

# Uuema eesti ilukirjanduse mitmikute loendid

*Uku Raudvere ja Kristel Uihoaed*

Tartu Ülikooli raamatukogu ja Eesti Rahvusraamatukogu

## Taust

Alates 01.01.2017 jõustus autoriõiguse seaduse muudatus, mis võimaldab digitaalse objekti töötlemist teksti- ja andmekaeve eesmärkidel tingimusel, et selline kasutamine ei taotle ärilisi eesmärke ([AutÕS § 19 lg 31](#)). Samal ajal hakkas kehtima ka säilituseksemplari seadus, mille järgi on kirjastustel kohustus loovutada raamatukogudele lisaks füüsilistele ka digitaalsed teosed. Seega on uus olukord laiendanud uuema ja mitmekesisema uurimismaterjali kättesaadavust.

Käesolev töö on esimene katse seda materjali kasutades luua teadlaste ja arendajate jaoks vabalt kättesaadav ressurss.

## Ressurss

Mitmikud või n-grammid (täpsemalt vt nt [Wikipediast](#)) on esitatud failidena, ühe-, kahe- ja kolmekaupa. Info on failides tabulaatoriga eraldatud (*tab-delimited*) ja organiseeritud järgnevalt: sõnavorm või mitmik, selle sagedus kõigis allikates kokku, allikate arv, kus vorm või mitmik esines.

Näiteks

```
ma ei tea 1410 80
```

tähendab, et sõnakolmik *ma ei tea* esines materjalis 1410 korda 80 erinevas teoses.

Siin on esitatud uuema eestikeelse ilukirjanduse tekstide põhjal koostatud keeleressursid:

- sõnavormide sagedusloend (`token_1_grams.txt`),
- sõnavormide bigrammid (`token_2_grams.txt`),
- sõnavormide trigrammid (`token_3_grams.txt`),
- algvormide ehk lemmade sagedusloend (`lemm_1_grams.txt`),
- algvormide ehk lemmade bigrammid (`lemma_2_grams.txt`),
- algvormide ehk lemmade trigrammid (`lemma_3_grams.txt`).

## Koostamisõhimõtted

Peale teksti puhastamist on algmaterjalis ära märgitud lausete algused (`<s>`) ja lausete lõpud (`</s>`). Seejärel koostasime sõnavormide ehk tekstisõnade, bigrammide ja trigrammide loendid.

Kuna mitmikud sisaldavad ka lause alguse- ja lõpumärgendeid, siis pole võimalik lauseid kokku panna eelnevate ja järgnevate lausetega. Samuti jätsime välja kõik n-grammid, mis esinesid ainult ühes algtekstis, nii et ükski loendis sisalduv sõnajärjend pole kokkuviidav ühe konkreetse allikaga. Niisiis on tekst piisavalt fragmenteeritud ja seda pole võimalik algsel kujul taastada.

## Kasutatud teosed

- Ühe laulu lugu / Aare Tamm
- Loti naine / Aare Tamm
- Kaks elu: [ulmeromaan] / Aare Tamm
- Varjajad / Agnes Kolga
- Põgenemine. 1. osa / Alex Lepajõe
- Ajatu rahutus / Anett Tuisk
- Taevakirju / Argo Kasela
- Minu erootika saladus / Armin Kõomägi
- Kuuvirvendus / Arvo Möldri
- Olla keegi teine / Arvo Möldri
- Minu Armeenia: aus vastus / Brigitta Davidjants
- Kirjad tütrele, ehk, Kuidas ma emaks kasvasin ja oma last kogemata maha ei tapnud / Dagmar Lamp
- Juured ja võrsed / Eet Tuule
- Mine ja kaota ennast: ühe uitaja reisipäevik Euroopast / Egert Rohtla
- Lastekirju / Epp Petrone
- Leikude / Erni Kask
- Eesti gootika XX: jutte ja lugusid aastatest 1997-2017 / Ervin Õunapuu
- Hingeside: romaan / Gen Noa
- Eneseväarikusel pole sellega mingit pistmist / Heinrich Weinberg
- Fööniksi viimane lend / Helen Käit
- Mu koju tood sa / Heli Künnapas
- Saatmata kirjad / Heli Künnapas
- Liivalaia / Heljo Mänd
- Rünkrasked pilved / Helju Pets
- Kirjailm kastepiisas / Henn Risto Mikelsaar
- Doanizarre udulaam / Indrek Hargla
- Apteeker Melchior ja Gotlandi kurat: kriminaalromaan vanast Tallinnast / Indrek Hargla
- Kingin sulle valgust / Inga Raitar
- On nagu on / Jaak Jõerüüt
- Kohtumine rotveileritega / Jaan Mikweldt
- Roheline suits / Jüri Kolk
- Linnade põletamine / Kai Aareleid
- Vesi minu veskile: videvikulood / Kaie Bergmann
- @keitivilms: Eesti esimene säutsukogumik / Keiti Vilms

