

Märksõnastamise tänaest ja homsest: praktikad ja visioonid

Meelis Somelar
Rahvusraamatukogu
Kogude Arenduse osakond
29. november 2023

Millest tänaräägime

- Vormimärksõnad teemamärksõnaväljadel
- Edasised praktilised plaanid:
 - Üks ühine EMS-i väliste märksõnade väli
 - Ülearuste punktide kaotamine ajamärksõnadest
- Visioonid: kratid ja suured keelemudelid
 - Marta ja Annif
 - Aria , Bard, ChatGPT 3.0

Täitsa ulme!

e ms EESTI MÄRKSONASTIK
ESTONIAN SUBJECT THESAURUS

Valdkonnad | Ajamärksõnad | Vormimärksõnad | Valimid | EMS arvudes | Tellimärgid | Abi | Tagasiside |

Otsin: sõna algust sõna osa täpset vastet eesti keeles OTSI ENG

5 Algus 'ulme' Ülevaade

Prindi | E-post | Salvesta

Piira valdkonnaga

Lisa valimisse

- ulme KASUTA **ulmefilmid**
- ulmekirjandus**
- ulmefilmid**
- ulmejutud**
- ulmekirjandus**
- ulmeromaanid**

ulmefilmid (655)

ulmekirjandus (650)

ulmeromaanid (655)

ulmejutud (655)

ulme

Kasutada ainult kirjandusteaduses temaatilise märksõnana! Vormimärksõnana kasutada ulmeromaanid.

Täits a problema tiline ...

		<input checked="" type="checkbox"/> BIBLIOGRAPHIC	<input type="checkbox"/> AUTHORITY	<input type="checkbox"/> PATRON	<input type="checkbox"/> ITEM	<input type="checkbox"/> ORDER	<input type="checkbox"/> CHECKIN	<input type="checkbox"/> COURSE		
Index		SUBJECT (d)		ulmekirjandus						
#		SUBJECT							# ENTRIES	
<input checked="" type="checkbox"/>	1	b	655 4 aulmekirjandus						1	
	2	a	450 aulmekirjandus						1	
<input checked="" type="checkbox"/>	3	b	654 9 aulmekirjandus						3	
<input checked="" type="checkbox"/>	4	b	655 4 aulmekirjandus						98	
<input checked="" type="checkbox"/>	5	b	692 aulmekirjandus						24	
	6	a	150 aulmekirjandus 0https://ems.elnet.ee/id/EMS027390						1	
	7	a	550 aulmekirjandus 0https://ems.elnet.ee/id/EMS027390						4	
<input checked="" type="checkbox"/>	8	b	650 4 aulmekirjandus 0https://ems.elnet.ee/id/EMS027390						262	
<input checked="" type="checkbox"/>	9	b	655 4 aulmekirjandus 0https://ems.elnet.ee/id/EMS027390						2	
	10	a	550 wg aulmekirjandus 0https://ems.elnet.ee/id/EMS027390						3	
	11	a	550 wh aulmekirjandus 0https://ems.elnet.ee/id/EMS027390						2	


- Spacecraft in fact and fiction // cHarry Harrison, Malcolm Edwards.
- Эдем ;|bСолярис : [романы] ; Рассказы //cСтанислав Лем
- The fall of the House of Usher|c/Edgar Allan Poe ; with an afterword by R. P. Blackmur
- From utopia to apocalypse :|bscience fiction and the politics of catastrophe //cPeter Y. Paik

Läheb aina hullemaks!

