



DARIAH-EU  
Digital Research Infrastructure  
for the Arts and Humanities

# Creation of Open Data Resources: Benefits of Cooperation

**Kira Kovalenko**

Institute for Linguistic Studies, Russia  
Austrian Centre for Digital Humanities, Austria

**Eveline Wandl-Vogt**

Austrian Centre for Digital Humanities, Austria



# INSTITUTIONS: GENERAL INFORMATION



Austrian Centre for Digital Humanities  
Austrian Academy of Sciences (Vienna)

established in 2015

1 department

4 working groups,

director Dr. Karlheinz Mörth

50 researchers



Institute for Linguistic Studies,  
Russian Academy of Sciences  
(St. Petersburg)

established in 1921

6 departments

director Prof. N. Kazansky

120 researchers

# JOINT PROJECTS OF THE ACDH AND ILS



## main aims:

- ✿ increase accessibility
- ✿ increase interoperability
- ✿ increase reusability
- ✿ enrich dictionary data
- ✿ interlink dictionary data
- ✿ create new workspaces
- ✿ open up dictionaries  
for research process and public curiosity

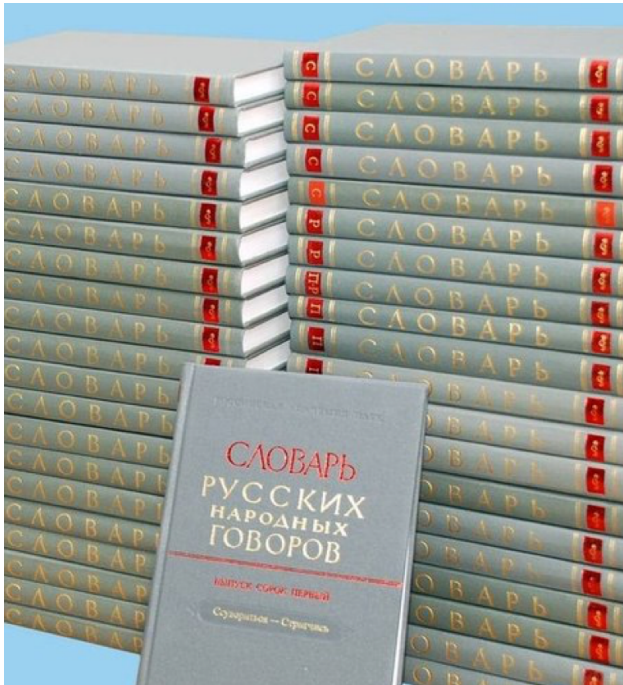


✿ A Dictionary of Russian Dialects:  
infrastructure for compilers and researchers

