



Euroopa Liit  
Euroopa  
Regionaalarengu Fond



Eesti  
tuleviku heaks

# Automaatse märksõnastamise KRATT: prototüübi kasutusjuhend

# Sisukord

<b>Sisukord</b>	<b>1</b>
<b>1. Märksõnastamise töövoog</b>	<b>2</b>
<b>2. Parameetrid ja funktsionaalsused</b>	<b>4</b>
2.1. Lehekülgede arv	4
2.2. Märksõnade lävend	5
2.3. Keeled	6
2.4. Märksõnade eemaldamine	7
2.5. Märksõnaplokkide avamine ja sulgemine	8
2.6. MARC kirje kopeerimine	10
<b>3. Veateated</b>	<b>11</b>
<b>4. Näide kasutuslooga</b>	<b>14</b>

# 1. Märksõnastamise töövoog

Ühe teaviku automaاتمärksõnastajaga KRATT märksõnastamise töövoog on järgnev:

1. Vali teaviku üleslaadimise viis.
  - URL-ilt laadimiseks vali ülemiselt ribalt “Teavik läbi URL-i” ning sisesta lahtrisse “Teaviku URL” teavikule vastav URL (vt. Joonis 1.1).
  - Faili üleslaadimiseks arvutist vali ülemiselt ribalt “Teavik läbi faili”. Seejärel kliki lahtrile “Teavik kasutaja arvutist” ning vali arvutist soovitud fail (vt. Joonis 1.2).



**Automaatmärksõnastaja**

Teavik läbi URL-i  Teavik läbi faili

Teaviku URL \*  
http://

Lehekülgede arv \*  
10

Märksõnasta

**Joonis 1.1.** Teaviku üleslaadimine URL-i kaudu.



**Automaatmärksõnastaja**

Teavik läbi URL-i  Teavik läbi faili

Teavik kasutaja arvutist \*  
Otsi

Lehekülgede arv \*  
10

Märksõnasta

**Joonis 1.2.** Teaviku üleslaadimine kasutaja arvutist.

2. Muuda vajadusel märksõnastajale edastatavate lehekülgede arvu. Selleks kliki lahtrile “Lehekülgede arv” ning sisesta soovitud lehekülgede arv. Loe lisa peatükist [2.1. Lehekülgede arv](#).
3. Vajuta nupule “Märksõnasta”.

4. Pärast “Märksõnasta” nupule vajutamist kuvatakse lehe paremal pool info märksõnastamise progressist (vt. Joonis 2).



Joonis 2. Info märksõnastamise progressi kohta.

5. Kui märksõnastaja protsessis tekib viga, kuvatakse kasutajale vastav veeteade ning töö lõpetatakse. Loe lisa peatükist [3. Veeteated](#).
6. Kui märksõnastaja lõpetab töö edukalt, kuvatakse kasutajale lehe vasakpoolses ääres sektordiagramm tuvastatud keeltest (loe lisa peatükist [2.3. Keeled](#)) ning lehe parempoolses ääres leitud märksõnad vaikelävendi 0.4 juures (vt. Joonis 3.1) või tekst “Määratud lävendi ületavad märksõnad puuduvad” (vt. Joonis 3.2).

**Automaatmärksõnastaja**

Teavik läbi URL-i  Teavik läbi faili

Teavik kasutaja arvustist \*  
401567

Lehekülgede arv \*  
10

Märksõnasta

Leitud keeled

inglise

eesti

Leitud märksõnad

Märksõnade lävend:

Märksõnu kokku 5 Selekteeritud märksõnu 5 Selekteerimata märksõnu 0

**Teemamärksõna**

- ühistud
- ühistunudus

**Kohamärksõna**

- Eesti

**Vormimärksõna**

- e-raamatud
- võrguväljaanded

Logo: Euroopa Liit, Euroopa Regionaalarengu Fond; Eesti tuleviku heaks

Joonis 3.1. Märksõnastaja väljund.



**Joonis 3.2.** Märksõnastaja väljund, kui määratud lävendi ületavad märksõnad puuduvad.

7. Soovi korral on kasutajal võimalik lävendit tõsta (kuvatakse vähem, aga potentsiaalselt täpsemaid märksõnu) või langetada (kuvatakse rohkem, aga potentsiaalselt vähem täpseid märksõnu). Loe lisa peatükist [2.2. Märksõnade lävend](#).
8. Soovi korral on kasutajal võimalik eemaldada linnuke sobimatute märksõnad eest - selekteerimata märksõnu tulemusse ei kaasata. Loe lisa peatükist [2.4. Märksõnade eemaldamine](#).
9. “Kopeeri” nupule vajutades kopeeritakse puhvrisse kõikide antud lävendi juures valitud märksõnade MARC kirje. Loe lisa peatükist [2.6. MARC kirje kopeerimine](#).

## 2. Parameetrid ja funktsionaalsused

Käesoleva peatüki eesmärgiks on anda ülevaade automaatmärksõnastaja parameetrite, väljundi ning erinevate funktsionaalsuste kohta. Automaatmärksõnastaja parameetrite hulka kuuluvad lehekülgede arv ning märksõnade lävend.