Index		SUBJECT (d)		kirikulaulud
#		SUBJECT		# ENTRIES
1	a	455 akirikulaulud		1
2	a	555 akirikulaulud		6
<input checked="" type="checkbox"/>	3	b	650 4 akirikulaulud	823
<input checked="" type="checkbox"/>	4	b	692 akirikulaulud	83
<input checked="" type="checkbox"/>	5	b	697 9 akirikulaulud	195
6	a	155 akirikulaulud 0https://ems.elnet.ee/id/EMS020551		1
<input checked="" type="checkbox"/>	7	b	650 4 akirikulaulud 0https://ems.elnet.ee/id/EMS020551	25
<input checked="" type="checkbox"/>	8	b	655 4 akirikulaulud 0https://ems.elnet.ee/id/EMS020551	1486
9	a	555 wg akirikulaulud		6
10	a	555 wh akirikulaulud		2

Index		SUBJECT (d)		palved
#		SUBJECT		# ENTRIES
<input checked="" type="checkbox"/>	1	b	655 4 apalved	12
2	a	450 apalved		1
<input checked="" type="checkbox"/>	3	b	655 4 apalved	649
4	a	150 apalved 0https://ems.elnet.ee/id/EMS022798		1
5	a	550 apalved 0https://ems.elnet.ee/id/EMS022798		7
<input checked="" type="checkbox"/>	6	b	650 4 apalved 0https://ems.elnet.ee/id/EMS022798	513
<input checked="" type="checkbox"/>	7	b	655 4 apalved 0https://ems.elnet.ee/id/EMS022798	24
8	a	550 wg apalved 0https://ems.elnet.ee/id/EMS022798		3
9	a	550 wh apalved 0https://ems.elnet.ee/id/EMS022798		1

Näide EMS-i toimetuse töölaualt

- õpikud vormimärksõna
- kõrgkooliõpikud vormimärksõna
- kutsekooliõpikud vormimärksõna
- kooliõpikud **teema** märksõna  **vormi** märksõna

Teeme asjad lihtsamaks

Hetke olukord: 1 norm(i kirje)

- 155 kirikulaulud|0<https://ems.elnet.ee/id/EMS020551>

Saabuv olukord: 2 norm(i kirjet)

- 150 kirikulaulud|0<https://ems.elnet.ee/id/EMS020551>
- 155 kirikulaulud (vormimärksõna)|0<https://ems.elnet.ee/id/EMS020551>

Mis muutub, mis jääb samaks

- Sierrades saab valida 2 täheduselt lähedase märksõna vahel
- Vormimärksõnade eristamiseks lisandub neile sulgudesse täpsustus (vormimärksõna)
- Mõlema märksõna URI-d viitavad samale EMS-i mõistele
- Muutus puudutab vaid vormimärksõnu
- Teema -, aja-, kohamärksõnade puhul ei muutu midagi

Edasised praktilised plaanid

- Üks ühine EMS-i väliste märksõnade väli
- Ülearuste punktide kaotamine ajamärksõnadest

d	650	4	templid (pühamud) https://ems.elnet.ee/id/EMS019678
d	651	4	Vana-Kreeka https://ems.elnet.ee/id/EMS011079
d	648	4	<u>13. sajand eKr</u>
d	655	4	DVD-d
d	655	4	ajaloolised filmid
d	655	4	dokumentaalfilmid

Veel tähelepanekuid

- Ärge kasutage välja 653! Paar nädalat tagasi oli selliseid kirjeid 59! Väli 653 sisu kantakse väljale 648.

Vaatame artikli märksõnastamise tulemusi!

Milliseid märksõnu kasutaksite Kukepea-artikli märksõnastamiseks?

Wordcloud Poll 24 responses 10 participants



Suurte keelemudelite eelsed lahendused

- Automaatse märksõnastamise kratiid
 - <https://marta.nlib.ee>
- Soomlaste Annif – automaatse märksõnastamise tööriist
 - <https://annif.org>

MARTA: Artiklite automaatse märksõnastamise krati prototüüp

Artikli valimine

Tekst Fail URL ⓘ

Sisesta või kleebi tekst...

Küsimuste või probleemide korral kontakteeru Eesti Rahvusraamatukoguga: data@nlib.ee

annif Tool for automated subject indexing and classification

Choose a controlled subject vocabulary and train Annif on already indexed documents – it can then suggest subjects for new documents!