- Küpsiseparadiis 2, ehk, Maskid & maskotid / Kerttu Rakke
- Kas keegi kuuleb mind? / Ketlin Priilinn
- Tupiktänavas / Kiiri Saar
- Ajapüüdja: mälestusi Kokora mailt / kogunud Andrus Kasemaa
- Minu Tšehhi: parajad Švejkid / Kristel Halman
- KrattPunktKomm / Köster
- Vikerkaare taga ja ees / Lea Jaanimaa
- Novembriöö kirjad / Mairi Laurik
- Röntgennägemisega mees / Margit Peterson
- Minu Iisrael: kuidas tõlkida sabbatit / Margit Prantsus
- Õhku joonistatud naeratus / Margit Sarapik
- Nõiatants / Margus Sanglepp
- Kuu, räägi minuga / Maria Lepmaa
- Saladus sinu sees / Maria Lepmaa
- Teine elu / Marilin Karu
- Ja sada surma: romaan / Marje Ernits
- Lenda, lenda, lepalind / Marje Ernits
- Reeturlik metsavend : Krolli esimene juurdlus / Martin Kukk
- Viimsepäeva laupäeva hommikul: novellid / Martti Kalda
- Elukargus / Mati Soonik
- Kui tee viis üle jõe / Merilin R. Jürjo
- Mõtted / Mihkel Mutt
- Sinu enda valitud teel / Mihkel Tattar
- Minu helesinine elu / Märt Aimla
- Venna arm / Oliver Berg
- Rääkimine vaikimine nutmine: vabaduse seletamatu ilu / Pille Önnepalu
- 847: kaheksasada nelikümmend seitse / Piret Lai
- San Agustini vereohvrid / Reeli Reinaus
- Muidumiljonärid / Reet Kudu
- Peidust välja / Reet Made
- Savimäe: romaan ootamatust pärandist / Rein Pöder
- Sõjaväeluure erukolonelleitnant Raimo Siig ja teised Väikelinna kangelased / Rõgiwald Pääbo
- Mees minevikust / Saale Väester
- Armuadralsündind / Sende Lipu
- Kord õide puhkeb iga aed / Siiri Laidla
- Purunenud elud. Eneken / Silja Vaher
- Käbi ja känd / Stepan Karja
- 572 sketši neljas võte / Taavi Novek
- Minu Island: tule ja jää sümfoonia / Tarvo Nõmm
- Kevadsulased nitševood: nitševood ja aastaajad. II / tekst ja karikatuurid: Raimo Aas
- Kullakene / Tiit Sepa
- Armastus pärast surma / Toomas Raudam
- Kustunud suitsud Emajõe ääres / Toots Normann

- Targa ja rumala jutud / Triin Soomets
- Liivakella helinal / Tuule Lind
- Aristarkhovi meetod / Tuuli Tolmov
- Taevatee haigla langeb koomasse / Uido Truija
- Neverland : [romaan inimestevahelistest suhetest] / Urmas Vadi
- Nädalalõpp: ühe vana mehe lugu / Valter Väljap
- Vabatahtlikuna Rios / Vivian Tiiman
- Kaks küünalt / Ülle Solovjova
- Sina ei pea mitte imestama / Aleksander Torjus
- Virtuaalne kogumik, ehk, "Algernoni" toimetus soovitab / Indrek Hargla, Kristjan Sander, Karen Or-lau, Maniakkide Tänav, Mart Raudsaar, Freya Ek, Juhan Habicht, Siim Veskimees, André Trinity, Elläi Tuulepäälse, Meelis Friedenthal, Erkki Kõlu, Marat Faizijev, August Chang, Taivo Rist, Laur Kraft, Eesti Ulmeühing
- Tööraamat / Katrin Kaarep
- Jääminek / Mathura [kirjutanud ja kujundanud]
- Трудовая книжка = Tööraamat / Merilin Pärl ja Katrin Kaarep
- Talikarged nitševood: nitševood ja aastaajad / Raimo Aas [tekst ja karikatuurid:]
- Homse maailma kirjanikud: esimesed 10 aastat / Sass Henno