### 2.1. Lehekülgede arv

Parameetri nimetus	Parameetri funktsioon	Vaikeväärtus
Lehekülgede arv	Väljendab, mitu juhuslikult valitud lehekülge teavikust märksõnastajale edastatakse.	10

## Kirjeldus

“Lehekülgede arv” väljendab, mitu juhuslikult valitud lehekülge teavikust märksõnastajale edastatakse. See tähendab, et kui teaviku tegelik lehekülgede arv on 500, märksõnastajale edastatakse aga 10 lehekülge, leitakse märksõnad üksnes 2% teose sisu põhjal. Kuna märksõnastamiseks kasutatavad leheküljed valitakse juhuslikult, ei pruugi lisaks kõik neist olla representatiivsed teose sisule. Seega peaks kasutaja märksõnastajale saadetavate lehekülgede arvu vajadusel suurendama - eriti oluline on see näiteks esseekogumike põhjal, kus teaviku alamosade sisud võivad olla täiesti erinevad.

Siiski tuleb tähele panna, et koos märksõnastajale edastatavate lehekülgede arvuga kasvab ka teaviku märksõnastamise aeg. Vaikeväärtus 10 on ühtlase sisulise jaotusega teoste puhul üldjuhul piisav arv.

Vaata lisaks näitlikku ülevaadet lehekülgede arvu ja märksõnade lävendi seosest peatükist [4. Näide kasutuslooga](#).

## 2.2. Märksõnade lävend

Parameetri nimetus	Parameetri funktsioon	Vaikeväärtus
Märksõnade lävend	Väljendab, kui suurest osast märksõnastajale edastatud lehekülgedest peab olema märksõna tuvastatud.	0.4

## Kirjeldus

Märksõnade lävend väljendab, kui paljude märksõnastamisele saadetud lehekülgedega peab iga märksõna minimaalselt seotud olema. Näiteks kui märksõnastamiseks edastatakse 10 lehekülge, siis väljendab lävend 0.4, et iga kuvatud märksõna esineb vähemalt neljas ( $0.4 \cdot 10$ ) märksõnastamisele saadetud leheküljes. Kui märksõnastamiseks edastatakse aga 500 lehekülge, siis väljendab lävend 0.4, et iga kuvatud märksõna esineb vähemalt ( $0.4 \cdot 500$ ) 200 leheküljes jne. Seega tähendab lävendi väärtus 1, et kuvatavad märksõnad esinevad igas märksõnastamisele edastatud leheküljes (praktiliselt üsna haruldane) ning lävend 0, et iga kuvatud märksõna esineb vähemalt ühes märksõnastamisele edastatud leheküljes (sõltumata märksõnastajale edastatud lehekülgede arvust).

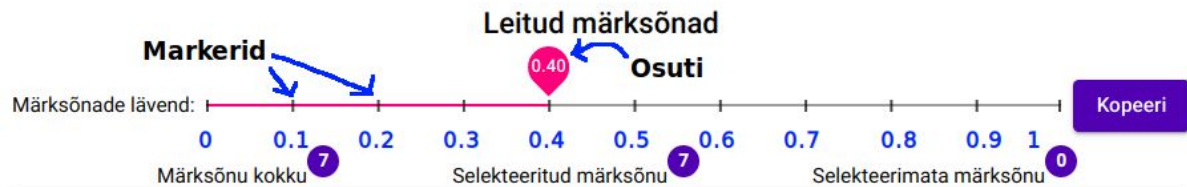
Seega ei pruugi lävendi (eriti väikeste intervallidega) muutmine tuua kaasa muutusi kuvatud märksõnade hulgas. Tabelis 1 näeme, et lävendi 0.8 ületab ainult märksõna “Eesti”. Kui lävend langetada aga 0.7 peale, siis on “Eesti” endiselt ainus märksõna, mis selle ületab, kuna järgmiste märksõnade suhteline sagedus on sellest madalam (0.6).

## Lävendi muutmine

Lävendi muutmiseks on kaks võimalust:

1. Klõkkida hiirega osuti peal ning nihutada seda soovitud suunas soovitud väärtusele.
2. Klõkkida lävendi ribal soovitud lävendi peale, mille järel osuti ise sinna nihkub.

Lävendiribal olevad markerid tähistavad vasakult paremale liikudes väärtuseid 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 ning 1.



Joonis 4. Lävendiriba.

Vaata lisaks näitlikku ülevaadet lehekülgede arvu ja märksõnade lävendi seosest peatükist [4. Näide kasutuslooga](#).