HOW TO USE ANNIF

- Choose subject vocabulary
- Prepare a corpus from training data
- Load the vocabulary and train a model
- Suggest subjects for new documents

Kuidas märksõnastas Annif?

INPUT TEXT

Kukepea võitjal läksid püksid kasvama
Aprillis Lääne-Virumaa metsades toimunud suurte masinate offroadvõistlusel Kukepea põristasid sõidukid, mis pole loodud parketi peal veeremiseks.
Yabaklassis võidutses meeskond Samru. ATV-arvestuses võitis tiim Andromeda Racing. TR-1 parimaks osutus Lendav Draakon. TR-3 parima meeskonna Boatworld moodustasid Haljalast pärit Erkki Salak ning pealinnamees Hardo Mere, kes rallisid maastikul omavalmistatud masinaga H6. Nimi tähendab, et tegemist on Hardo ehitatud kuuenda sõidukiga. Kogu võistluse raskeim osa oli Erkki sõnul autos vastu pidamine. "Mul tekkisid füüsiliselt tõsised probleemid. Suurte kiiruste juures pidi kõvasti kinni hoidma, et autost välja ei lendaks. Käe- ja kaelalihased andsid pärast tunda ja käed on siiaaani rakkus, nahk on maas," meenutas vintsimees kümnetunnist võistluspäeva. Kõige keerulisem punkt sattus meestel kohe esimeseks. Kui vintsimees uppuvast sõidukist välja hüppas, sattus ta rinnuni vette. "Noh, mis vesi. See oli üks rinnuni jama. Peal oli 30 sentimeetrit läbikasvanud kamaralaadset toodet. Jalgupidi seal peal käia ei saanud. Ja selle all oli vesi. Väga vedel mant. Roomata ei olnud võimalik, kõndida ei olnud võimalik, ujuda ei olnud võimalik," kõneles Erkki. "Ma ei saanud päris hästi aru, et kui kuskilt ei ole ennast edasi lükata, siis kuidas ma liigun. Süda läks pahaks ja lihtsalt ei jaksanud ennast tõmmata," kirjeldas vintsimees keerulist trossivedamise katsumust. Erkki pani võistlusteks jalga uhiuued jooksupüksid, mis esimeses punktis kurijade ohakate meelevalda sattusid ja pärast üsna siili moodi välja nägid. Pesukaussi likku visatud riided olid järgmiseks päevaks haliendama läinud ning ühes uute rohetavate retuusidega rändasid prügikasti ka vanaema kootud villased sokid. Haljala mees oli Kukepeal võistlustules esimest korda. Järgmiselgi aastal oleks ta hea meelega võistleja rollis. "Palju

PROJECT (VOCABULARY AND LANGUAGE)

YSO NN ensemble Finnish

MAX # OF SUGGESTIONS

10 15 20

Get suggestions →

annif

SUGGESTED SUBJECTS

Länsi-Virumaa	Lääne-Virumaa
Viro	Eesti
historia	kultuuripärand
kultuuriperintö	nimed
nimet	vesi
vesi	reovesi
jätevesi	piim
kokemukset	võistlused
maito (lehmänmaito)	
urheilu	

Kuidas märksõnastas Annif?

