammatikaparanduste arv -0,52</li><li>• Puuduvad kirjavahemärgid -0,41</li><li>• Kobarparandused -0,41</li><li>• Sõnaasendused -0,37</li></ul> <p>Keerukus/mitmekesisus:</p> <ul style="list-style-type: none"><li>• Keskmine sõnapikkus 0,43</li><li>• Guiraud' indeks 0,41</li><li>• Maasi indeks -0,39</li><li>• MTLD indeks 0,39</li><li>• Uberi indeks -0,37</li><li>• Harvade sõnade osakaal 0,36</li><li>• Korrigeeritud verbivarieeruvus 0,34</li><li>• Leksikaalne tihedus 0,32</li></ul>



# Järeldusi, küsimusi

- 2024. ja 2025. a kirjandite hindamisel mängisid rolli sarnased tunnused – stabiilsus tähendab usaldusväärsust
- Õpetajad eristasid analüütilisemaid kategooriaid, sõnastuse ja lausestuse hindamisel olulised tunnused kattusid mõlemal aastal
- Kas ja milliseid tunnuseid hindajatele kuvada võiks?
- Kas sama tunnuse alusel võib hinnata mitut aspekti?





# *Arutelu*



TALLINNA ÜLIKOOL

Digitehnoloogia  
instituut



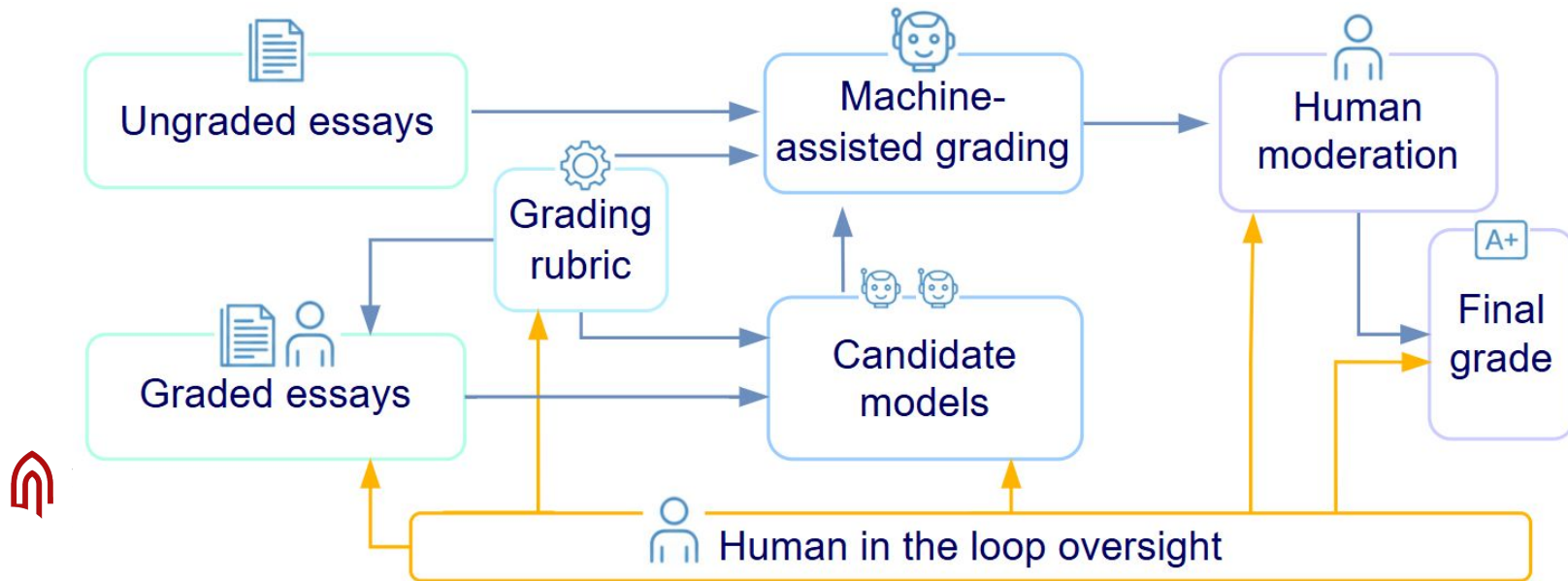
TALLINNA ÜLIKOOL

Humanitaarteaduste  
instituut

# Meie ettepanek: hübriidhindamine

Vt: *Machine-Assisted Grading of Nationwide School-Leaving Essay Exams with LLMs and Statistical NLP*  
<https://arxiv.org/abs/2601.16314>

Tehisaru ja keeletehnoloogia võimaldavad kirjandite jm tekstide kiiret, täpset ja efektiivset hindamist. Inimene jääb alati lõpp-otsustajaks, aga sellise lahenduse integreerimine hindamise protsessi hoiaks oluliselt kokku raha ja aega.



# Meie ettepanekud

- Hübriidhindamise katsetused õpetajatega
  - Kaaluda hübriidhindamise integreerimist EISis
  - Suurema hulga ja vabavaraliste keelemudelite hindamiseks sobivuse evalveerimine (koostöö EKiga)
  - Mitme aasta kirjandite hindamine sama juhendi alusel + uuem keeletehnoloogia → täpsemad ja paindlikumad statistilised mudelid
- Kirjandite keeletunnuste ja alusteksti mõju kvalitatiivne analüüs
- Samal ajal arendada kirjutamise hindamise alast teadlikkust õpetajatel



Esileht



Ülesande loomine



Hindamine



Varem hinnatud



Hindamissüsteemid

## Ülesande loomine

Pealkiri

Klass

Tähtaeg

dd/mm/yyyy



Hindamissüsteem

Ülesande kirjeldus

Ülesande tekst tuleb ka hindamissüsteemist. Siin saab seda täiendada.