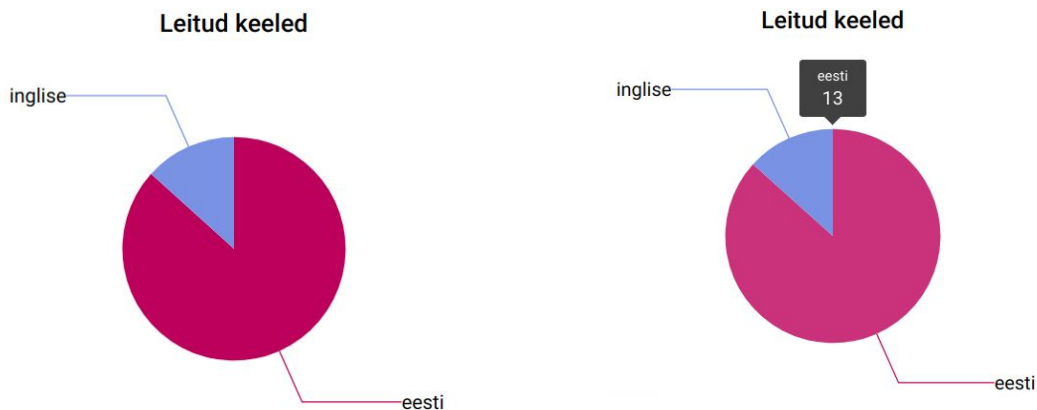
## 2.3. Keeled

Keeled tuvastatakse **ainult märksõnastamiseks edastatud lehekülgedest**. See tähendab, et sektordiagramm ei pruugi anda adekvaatset ülevaadet tervikteose keelte jaotumisest. Näiteks juhul kui teose pikkuseks on 500 lehekülge, märksõnastajale edastatakse aga 10, siis kuvatakse sektordiagrammil üksnes märksõnastamisele saadetud 10 lehekülje keeleline jaotus. Kui kasutaja soovib näha tervikteose keelelist jaotust, tuleks tal märksõnastajale edastatavate lehekülgede arvuks sisestada teose täielik pikkus (NB! see võib olla ajakulukas protsess).

Keelelise jagunemise sektordiagrammi eesmärk on anda ülevaade märksõnastatud lehekülgede keelelisest jagunemisest, kuna märksõnastaja on treenitud valdavalt eestikeelsete teoste peal. Kui sektordiagrammil kuvatud keelte jaotuses on märgatav osa aga mõnel teisel keelel, ei pruugi märksõnastaja pakutud märksõnad olla kuigi täpsed.

Märksõnastajale edastatud lehekülgedelt tuvastatud iga keele esinemise täpse arvu saab teada sektordiagrammil hiirega vastava keele sektorile liikudes (vt. Joonis 5.1 ja Joonis 5.2).

Tuvastatud keelte kuvamisel eristatakse järgnevaid keeli: **eesti**, **inglise**, **vene**. Kui tuvastatud keel ei kuulu nimetatud kolme hulka, kuvatakse selle sildiks “**toetamata**”.

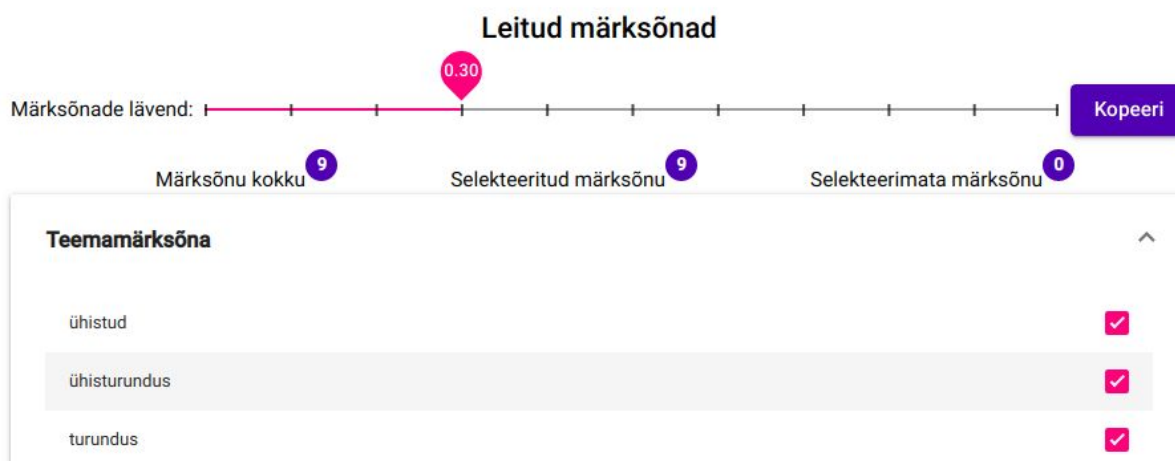


**Joonis 5.1.** Tuvastatud keelte jaotus 15 lehekülje märksõnastajale edastamisel.

**Joonis 5.2.** Sektordiagrammil sektoritele hiirega peale liikudes on näha vastavas keeles lehekülgede täpne arv. Antud juhul näeme, et 15 märksõnastamisele saadetud leheküljest oli eestikeelseid kokku 13.