INPUT TEXT

Kukepea voittajalla meni housut kasvamaan
Huhtikuussa Länsi-Virumaan metsissä järjestetyssä suurten koneiden offroad-kilpailussa Kukepea pörsähtivät ajoneuvot, jotka eivät ole luotu parketilla vierimiseen.
Vapaassa luokassa voitti joukkue Samru. ATV-luokassa voitti tiimi Andromeda Racing. TR-1 parhaaksi osoittautui Lentävä Lohikäärme. TR-3 parhaan joukkueen Boatworld muodostivat Haljalasta kotoisin olevat Erkki Salak ja pääkaupunkilainen Hardo Mere, jotka rallivat maastossa itsevalmistetulla koneella H6. Nimi tarkoittaa, että kyseessä on Hardon rakentama kuudes ajoneuvo. Koko kilpailun vaikein osa oli Erkin mukaan autossa pysyminen. "Minulle tuli fyysisesti vakavia ongelmia. Suurilla nopeuksilla piti pitää kovasti kiinni, ettei autosta lentäisi ulos. Käsi- ja kaulalihakset antoivat jälkeinpäin tuntua ja kädet ovat vieläkin rakkuloilla, iho on maassa", muisteli vinssimies kymmentuntista kilpailupäivää. Vaikein piste osui miehille heti ensimmäiseksi. Kun vinssimies hyppäsi uppoavasta ajoneuvosta ulos, hän joutui rintaan asti veteen. "No, mikä vesi. Se oli yksi rintaan asti jama. Päällä oli 30 senttimetriä läpikasvanutta kamaramaista tuotetta. Jalkapohjalla siellä päällä ei voinut käydä. Ja sen alla oli vettä. Erittäin nestemäinen mantteli. Ryömiä ei ollut mahdollista, kävellä ei ollut mahdollista, uida ei ollut mahdollista", kertoi Erkki. "En saanut ihan hyvin selvää, että jos mistään ei ole itseään työntää eteenpäin, niin miten minä liikun. Sydän meni pahaksi ja en yksinkertaisesti jaksanut vetää itseäni", kuvasi vinssimies vaikeaa köysivedon koetusta. Erkki laittoi kilpailuiksi jalkaansa upouudet juoksu housut, jotka ensimmäisessä pisteessä joutuivat ilkeiden ohdakkeiden armoille ja näyttivät jälkeinpäin melko siilimäisiltä. Pesuvätiin likoamaan heitetty vaatteet olivat seuraavaksi päiväksi vihertyneet ja uusien vihreiden retuusien kanssa päätyivät roskakoriin myös isoäidin neulomat

PROJECT (VOCABULARY AND LANGUAGE)

YSO NN ensemble Finnish

MAX # OF SUGGESTIONS

10 15 20

Get suggestions →

annif

SUGGESTED SUBJECTS

<input type="checkbox"/> kilpailut	võistlused
<input type="checkbox"/> urheilijat	sportlased
<input type="checkbox"/> ajoneuvot	sõidukid
<input type="checkbox"/> autot	autod
<input type="checkbox"/> kilpaurheilu	võistlusport
<input type="checkbox"/> kokemukset	kogemused
<input type="checkbox"/> historia	spordivõistlused
<input type="checkbox"/> urheilukilpailut	parkett
<input type="checkbox"/> parketti	parkett
<input type="checkbox"/> metsät	metsad

Kuidas märksõnastas Marta?

MARTA: Artiklite automaatse märksõnastamise krati prototüüp

võistlused ✓ esimehed ✓ vulkaanid ✓ meeskonnad ✓

meeskonnad ✓ orienteerujad ✓ meistrivõistlused ✓ autoralli ✓ kiirus ✓ jooksjad ✓ jooksmine ✓

Minimaalne tõenäosus

liiklusvahendid ✓ võistlused ✓ kukepea erkki Haljala ✓ vintsimees meeskonnad ✓ masinad ✓ hardo boatworl vabaklass veeremine parkett ✓