# A DICTIONARY OF RUSSIAN DIALECTS



## general information



since 1965

48 volumes

> 300,000 entries

chief editor

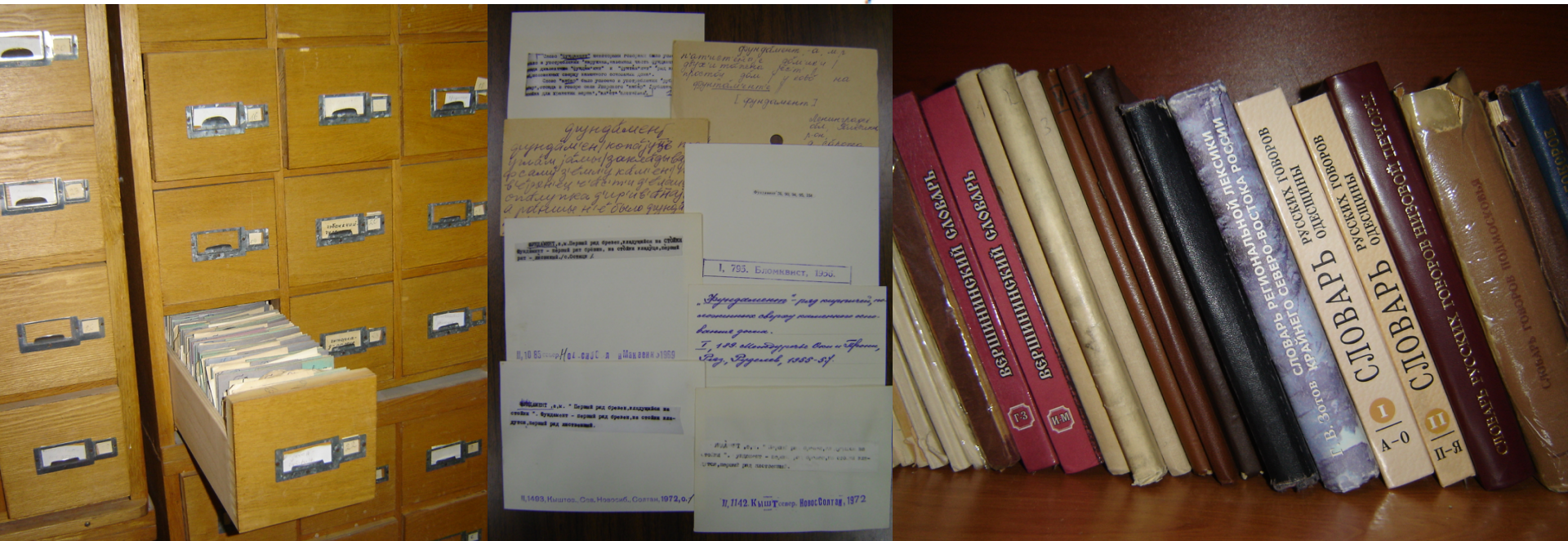
prof. Sergey Myznikov

8 compilers

# A DICTIONARY OF RUSSIAN DIALECTS



**Тычььяк, м., собир.** Колья, палки и т. п. для ограды, изгороди. Загородили тычььяком. Калинин., 1972.



# A DICTIONARY OF RUSSIAN DIALECTS



online representation

Институт лингвистических исследований РАН

конференции  
аспирантура  
библиотека

отделы  
personalia  
история

Словарь русских народных говоров

- Выпуск 1. А. 1965.
- Выпуск 2. Ба-Блазниться. 1966.
- Выпуск 3. Блазничка-Бяшутка. 1968.
- Выпуск 4. В-Военки. 1969.
- Выпуск 5. Военство-Выростковый. 1970.
- Выпуск 6. Выросток-Гон. 1970.
- Выпуск 7. Гона-Депеть. 1972.
- Выпуск 8. Дер-Ерепениться. 1972.
- Выпуск 9. Ерепена-Заглазеться. 1972.
- Выпуск 10. Заглазки-Заросить. 1974.
- Выпуск 11. Зароситься-Зубрёнка. 1976.
- Выпуск 12. Зубрёха-Калумаги. 1977.
- Выпуск 13. Калун-Кобза. 1977.
- Выпуск 14. Кобзарик-Корточки. 1978.
- Выпуск 15. Кортусы-Куделюшки. 1979.
- Выпуск 16. Куделя-Лесной. 1980.

РОССИЙСКАЯ АКАДЕМИЯ НАУК  
ИНСТИТУТ ЛИНГВИСТИЧЕСКИХ ИССЛЕДОВАНИЙ

## СЛОВАРЬ РУССКИХ НАРОДНЫХ ГОВОРОВ

ВЫПУСК СОРОК ШЕСТОЙ

Тычак — Ужоля



Санкт-Петербург  
«НАУКА»  
2013

<https://iling.spb.ru/vocabula/srng/srng.html>

ReferentIn: K. Kovalenko, E. Wanld-Vogt Thema: Creation of Open Data Resources...

36

Тюрокаться

тюркает, вот-вот разорвет. Чаго-дош. Волгог., 2005.

3. При игре в прятки — консуется рукой другого играющего. В хоронки играли: который затюркает, в нево-ды буди; тюркает, который идет. Медвежьегор. Карел., 2005.

**Тюрокаться**, несом. Стучать во что-л. (чтобы открыли). Те сваты было тюркались в ворота, да мы с кем и ушли. Клян. Краснояр., 1985.

**Тюрки**, мн. Мужское нижнее белье, кальсоны. Морд., 2006.

**Тюрлик**, м. Птица степной кулик, крошкеш. Тюрлик — птичка здесь водится, он кричит, когда воде прибывает. Р. Урал, 1976.

**Тюря**, ж. То же, что 1. Тюря (в 1-м знач.). Крошит matka хлеб, ложит масло с салью, ешьте тюрлю. Одесск., 2001.

**Тюрма**, ж. 1. Место заключения преступников или подсудимых, тюрьма. Крепка тюрма, да черт ей рай! От сумы да от тюрьмы не отвяжись (как раз угодись). В тюрму широка дорога, а из тюрмы тесна. Даль. Слов. Акад. 1963 [устар.].

2. Изба, построенная окнами в противоположную югу сторону. Дана стропило окнами на юг, избу, построено окнами в другую сторону, называют тюрмою. Котельн. Вят., Архив РГО.

**Тюрмал**, м. Костяной наконечник шеста, которым управляют оленями в упряжке. В правой руке оужа, а в левой хорей, на конце нахорейник, приколочивают тюрмал, это нахорейник, чтобы не проколоть олена. Низ. Печора, 2005.