## 2.4. Märksõnade eemaldamine

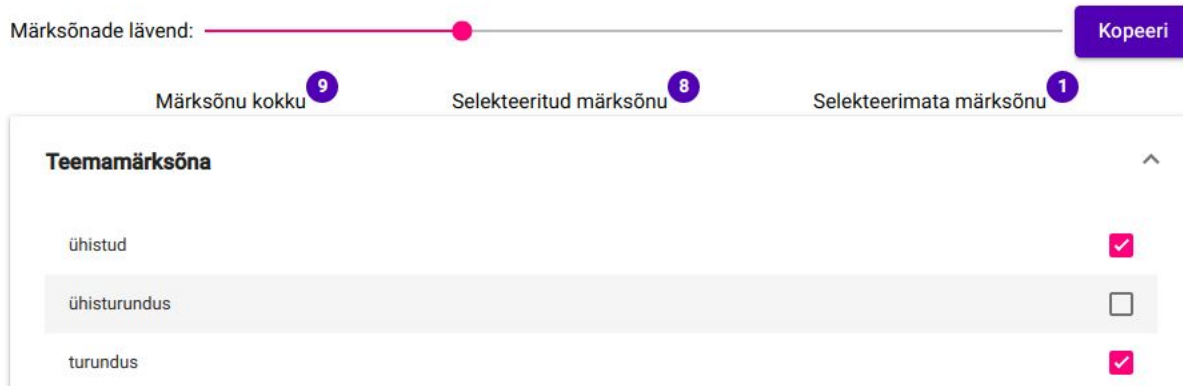
Kõik kuvatud märksõnad on vaikimisi selekteeritud. Kui kasutajale tundub, et kuvatud märksõnade hulgas on valede märksõnade osakaal väga suur, on soovitatav esialgu lävendiribal lävendit suurendada (vt. peatükk [2.2. Märksõnade lävend](#)). Kasutajal on võimalus märksõnade valikut piirata aga ka märksõna kõrval asuvas kastis oleva linnukese eemaldamisel (vt. Joonis 6.1 ja Joonis 6.2). Märksõnade selekteeritus säilib ka lävendi muutmisel - see tähendab, et kui lävendi 0.3 juures märksõna X kõrvalt linnuke eemaldada ning seejärel lävend 0.1 peale seada, jääb märksõna X endiselt selekteerimata (vt. Joonis 6.1 - Joonis 6.3). Märksõnade eemaldamisel kajastub tulemus ka märksõnade kohal asuvates indikaatorites, mis väljendavad selekteerimata ja selekteeritud märksõnade arve.



**Joonis 6.1.** Märksõna "ühistorundus" on selekteeritud lävendi 0.3 juures.



## Leitud märksõnad



**Joonis 6.2.** Märksõna “ühisturundus” kastist eemaldatakse lävendi 0.3 juures linnuke. Selekteerimata märksõnade arv tõuseb ühe võrra ning selekteeritud märksõnade arv kahaneb ühe võrra.

## Leitud märksõnad



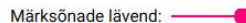
**Joonis 6.3.** Lävendi langetamisel 0.1 peale säilib märksõna “ühisturundus” selekteerituse staatus.

## 2.5. Märksõnaplokkide avamine ja sulgemine

Kõik märksõnad on nende sisu alusel grupeeritud nelja ploki: **Teemamärksõna**, **Kohamärksõna**, **Ajamärksõna** ning **Vormimärksõna**. Kui mõnest nimetatud märksõnaplokist (antud lävendi juures või üldse) ühtegi märksõna ei tuvastatud, ei kuvata ka vastavat ploki.

Iga kuvatud ploki on võimalik avada ja sulgeda klikkides ploki kõrval asuvale noolele (vt. Joonis 7.1. ja Joonis 7.2.). Ploki sulgemine on mõeldud üksnes mugavamaks navigeerimiseks ning ei mõjuta märksõnade selekteeritust.

## Leitud märksõnad

Märksõnade lävend:  Kopeeri


Märksõnu kokku **67**    Selekteeritud märksõnu **67**    Selekteerimata märksõnu **0**

**Teemamärksõna** ↗ ⌵

ühistud	<input checked="" type="checkbox"/>
ühisturundus	<input checked="" type="checkbox"/>
turundus	<input checked="" type="checkbox"/>
ühistegevus	<input checked="" type="checkbox"/>
toiduturg	<input checked="" type="checkbox"/>
turustus	<input checked="" type="checkbox"/>
eksporditurundus	<input checked="" type="checkbox"/>
turunduskommunikatsioon	<input checked="" type="checkbox"/>
otseturundus	<input checked="" type="checkbox"/>

Joonis 7.1. Märksõnaplokk "Teemamärksõna" on avatud.