Hybrid Tagger 6/15

Minimaalne tõenäosus

võistlused ✓ autod ✓ sõjamängud ✓ liiklusohutus ✓ vargused ✓ üliõpilaspäevad ✓

Isikunimed 2/4

Minimaalne sõnakorduste arv

Erkki Salak Hardo Mere

Kohamärksõnad 13/13

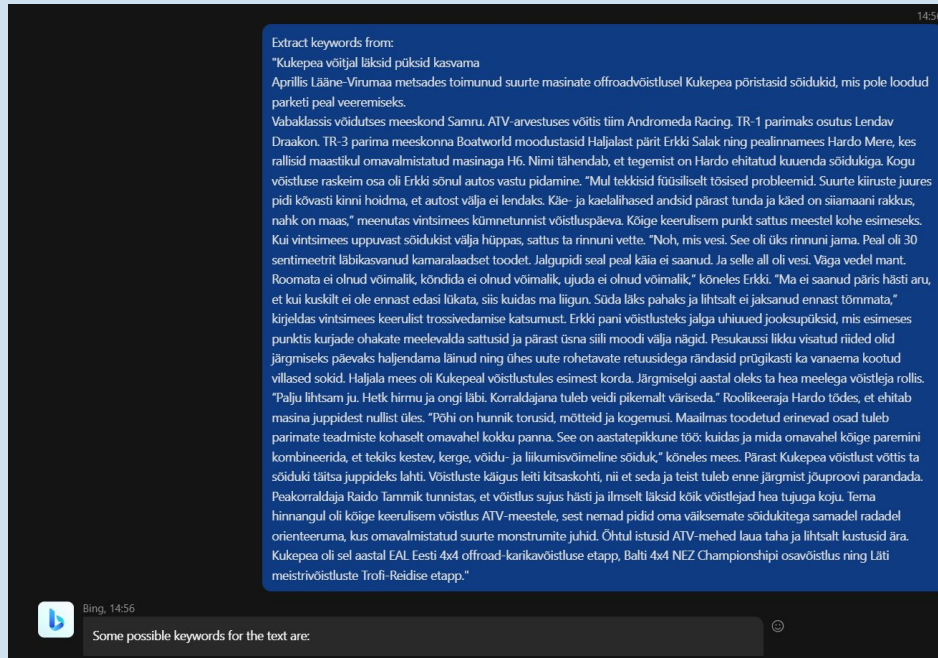
Minimaalne sõnakorduste arv

Haljala ✓ Trofi-Reidis ATV-arvestus Läti ✓ Eesti ✓ Lääne-Virumaa ✓ Balti Baltimaad ✓ Euroopa ✓ Läänemere maad ✓ Põhja-Euroopa ✓ maakonnad ✓

Vigased pruudid
ehk mida suudab teha taip



Kuidas märksõnastas Bing (Skype)?