**Тюрмэк**, м. Растение [какое?], употребляемое как краситель. «Красильное». Наурская Терск. казан., 1908.

**Тюрмо** и **тюрмо**, ср. Трюмо. □ Тюрмб. Любин. Яросл., 1968. □ Тюрмб. Тюрмо, а тогда (наме-

вали) зеркала. Р. Урал, 1976.

**Тюрмош**, м. Вязаная льна. «Мятого, но не трепаного, весом примерно с 1 пуд». Этот тюрмоши завязки и ладно. Опоч. Пск., 1957.

**Тюрнэс**, м. Кормовая репа, турнепс. Тюрнэс, она, знаешь, тюра и называется турнепс. Пинеж. Арх., 1960.

**Тюрнэс**, м. Кормовая репа, турнепс. Тюрнэс как редька, редька только сорская, а тюрнэс не. Пинеж. Арх., 1960.

1. **Тюрник**, м. 1. Любитель тюри. Даль.

2. О глупом человеке. Тюрниками мы глупых людей зовем. Морд., 2006.

3. Прозвище. «За то, что летом они не топят печей, а едят одну тюрю». Наровн. Пеня., Архив РГО.

2. **Тюрник** и **тюрник**, м. То же, что тюрник. □ Тюрник. Тюрник — дырка серая птица, в степе ее гнезды, с курами привычка ходить вместе, с осени улетает. Р. Урал, 1976. □ Тюрник. «Название произошло от звукоподражания крику птицы». Уральск., Карлов, 1908.

**Тюрнуться**, сов. Ввалиться, понасть во что-л. Одна корова передом тюрнулась [в глинистую почву], передице-то ноги уторкнули, а заднице вверх. Пинеж. Арх., 1958.

1. **Тюрбк**, м. То же, что 1. Тюрбк (в 1-м знач.). Выюшка-то? Как на клубок вьешь, а потом наивыешь. Чадоды-то ее по дурному звали — тюрбк. Ср. Прииртыше, 1993.

2. **Тюрбк**, м. Свернутая наподобие воронки береста или бумага. Вят., 1892.

**Тюрбонница**, ж. Затворница. Matka ная у тюрбонцы, как тюрбонница. Одесск., 2001.

**Тюрбомок**, м. Растение Tagaxum officinale Wegg., сем. слож-



# A DICTIONARY OF RUSSIAN DIALECTS



## methods & standards:

- ✿ TEI P5
- ✿ Lemon model
- ✿ Ontolex core model
  - with extentions:
    - ✿ synsem
    - ✿ decomposition
    - ✿ variation and translation
    - ✿ linguistic metadata



# JOINT PROJECTS OF THE ACDH AND ILS

---



- ✿ A Dictionary of Russian Dialects:  
infrastructure for compilers and researchers
- ✿ Russian Manuscript Lexicons:  
infrastructure for researchers