## Leitud märksõnad

Märksõnade lävend:  Kopeeri

Märksõnu kokku **67**    Selekteeritud märksõnu **67**    Selekteerimata märksõnu **0**

**Teemamärksõna** ↗ ⌵

**Kohamärksõna** ⌵

Eesti	<input checked="" type="checkbox"/>
Järvamaa	<input checked="" type="checkbox"/>
Siber	<input checked="" type="checkbox"/>
NSV Liit	<input checked="" type="checkbox"/>

**Ajamärksõna** ⌵

1910-ndad	<input checked="" type="checkbox"/>
-----------	-------------------------------------

**Vormimärksõna** ⌵


e-raamatud	<input checked="" type="checkbox"/>
võrguväljaanded	<input checked="" type="checkbox"/>
infotrükised	<input checked="" type="checkbox"/>

Joonis 7.2. Märksõnaplokk "Teemamärksõna" on suletud.

## 2.6. MARC kirje kopeerimine

Kui kasutaja on sobivad märksõnad leidnud, tuleb MARC kirje puhvrissse kopeerimiseks vajutada "Kopeeri" nuppu. Seejärel on valitud märksõnade MARC kirje võimalik pasteerida soovitud kohta (vt. Joonis 8.1 ja Joonis 8.2).

**Leitud märksõnad**

Märksõnade lävend:  Kopeeri

Märksõnu kokku 8    Selekteeritud märksõnu 8    Selekteerimata märksõnu 0

**Teemamärksõna** ^

tervishoiuteenused	<input checked="" type="checkbox"/>
ravi	<input checked="" type="checkbox"/>
alkoholism	<input checked="" type="checkbox"/>
alkoholi liigtarvitamine	<input checked="" type="checkbox"/>
terviseuuringud	<input checked="" type="checkbox"/>
sõltuvushäired	<input checked="" type="checkbox"/>
tervishoid	<input checked="" type="checkbox"/>

---

**Kohamärksõna** ^

eesti	<input checked="" type="checkbox"/>
-------	-------------------------------------

Joonis 8.1. Valitud märksõnad

d	650 9	tervishoiuteenused
d	650 9	ravi
d	650 9	alkoholism
d	650 9	alkoholi liigtarvitamine
d	650 9	terviseuuringud
d	650 9	sõltuvushäired
d	650 9	tervishoid
d	651 9	eesti

Joonis 8.1. Joonisel 8.2 kuvatavate valitud märksõnade kopeeritud MARC kirje.

### 3. Veateated

Käesolev peatükk annab ülevaate võimalikest kasutajale kuvatavatest veateadetest ning nende tekkimise põhjustest.

Rakendenduses ilmuvad veateated lehe alaosa keskele (vt. Joonis 9).

Eestikeelne veateade	Ingliskeelne veateade	Tähendus
Faili formaat ei ole toetatud.	Unsupported file format.	Teaviku faili tuvastatud formaat ei kuulu toetatud formaatide hulka. Toetatud formaadid on application/epub+zip, text/html, text/plain ja application/pdf.
Faili laiend ei ole toetatud.	Unsupported file extension.	Teaviku faili tuvastatud laiend ei kuulu toetatud laiendite hulka. Toetatud laiendid on .epub, .html, .txt ja .pdf.
Faili PDF-ks konverteerimisel tekkis viga.	Error in converting file to PDF.	Kratt üritab kõik märksõnastamisele saadetud failid konverteerida PDF-iks, faili iseärasuste tõttu aga PDF-iks konverteerimine ebaõnnestus.
PDF-i pakkimisel tekkis viga.	Error in compressing PDF.	Kratt üritab kettaruumi säästmiseks kõik konverteeritud PDF-id kokku pressida, faili iseärasuste tõttu aga kokkupressimine ebaõnnestus.
PDF-i segmenteerimisel tekkis viga	Error in splitting PDF into segments.	Faili iseärasuste tõttu ei õnnestu Kratil failist segmente (lehekülgi) eraldada.
Segmentidest juhusliku alamhulga valimine ebaõnnestus.	Error in selecting random subset.	Kõikide eraldatud segmentide (lehekülgede) hulgast märksõnastajale edastatavate juhuslike lehekülgede valimine

		ebaõnnestus (Krati-sisene viga).
PDF-i segmentidest teksti eraldamine ebaõnnestus.	Error in extracting text from PDF segments.	Valitud segmentidest (lehekülgedest) teksti eraldamine ebaõnnestus. Tõenäoliseks vea põhjuseks on ühenduse probleem Tika serveriga.
Kvaliteedikontroll ebaõnnestus. Pädevaid segmente ei leitud.	Error in quality validation. No valid segments detected.	Ühestki märksõnastamiseks valitud leheküljest ei tuvastatud valideeritud teksti. See viga tuleb kõige sagedamini esile rohkelt pilte, jooniseid jms mittetekstilist sisu sisaldavate dokumentide puhul.
Keeletuvastus ebaõnnestus.	Error in language detection	Dokumendist keele tuvastamisel tekkis viga. Tõenäoliseks veapõhjuseks on ühenduse katkemine MLP teenusega või viga MLP teenuses.
Märksõnastamine ebaõnnestus.	Tagging failed.	Märksõnastamine Texta TK kaudu ebaõnnestus. Tõenäoliseks vea põhjuseks on ühenduse katkemine Texta TK liidesega või viga Texta TK liideses.
Märksõnade vormistamine ebaõnnestus.	Error in tag formatting.	Märksõnade Krati väljundi jaoks vormistamine ebaõnnestus.
Kas fail või faili URL on nõutud.	Either file or file_url is required	Märksõnastamisele saadetav teavik on sisestamata.
URL-ilt faili alla laadimine ei õnnestunud. Kontrolli URL-i sobivust	URL upload unsuccessful, check your URL.	Faili ei õnnestu URL-ilt alla laadida. Võimalikeks vea põhjusteks on ebatavalised sümbolid URL-i nimes või vigane viide.