Sõnade arv

Alustekstid

SISESTA OMA TEKST

KUIDAS OLLA HEA SÖBER?

PERE JA KODU

Esita ülesanne

Sõprus on üks olulisemaid suhteid inimese elus. Tõelist sõpra iseloomustavad eelkõige usaldusväärsus ja ausus. Hea sõbraga on alati meeldiv aega veeta, sest teil on sageli ühised huvid ja arusaamad. Samuti on oluline, et sõber oskaks hoida saladusi ning oleks olemas nii rõõmsatel kui ka rasketel hetkedel. Just selline vastastikune toetus teeb sõprusest tugeva ja püsiva suhte. Sõprussuhted võivad tekkida juba lapseas ja mõned neist kestavad aastaid. Lapsena veedetakse sõpradega palju aega koos ning sõprust peetakse sageli väga tähtsaks ja lahutamatuks osaks oma elust. Noorukieas muutuvad suhted aga mitmekesisemaks. Inimesed õpivad paremini mõistma nii iseennast kui ka teisi. Samal ajal laieneb ka sõpruskond ning sageli veedetakse aega nii samast kui ka vastassoost sõpradega. Sellistes seltskondades võivad tekkida ka esimesed romantilised suhted. Sõprus õpetab inimesele mitmeid olulisi oskusi. Näiteks õpitakse hoolima teisest inimesest, teda usaldama ja austama. Samuti õpitakse arvestama sõbra soovide ja vajadustega. Need oskused on väga vajalikud ka hilisemates suhetes, näiteks armastussuhtes ja perekonnas. Seega võib öelda, et sõprus valmistab inimest ette tulevasteks lähedasteks suheteks. Lisaks mõjutavad sõprussuhted inimese enesehinnangut. Terved ja toetavad sõprussuhted annavad enesekindlust ning aitavad tunda end väärtuslikuna. Kui inimene teab, et tal on sõbrad, kes teda mõistavad ja toetavad, annab see palju energiat ja julgust igapäevaeluga toime tulla. Samas kaasneb sõprusega ka vastutus. Oluline on hoida sõbra usaldust ning olla ise aus ja hooliv. Samuti tuleb märgata, kui suhtes esineb ebaausust, ärakasutamist või teise inimese tahte pealesurumist. Sellised käitumisviisid ei kuulu tervete sõprussuhete juurde. Tõeline sõprus põhineb vastastikusel austusel, hoolimisel ja usaldusel. Just need väärtused aitavad sõprusel püsida tugevana ja tähendusrikkana kogu elu vältel.

## Teemaarendus

Pealkiri puudub.

Kontekst – 3 p

Sissejuhatus avab sõpruse teema ja tausta selgelt („Sõprus on üks olulisemaid suhteid inimese elus..“).

Probleem ja vastus – 2 p

Tekstis on üldine väide sõpruse olulisusest, kuid selge probleemipüstitus või küsimus puudub. Lõpetus seostub teemaga, kuid pigem

## Lausestus ja sõnastus

Lausestuse selgus – 3 p

Laused on selged, loogilised ja erineva struktuuriga.

Sõnavalik – 3 p

Sõnavara on sobiv ja kirjakeelne.

## Õigekiri ja vormistus

Kirjavahemärgid – 3 p

Vigu praktiliselt ei esine.

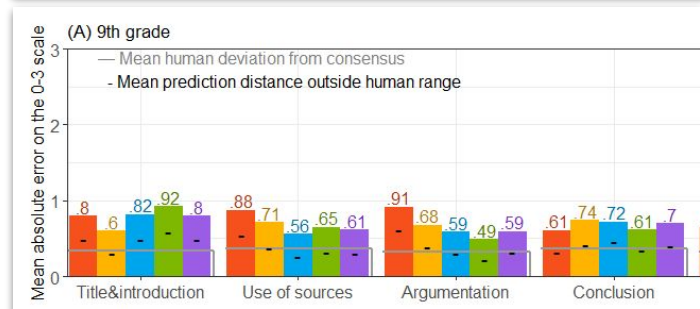
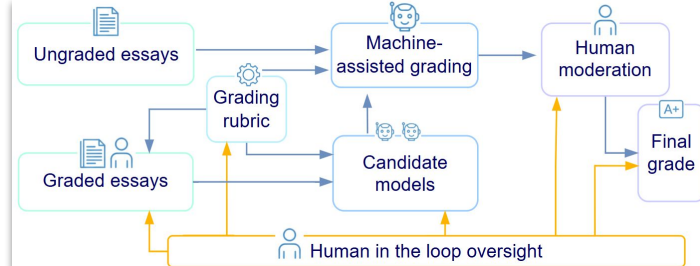
Õigekiri – 3 p

Ortograafia on korrektne.

Lühenäide ja vormistus – 3 p

# Kokkuvõtteks

- Tehisaru kasutamine õpilasteksti hindamisel on võimalik, tegelikuks tarvituselevõtuks tuleb uurimist ja arendustööd jätkata, kaasata õpetajad.
- Potentsiaal suuremaks riiklikuks arenduseks, koostöök: tehisaru usaldusväärsus jne.
- Hindamise kõrval ka kirjutamisõpetuse arendus, materjalid, koolitused.



Õigekiri ja vormistus

Punktid 9

Kirjavahemärgid - 3 p  
Vigu praktiliselt ei esine.

Õigekiri - 3 p  
Ortograafia on korrektne.

Liian... is... m... - 3 p

*Machine-Assisted Grading of Nationwide School-Leaving Essay Exams with LLMs and Statistical NLP*

<https://arxiv.org/abs/2601.16314>