# RUSSIAN MANUSCRIPT LEXICONS



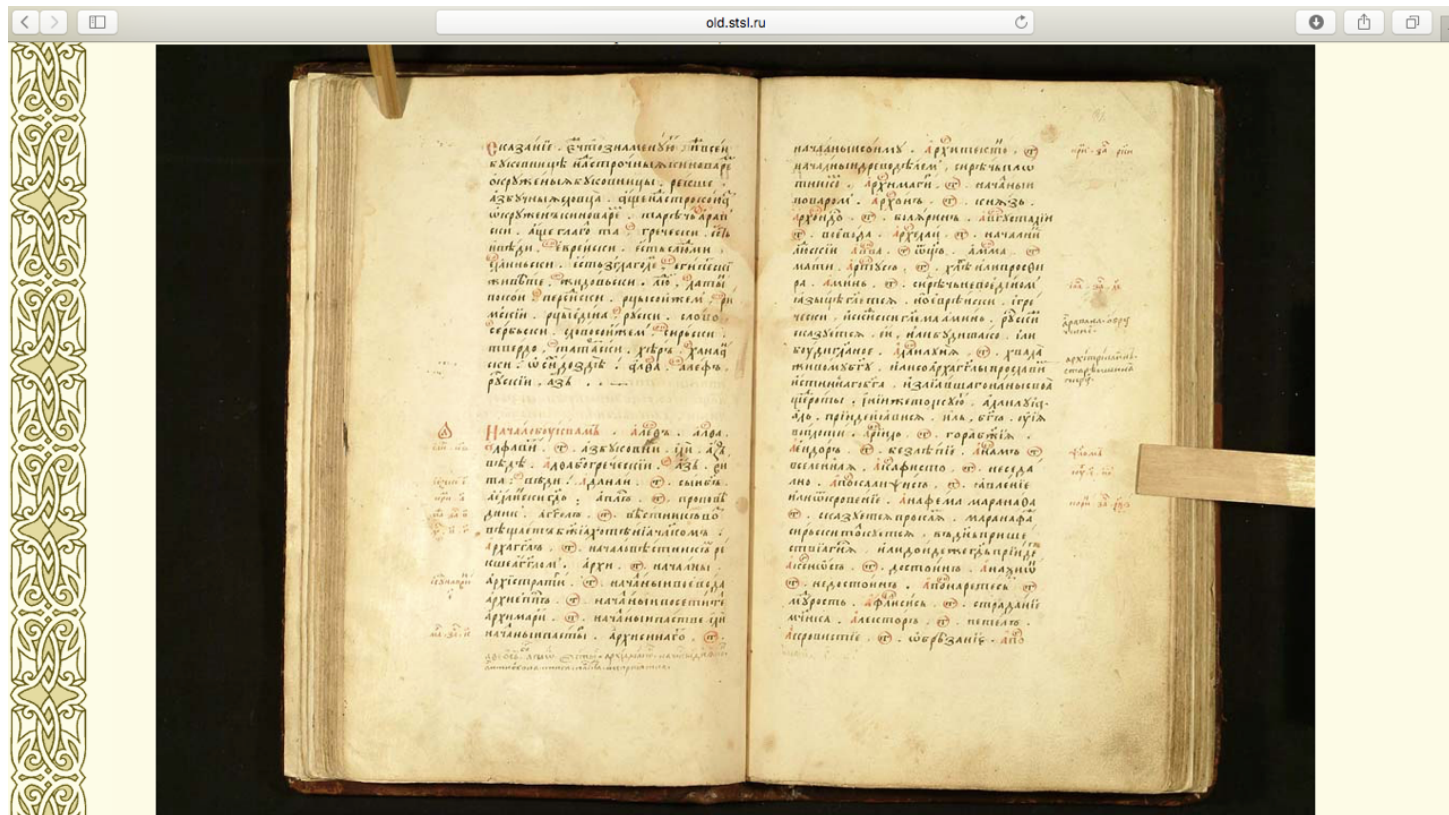
## general information

- ✿ created in the middle of the 16<sup>th</sup> century
- ✿ alphabet arrangement
- ✿ close connection to the text
- ✿ more than 9 types
- ✿ from 700 up to 16 000 word entries
- ✿ foreign words from Greek, Latin, Hebrew, Church Slavonic, Ruthenian, Tatar, Arab, and German origin
- ✿ about 150 lexicons
- ✿ important source for historical lexicography

# RUSSIAN MANUSCRIPT LEXICONS



online representation



<http://old.stsl.ru/manuscripts/medium.php?col=5&manuscript=035&pagefile=035-0084>

# RUSSIAN MANUSCRIPT LEXICONS



## methods & standards

- ✿ TEI P5
- ✿ *cr\_xq*: a standards-based, fully configurable publication framework for XML data
- ✿ full-text search and field-specific searches
- ✿ synoptic view of facsimile and text
- ✿ computer-assisted collation and stemma creation
- ✿ facilitates creation of various indexes of tagged information in the ingested resources (lemma list, index of translations by language etc.)
- ✿ standards: METS, FCS-SRU, currently mostly used for TEI content (e.g. Austrian Baroque Corpus ABaC:us <http://acdh.oeaw.ac.at/abacus>)



- ✿ A Dictionary of Russian Dialects:  
infrastructure for compilers and researchers
- ✿ Russian Manuscript Lexicons:  
infrastructure for researchers
- ✿ Russian Plant Names  
in the Diachronic Aspect (11-17 cc.):  
database with a search engine



## general information

- ✿ starts in 2017
- ✿ time frame: 11-17 cc.
- ✿ team: researchers with different backgrounds
- ✿ sources:
  - primary (manuscripts, printed books of the 11-17 cc.)
  - secondary (historical dictionaries, modern researches, etc.)
- ✿ database with a search engine
- ✿ interoperability with the Austrian plant names database (exploreAT!)
- ✿ contribution to the project Biodiversity and Linguistic Diversity (<https://reconcile.eos.arz.oeaw.ac.at>)
- ✿ collaborative Knowledge Discovery Environment



## information structure

- \* Old Russian name
- \* modern Russian name
- \* foreign etymon (if loanword)
- \* type of representing a foreign phytonym (translation, transcription, transliteration, loan translation, hyperon)
- \* Latin name
- \* example of the use from text/dictionary
- \* bibliographical information (if printed: author, title, place of publication, date, page, genre; if manuscript: genre, author, name, location, page, etc.)
- \* use of the plant
- \* symbolical meaning

# INTERNATIONAL RESOURCES



## EUROPEANA

OpenSearch X Add a search term

REFINE YOUR SEARCH

MEDIA

- Image (300)
- Only items with links to media

CAN I USE IT?