## Automaatmärksõnastaja

Teavik läbi URL-i

Teavik läbi faili

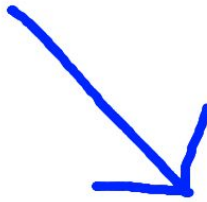
Teavik kasutaja arvutist \*

Screenshot from 2020-03-10 17:32-51.png

Lehekülgede arv \*

10

Märksõnasta



Faili formaat ei ole toetatud.

Sulge

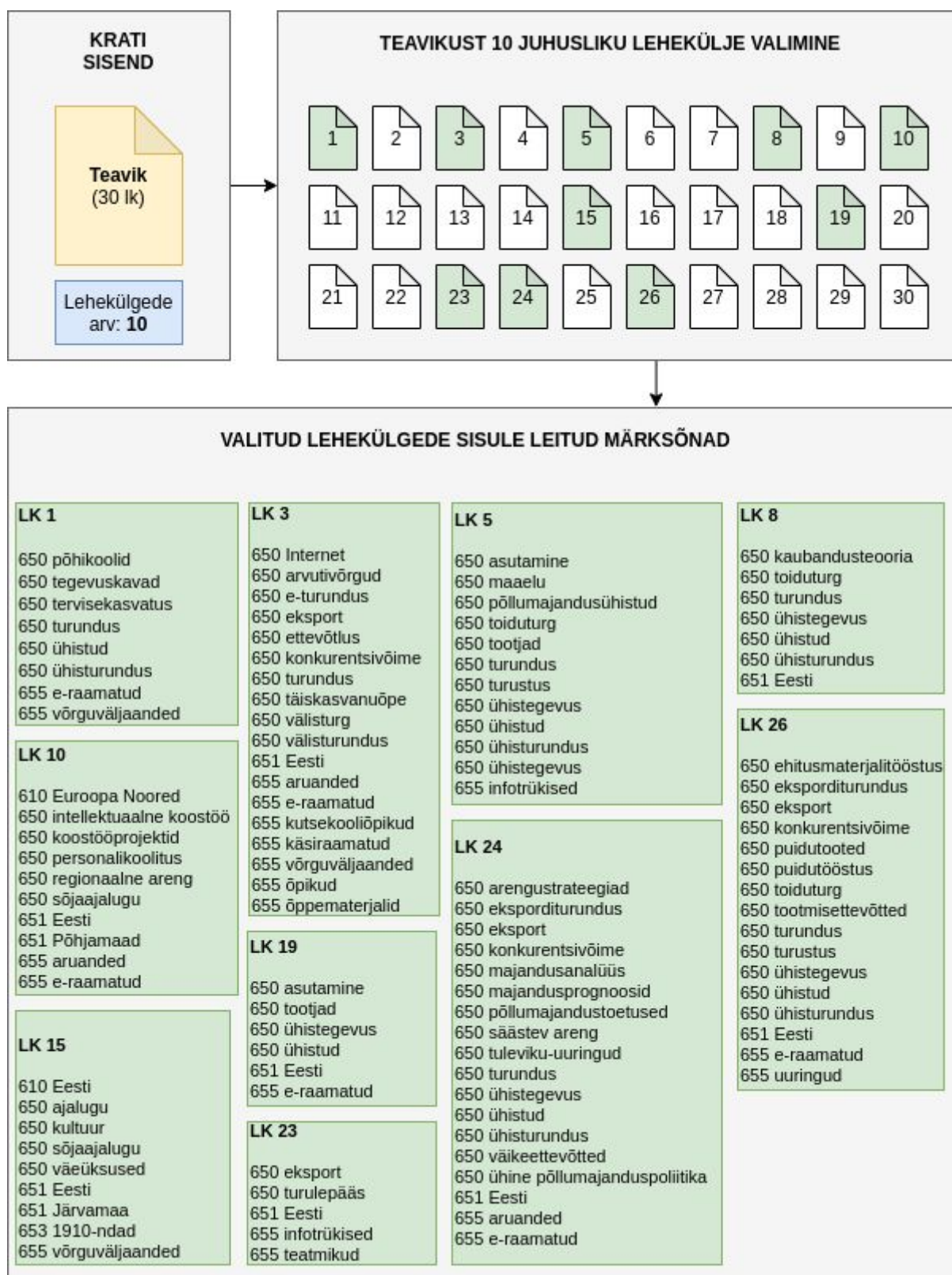
**Joonis 9.** Mittetoetatud .png formaadis faili märksõnastamisel kuvatakse kasutajale teade "Faili formaat ei ole toetatud".