14:56

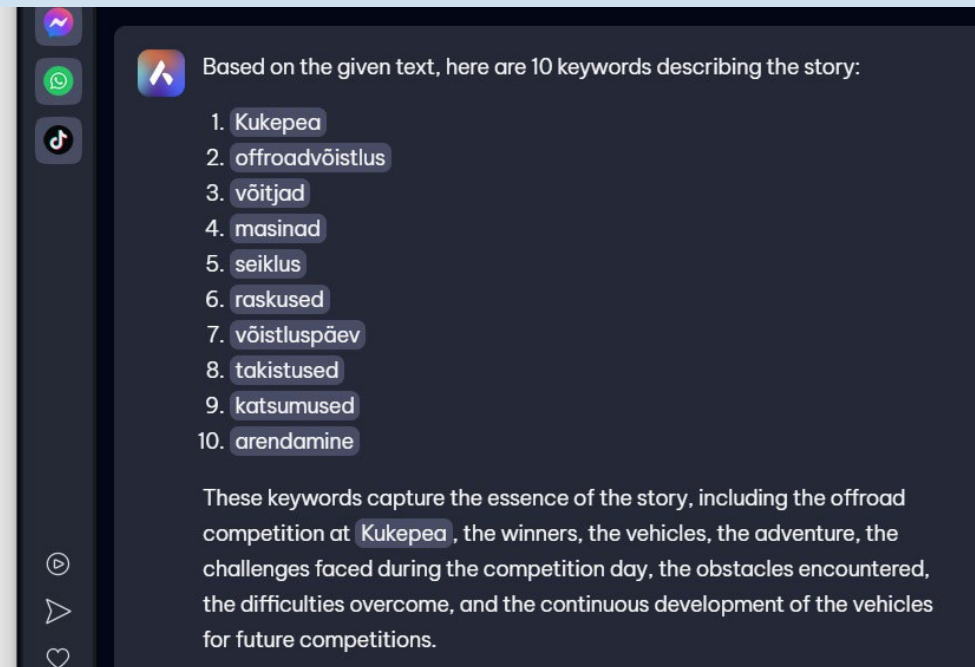
Extract keywords from:
"Kukepea võitjad läksid püksid kasvama
Aprillis Lääne-Virumaa metsades toimunud suurte masinate offroadvõistlusel Kukepea põristasid sõidukid, mis pole loodud parketi peal veeremiseks.
Vabaklassis võidutses meeskond Samru. ATV-arvestuses võitis tiim Andromeda Racing. TR-1 parimaks osutus Lendav Draakon. TR-3 parima meeskonna Boatworld moodustasid Hajjalast pärit Erkki Salak ning pealinnamees Hardo Mere, kes rallisid maastikul omavalmistatud masinaga H6. Niimi tähendab, et tegemist on Hardo ehitatud kuenda sõidukiga. Kogu võistluse raskeim osa oli Erkki sõnul autos vastu pidamine. "Mul tekkisid füüsiliselt tõsised probleemid. Suurte kiiruste juures pidi kõvasti kinni hoidma, et autost välja ei lendaks. Käe- ja kaelalihased andsid pärast tunda ja käed on siiani rakkus, nahk on maas," meenutas vintsimees kümnetunnist võistluspäeva. Kõige keerulisem punkt sattus meestel kohe esimeseks. Kui vintsimees uppuvast sõidukist välja hüppas, sattus ta rinnuni vette. "Noh, mis vesi. See oli üks rinnuni jama. Peal oli 30 sentimeetrit libikasvanud kamaralaadset toodet. Jalgu pidi seal peal käia ei saanud. Ja selle all oli vesi. Väga vedel mant. Roomata ei olnud võimalik, kõndida ei olnud võimalik, ujuda ei olnud võimalik," kõneles Erkki. "Ma ei saanud päris hästi aru, et kui kuskilt ei ole ennast edasi lükata, siis kuidas ma liigun. Sõida läks pahaks ja lihtsalt ei jaksanud ennast tõmmata," kirjeldas vintsimees keerulist trossivedamise katsumust. Erkki pani võistlusteks jalga uhiuued jooksupüksid, mis esimeses punktis kurjade ohakate meelevadla sattusid ja pärast üsna siili moodi välja nägid. Pesukaussi liikku visatud riided olid järgmiseks päevaks haljendama läinud ning ühes uute rohetavate retuusidega rändasid prügikasti ka vanaema kootud villased sokid. Hajjala mees oli Kukepeal võistlustules esimest korda. Järgmiselgi aastal oleks ta hea meelega võistleja rollis. "Palju lihtsam ju. Hetk hirmu ja ongi läbi. Korraldajana tuleb veidi pikemalt väriseda." Roolikeeraja Hardo tõdes, et ehitab masina juppidest nullist üles. "Põhi on hunnik torusid, mõtteid ja kogemusi. Maailmas toodetud erinevad osad tuleb parimate teadmiste kohaselt omavahel kokku panna. See on aastatepikkune töö: kuidas ja mida omavahel kõige paremini kombineerida, et tekiks kestev, kerge, võidu- ja liikumisvõimeline sõiduk," kõneles mees. Pärast Kukepea võistlust võttis ta sõiduki täitsa juppideks lahti. Võistluste käigus leiti kitsaskohti, nii et seda ja teist tuleb enne järgmist jõuproovi parandada. Peakorraldaja Raido Tammik tunnistas, et võistlus sujus hästi ja ilmselt läksid kõik võistlejad hea tujuga koju. Tema hinnangul oli kõige keerulisem võistlus ATV-meestele, sest nemad pidid oma väiksemate sõidukitega samadel radadel orienteeruma, kus omavalmistatud suurte monstrumite juhid. Öhtul istusid ATV-mehed laua taha ja lihtsalt kustusid ära. Kukepea oli sel aastal EAL Eesti 4x4 offroad-karikavõistluse etapp, Balti 4x4 NEZ Championshipi osavõistlus ning Läti meistrivõistluse Trofi-Reidise etapp."

