



# 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 Mörtl

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

# JOINT PROJECTS OF THE ACDH AND ILS

---

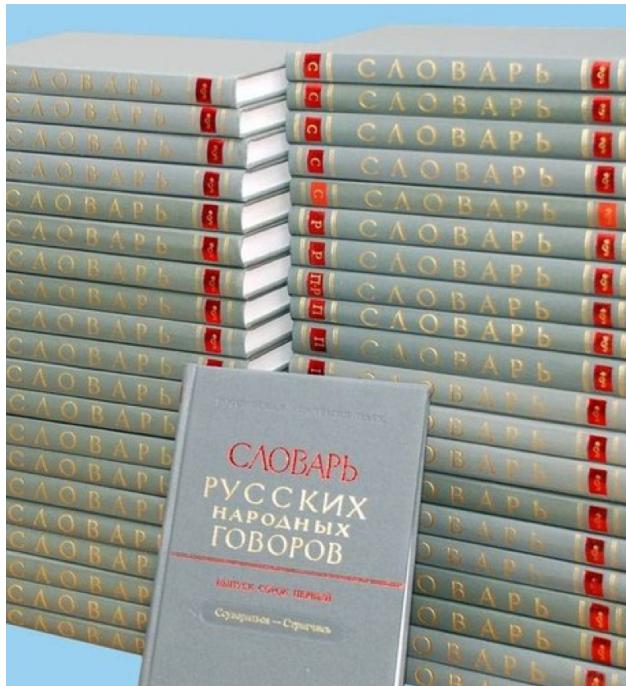


❖ A Dictionary of Russian Dialects:  
infrastructure for compilers and researchers

# A DICTIONARY OF RUSSIAN DIALECTS



## general information

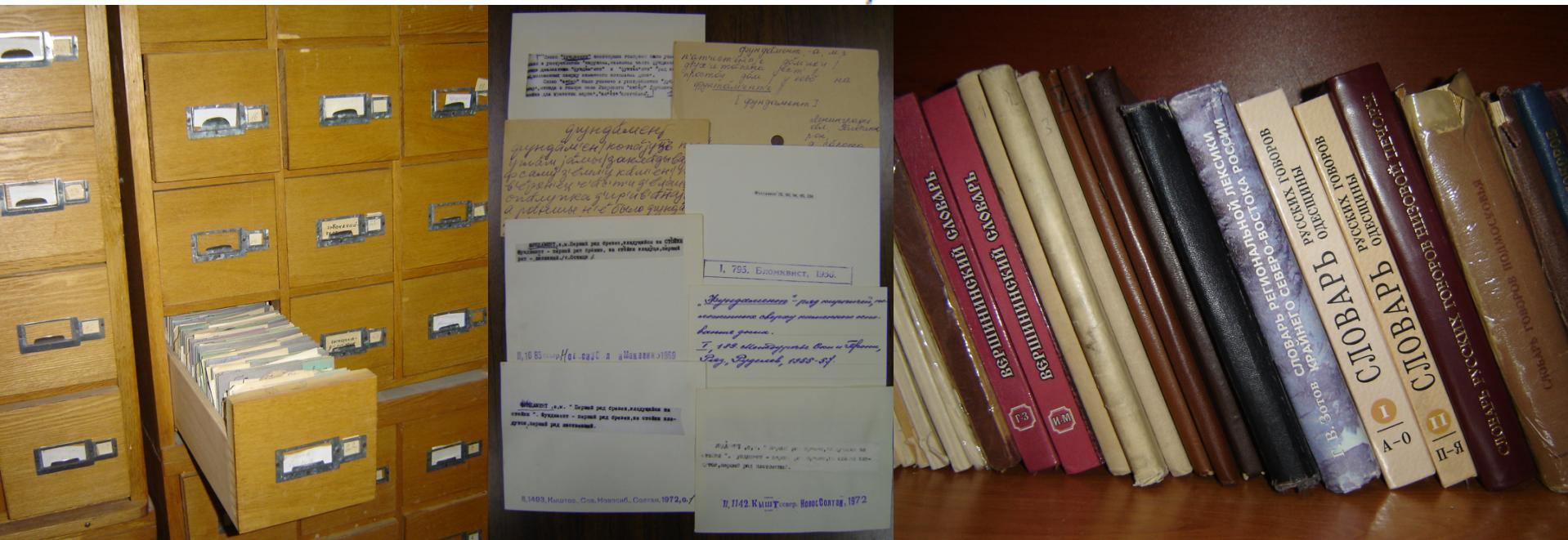


since 1965  
48 volumes  
> 300,000 entries  
chief editor  
prof. Sergey Myznikov  
8 compilers

# A DICTIONARY OF RUSSIAN DIALECTS



**Тычъяк, м., собир.** Колья, палки и т. п. для ограды, изгороди. *Загородили тычьяком.* Калин., 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. Блазнишка-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. Ерепеня-Заглазеться. 1972.