## 4. Näide kasutuslooga

Joonised 10.1 ja 10.2 koos Tabeliga 1 annavad lihtsustatud ülevaate, mis toimub märksõnastamise ajal "Krati" kõhus. Näite eesmärgiks on aidata kasutajal "Lehekülgede arvu" ja "Märksõna lävendi" parameetreid ning nendevahelist seost paremini mõista, et soodustada nende efektiivsemat kasutamist.

### Kasutusloo kirjeldus

Märksõnastajale edastatakse märksõnastamiseks 30-leheküljeline teavik koos märksõnastamiseks kasutatavate lehekülgede arvuga 10. Teavikust valitakse juhuslikult 10 lehekülge ning iga valitud lehekülge märksõnastatakse eraldiseisvalt (vt. Joonis 10.1). Seejärel koondatakse kõigi 10 märksõnastatud lehekülje märksõnad kokku ning reastatakse esinemissageduse alusel (vt. Tabel 1). Märksõna suhteline sagedus väljendab, mitme märksõnastamisele saadetud leheküljega kõigist lehekülgedest oli märksõna seotud. Näeme, et kõige sagedasem oli kohamärksõna "Eesti" olles seotud 8 leheküljega 10-st (suhtelise sagedusega  $8/10 = 0.8$ ). Järgnevad teemamärksõna "turundus", teemamärksõna "ühistud" ning vormimärksõna "e-raamatud", mis tuvastati 6 leheküljelt 10-st (suhtelise sagedusega  $6/10 = 0.6$ ). Vaikelävendi 0.4 ületavad seega lisaks juba nimetatud märksõnadele veel teemamärksõnad "ühisturundus", "ühistegevus" (suhtelise sagedusega  $5/10=0.5$ ) ning "eksport" (suhtelise sagedusega  $4/10 = 0.4$ ) (vt. Tabel 1 ja Joonis 10.2).