Bing, 14:56

Some possible keywords for the text are:

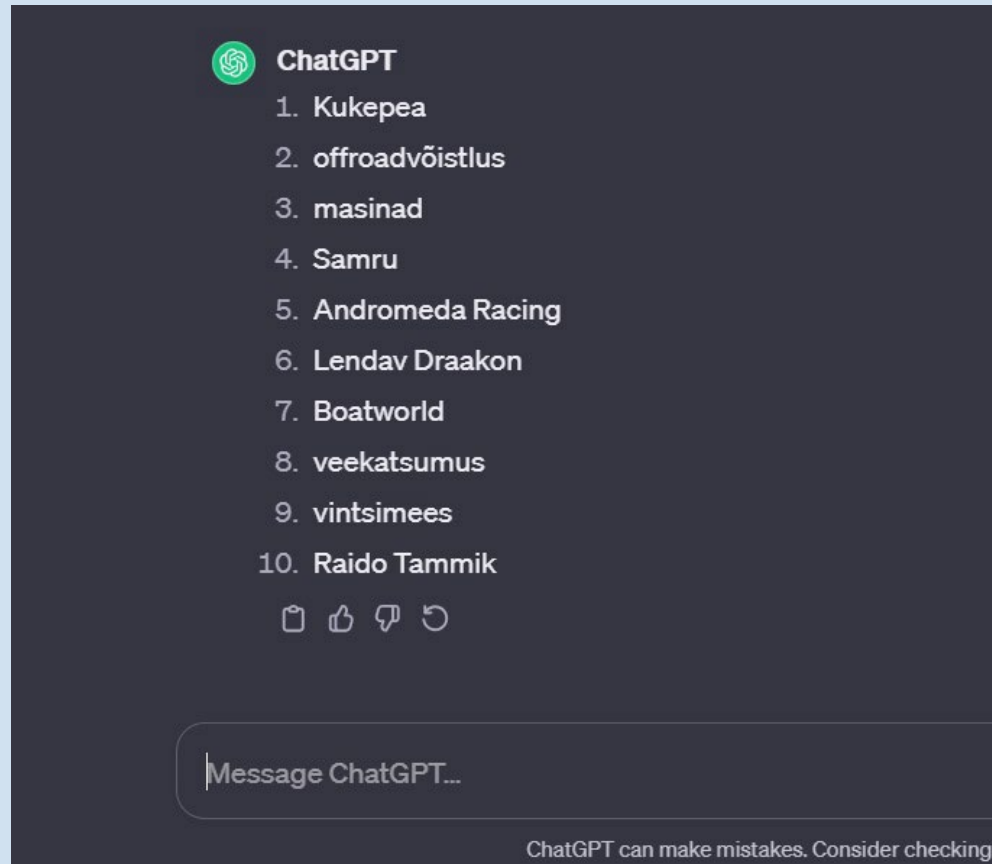
- Kukepea
- offroad -võistlus
- omavalmistatud masin
- Erkki Salak
- Hardo Mere
- vintsimees
- ATV -arvestus
- Lendav Draakon
- Samru
- Boatworld

Kuidas märksõnastada Aria (Opera)?



- Kukepea
- offroadvõistlus
- võitjad
- masinad
- seiklus
- raskused
- võistluspäev
- takistused
- katsumused
- arendamine

Kuidas märksõnastas ChatGPT?