- Yes With Restrictions (165)
- Yes With Attribution (123)
- Only With Permission (12)

COUNTRY

- United Kingdom (187)
- Austria (87)
- The Netherlands (8)
- Slovakia (6)
- Estonia (4)

More

LANGUAGE

- English (187)
- Deutsch (85)
- Nederlands (8)
- Slovenský (6)
- Eesti (4)

More

AGGREGATOR

- OpenJpt (296)
- Athena (4)

INSTITUTION

1 - 12 of 300 results

Per page: 12

Taraxacum ceratophorum (Ledeb.) DC.

Taraxacum ceratophorum (Ledeb.) DC.

Taraxacum ceratophorum (Ledeb.) DC.

Taraxacum ceratophorum Ledeb.

Taraxacum ceratophorum Ledeb.

Taraxacum ceratophorum (Ledeb.) DC.

Taraxacum ceratophorum Ledeb.

Taraxacum ceratophorum Ledeb.

Taraxacum ceratophorum (Ledeb.) DC.

Taraxacum ceratophorum Ledeb.

Taraxacum ceratophorum Ledeb.

Taraxacum ceratophorum (Ledeb.) DC.

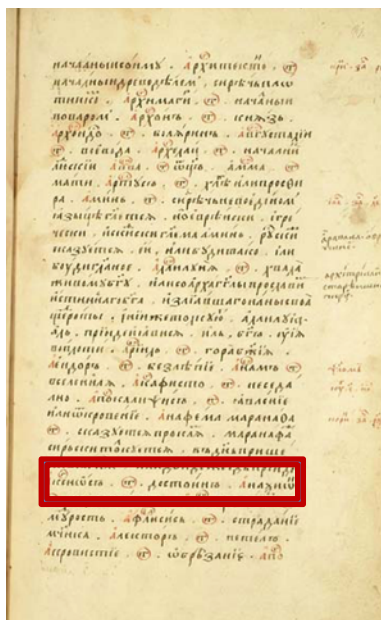




# INTERNATIONAL RESOURCES



## EUROPEANA



Technical metadata

**Title** Taraxacum ceratophorum (Ledeb.) DC.

**Classifications**

**Type:** Specimen, <http://dbpedia.org/resource/Specimen>

**Subject:** Одуванчик рогоносный, Одуванчик рогатый, Одуванчик алтайский, Одуванчик хамарский, Одуванчик короткорогий, Одуванчик Шамиссо, Одуванчик Хультена, Одуванчик ключевской, Одуванчик сахалинский, Одуванчик Малеза, Одуванчик короткорогатый, Одуванчик карагинский, <http://openup.nhm-wien.ac.at/commonNames/references/scientificName/141536>

**Has Type:** <http://rs.tdwg.org/dwc/dwctype/Specimen>

**Concept:** Одуванчик рогоносный, Одуванчик рогатый, Одуванчик алтайский, Одуванчик хамарский, Одуванчик короткорогий, Одуванчик Шамиссо, Одуванчик Хультена, Одуванчик ключевской, Одуванчик сахалинский, Одуванчик Малеза, Одуванчик короткорогатый, Одуванчик карагинский, Specimen

**Time** Creation Date:

**FIND OUT MORE**

Europeana is displaying this Item via the [Biologisches Zentrum Der Oberoesterreichischen Landesmuseen](#). Visit their site for more information.

**CAN I USE IT?**

Yes with attribution  
 CC BY-SA

**SHARE**



# BENEFITS OF COOPERATION



- ✿ collaborative approach, establish sustainable workflows
- ✿ shared use of unified or de facto standards and infrastructures; develop new standards
- ✿ experimenting with new methods and emerging technologies
- ✿ open up new data
  - ✿ dictionaries
  - ✿ manuscripts
  - ✿ lexical data
  - ✿ cultural data
- ✿ connection of data to the global resources
- ✿ more and better results for both cooperation partners
- ✿ sharing failures, risks and furthering learning and innovation
- ✿ improving competitiveness and visibility
- ✿ SHARING DIGITAL TRANSFORMATION AND INNOVATION





Thank you  
for your attention!