**Joonis 10.1.** 30-leheküljelisest teavikust 10 juhusliku lehekülje valimine ning antud lehekülgedele leitud märksõnad.

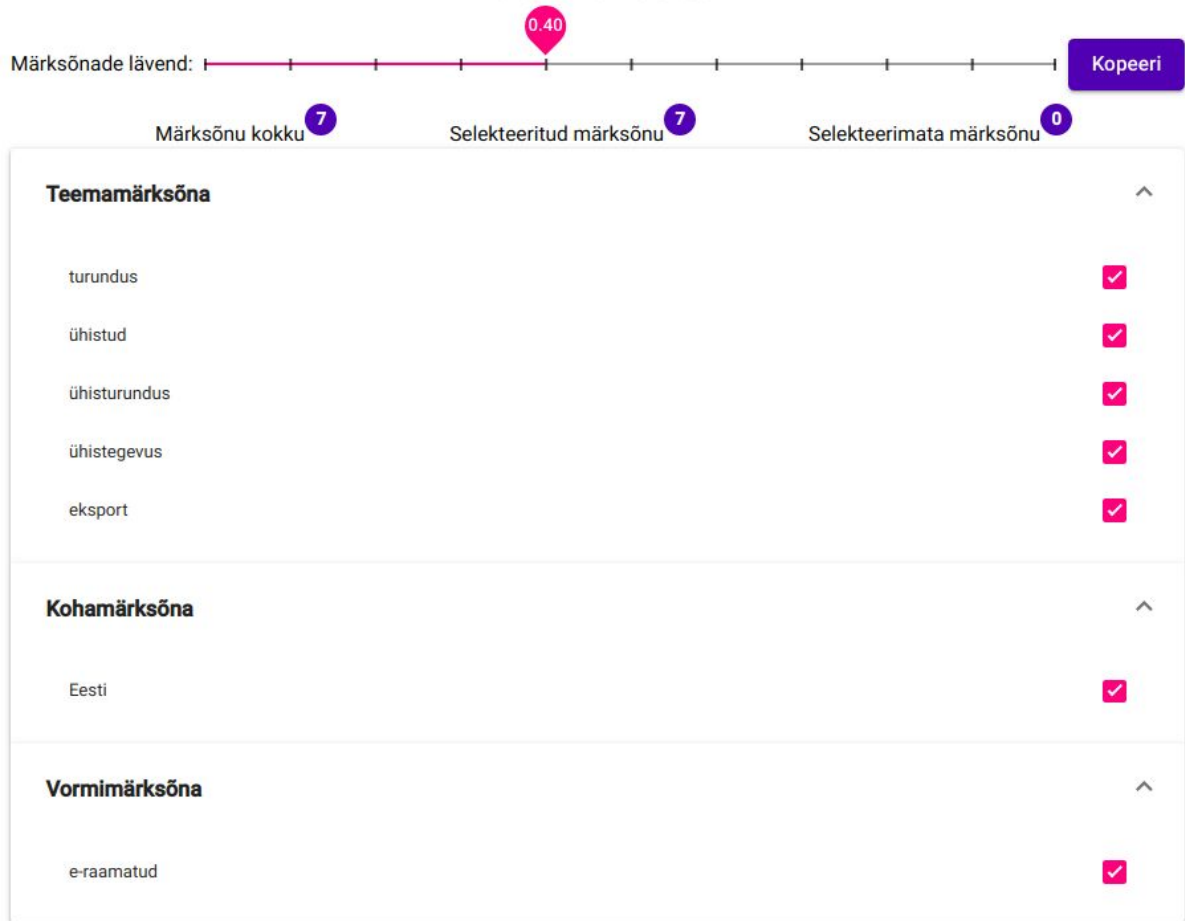


	<b>Märksõna</b>	<b>Märksõna tüüp</b>	<b>Lehekülgede arv</b>	<b>Suhteline sagedus</b>
1.	Eesti	kohamärksõna	8	0.8
2.	turundus	teemamärksõna	6	0.6
3.	ühistud	teemamärksõna	6	0.6
4.	e-raamatud	vormimärksõna	6	0.6
5.	ühistorundus	teemamärksõna	5	0.5
6.	ühistegevus	teemamärksõna	5	0.5
7.	eksport	teemamärksõna	4	0.4
8.	võrguväljaanded	vormimärksõna	3	0.3
9.	konkurentsivõime	teemamärksõna	3	0.3
10.	aruanded	vormimärksõna	3	0.3
11.	toiduturg	teemamärksõna	3	0.3
12.	asutamine	teemamärksõna	2	0.2
13.	tootjad	teemamärksõna	2	0.2
14.	turustus	teemamärksõna	2	0.2
15.	sõjaajalugu	teemamärksõna	2	0.2
16.	eksporditurundus	teemamärksõna	2	0.2
17.	tegevuskavad	teemamärksõna	1	0.1
18.	põhikoolid	teemamärksõna	1	0.1
19.	tervisekasvatus	teemamärksõna	1	0.1
20.	kutsekooliõpikud	vormimärksõna	1	0.1
21.	käsiraamatud	vormimärksõna	1	0.1
22.	õppematerjalid	vormimärksõna	1	0.1
23.	ettevõtlus	teemamärksõna	1	0.1
24.	välisurg	teemamärksõna	1	0.1
25.	välisurundus	teemamärksõna	1	0.1
26.	õpikud	vormimärksõna	1	0.1
27.	arvutivõrgud	teemamärksõna	1	0.1
28.	täiskasvanuõpe	teemamärksõna	1	0.1
29.	Internet	teemamärksõna	1	0.1
30.	e-turundus	teemamärksõna	1	0.1
31.	ühistegevus	teemamärksõna	1	0.1
32.	maaelu	teemamärksõna	1	0.1
33.	põllumajandusühistud	teemamärksõna	1	0.1
34.	infotrükised	vormimärksõna	1	0.1
35.	koostööprojektid	teemamärksõna	1	0.1

36.	intellektuaalne koostöö	teemamärksõna	1	0.1
37.	regionaalne areng	teemamärksõna	1	0.1
38.	personalikoolitus	teemamärksõna	1	0.1
39.	Põhjamaad	kohamärksõna	1	0.1
40.	puidutööstus	teemamärksõna	1	0.1
41.	tootmisettevõtted	teemamärksõna	1	0.1
42.	uuringud	vormimärksõna	1	0.1
43.	ehitusmaterjalitööstus	teemamärksõna	1	0.1
44.	puidutooted	teemamärksõna	1	0.1
45.	arengustrateegiad	teemamärksõna	1	0.1
46.	majandusanalüüs	teemamärksõna	1	0.1
47.	tuleviku-uuringud	teemamärksõna	1	0.1
48.	ühine põllumajanduspoliitika	teemamärksõna	1	0.1
49.	säästev areng	teemamärksõna	1	0.1
50.	põllumajandustoetused	teemamärksõna	1	0.1
51.	väikeettevõtted	teemamärksõna	1	0.1
52.	majandusprognoosid	teemamärksõna	1	0.1
53.	kaubandusteooria	teemamärksõna	1	0.1
54.	ajalugu	teemamärksõna	1	0.1
55.	väeüksused	teemamärksõna	1	0.1
56.	1910-ndad	ajamärksõna	1	0.1
57.	Järvamaa	kohamärksõna	1	0.1
58.	kultuur	teemamärksõna	1	0.1
59.	teatmikud	vormimärksõna	1	0.1
60.	turulepääs	teemamärksõna	1	0.1
61.	infotrükised	vormimärksõna	1	0.1

**Tabel 1.** Kõik valitud 10 leheküljelt tuvastatud märksõnad esinemissageduse alusel sorteerituna. Rohelise taustaga sektsioon tabeli alguses tähistab vaikelävendi 0.4 ületavaid märksõnu.

## Leitud märksõnad



Joonis 10.2. Läveni 0.4 ületavad märksõnad "Krati" väljundis.