- Kukepea
- offroadvõistlus
- masinad
- Samru
- Andromeda Racing
- Lendav Draakon
- Boatworld
- veekatsumus
- vintsimees
- Raido Tammik

Märksõnastamise võrdlus

Annif:

- võistlused
- sportlased
- sõidukid
- autod
- võistlussport
- kogemused
- spordivõistlused
- parkett
- metsad

Marta:

- võistlused
- meeskonnad
- meistrivõistlused
- autoralli
- kiirus
- jooksjad
- autod
- Haljala
- Lääne-Virumaa

Bing:

- Kukepea
- offroad-võistlus
- omavalmistatud masin
- Erkki Salak
- Hardo Mere
- vintsimees
- ATV-arvestus
- Lendav Draakon
- Samru
- Boatworld

Aria:

- Kukepea
- offroadvõistlus
- võitjad
- masinad
- seiklus
- raskused
- võistluspäev
- takistused
- katsumused
- arendamine

ChatGPT 3.5:

- Kukepea
- offroadvõistlus
- masinad
- Samru
- Andromeda Racing
- Lendav Draakon
- Boatworld
- veekatsumus
- vintsimees
- Raido Tammik

Lõpp paistab

15:43

miks on hundi laut külm?



Bing, 15:43

Hundi laut ei ole alati külm, kuid see võib olla külm siis, kui hunt on sellest lahkunud ja keegi teine on sinna sisse läinud ning ukсед on jäänud lahti. **1** **2** Hunt on laia saagispektriga kiskja, toitudes kõige kättesaadavamatest toiduobjektidest antud alal ja ajal. Hundi toidusedel võib sisaldada nii suuri saakloomi nagu põder, hirv, metskits ja metssiga kui ka pisinärlisi, selgrootuid, raibet ja taimset toitu. **2**

Kas sa soovid teada midagi veel huntide kohta?

Learn more:

[1. looduskalender.ee](https://looduskalender.ee)

[2. looduskalender.ee](https://looduskalender.ee)

Lõpp paistab



Hundi laut võib olla külm mitmel põhjusel. Siin on mõned võimalikud põhjused:

1. Halb isolatsioon: Kui **hundi laut** on halvasti ehitatud või sellel puudub piisav isolatsioon, võib see põhjustada külma sissetungi. Külmaõhk saab läbi pragude või halvasti suletud akende ja uste sisse tulla ning hoida lauta jaheda.
2. Puudulik küttesüsteem: Kui **hundi lautil** puudub korralik küttesüsteem või soojustus, võib see olla põhjuseks, miks see on külm. Kui laudis puudub soojusallikas, ei suuda see piisavalt soojust hoida ja seetõttu võib see külmuda.
3. Ilmastikuolud: Väljaspool olevad temperatuurid ja ilmastikuolud võivad mõjutada ka **hundi lauda** temperatuuri. Kui väljas on väga külmad talveilmad või tuuline ilm, võib see mõjutada ka lauda sisetemperatuuri.
4. Ventilatsiooniprobleemid: Halb ventilatsioon võib samuti põhjustada **hundi lauda** külmumist. Kui õhk ei liigu korralikult või kui ventilatsioonisüsteem on defektne, võib see viia külmaõhu kogunemiseni laudas.

On oluline parandada **hundi lauda** külmumise probleemid, et tagada loomadele sobivad elamistingimused. Parandused võivad hõlmata isolatsiooni parandamist, küttesüsteemi paigaldamist või ventilatsiooni parandamist. Kui probleem püsib, võib olla vaja konsulteerida ekspertidega, näiteks loomakasvatusteadlastega, et leida lahendusi **hundi lauda** soojendamiseks.

Lõpp paistab

14:29

Mis juhtus Kalevipoja jalgadega



Bing, 14:30

Kalevipoja jalad said rängalt kannatada, kui ta püüdis Soome silda ehitada. Ta lõi endale haamriga varbad katki ja veri jooksis jalast välja. Ta sidus oma jalad kinni, kuid ei saanud haavu korralikult ravida. Ta läks tagasi Eestisse, kus ta võitles oma viimase lahingu Põhja kuninga vastu. Lahingu ajal murdusid tema jalad ja ta kukkus maha. Ta suri oma murtud jalgade tõttu, kuid enne seda jõudis ta oma mõõga Tammsaare rabasse visata. **1** **3**



Aitäh!

- Eesti
Rahvusraamatukogu
- Kogude Arenduse
Osakond
- Meelis Somelar
- meelis.somelar@rara.ee
29.11.2023