Выпуск 10. Заглазки-Заросить. 1974.

Выпуск 11. Зароситься-Зубренка. 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

36

Тюркаться

тюркает, вот-вот разорвет. Чагодыс. Волг., 2005.

3. При игре в прятки — коснуться рукой другого игрока. В городке играли когда-то злодеи, в деревне — бояре; тюркет, который ищет. Медведевогор. Карел., 2005.

Тюркаться, несов. Стучать во что-л. (чтобы открыли). Те скамьи долго тюркались ворота, да ни с чем и ушли. Юлии. Краснова., 1988.

Тюрики, мн. Мужское нижнее белье, кальсоны. Морд., 2006.

Тюрилик, мн. Птица степной кулик, кроншнеп. Тюрилик — птичка здесь водится, он кричит, когда водится прибывать. Р. Урал., 1976.

Тюриля, ж. То же, что 1. Тюрия (в 1-м знач.). Крошают мякоть хлеба, ложат масло с солью, ешьтъ тюрилю. Одесск., 2001.

Тюромъ, ж. 1. Место заключения преступников или подсудимых, тюрьма. Крепкая тюрьма, да чертъ ей радъ! От смысла да от тюрины не отпрекайся (как раз угодила). В тюришу широка дверь, а из тюрины тесна. Даль. Слов. Акад. 1963 [устар.].

2. Изба, построенная окнами на противоположную югу стороны. Дома строят окнами на юг, избу, построенную окнами в другую сторону, называют тюрию. Котельн. Вят., Архив РГО.

Тюромъл, ж. Костяной наконечник шеста, которым управляли оленями в упряжке. В правой руке вожжа, а в левой хорей, на конце хорейчик, прикалачивают тюромъл, это нахорейник, чтобы не прокололть оленя. Низ. Печора., 2005.

Тюремѣ, ж. Растение [какое?], употребляемое как краситель. «Красильное». Наурская Терск. казач., 1908.

Тюромъ и тюромъ, ср. Тюромъ. Тюромъ б. Любим. Яросл., 1968. Тюромъ б. Любим. Тюромъ, а тохда (называлось тюрок). Ср. Принртышие, 1993.

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

Тюромница, ж. Заворнича. Матка мок у птички, как тюромница. Одесск., 2001.

Тюромок, м. Растение Tagahancum officinale Wegg., сем. сложно-

вали) зеркало. Р. Урал., 1976.

Тюромощь, м. Вязанка лыни. «Мягкого, но трепанного, весом примерно с 1 пуд». Этот тюромощь заявляя и полно. Опоч. Пск., 1957.

Тюренѣц, м. Коровья репа, турнепс. Тюренец, она, знаешь, тура и называется тюренец. Пинеж. Арх., 1960.

Тюренѣс, м. Коровья репа, турнепс. Тюренес как редика, редика только горькая, а тюренес не. Пинеж. Арх., 1960.

1. Тюриник, ж. 1. Любитель тюри. Дасть.

2. О глупом человеке. Тюрикали мы глупых людей зовем. Мора., 2006.

3. Прозвище. «За то, что летом они не топят печей, а едят одну тюри». Нароч. Пенз., Архив РГО.

2. Тюриник и тюрионик, ж. То же, что тюрилик. Тюри и ник. Тюриник — дикая серая птица, в стете ее звезды, с кудрами привычна ходить вместе, к осени улетает. Р. Урал., 1976. Тюри и ник. «Название произошло от звукоподражания крику птицы». Уральск., Карлов., 1908.

Тюринуться, сол. Взаливаться, поспать в чём-то. Одна корова перед тем тюринулась [в глинистую почву], переднее-то ноги утопрекнули, а задние-то вверх. Пинеж. Арх., 1958.

1. Тюрок, м. То же, что 1. Тюри и ник (в 1-м знач.). Вьюска-то? Как на кладбище вьевши, а потом начавши. Чадончи-то ее по другому звали — тюрок. Ср. Принртышие, 1993.

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

Тюромница, ж. Заворнича. Матка мок у птички, как тюромница. Одесск., 2001.

Тюромок, м. Растение Tagahancum officinale Wegg., сем. сложно-

# A DICTIONARY OF RUSSIAN DIALECTS



methods & standards:

- ❖ TEI P5
- ❖ Lemon model
- ❖ Ontolex core model
  - with extention:
  - ❖ 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



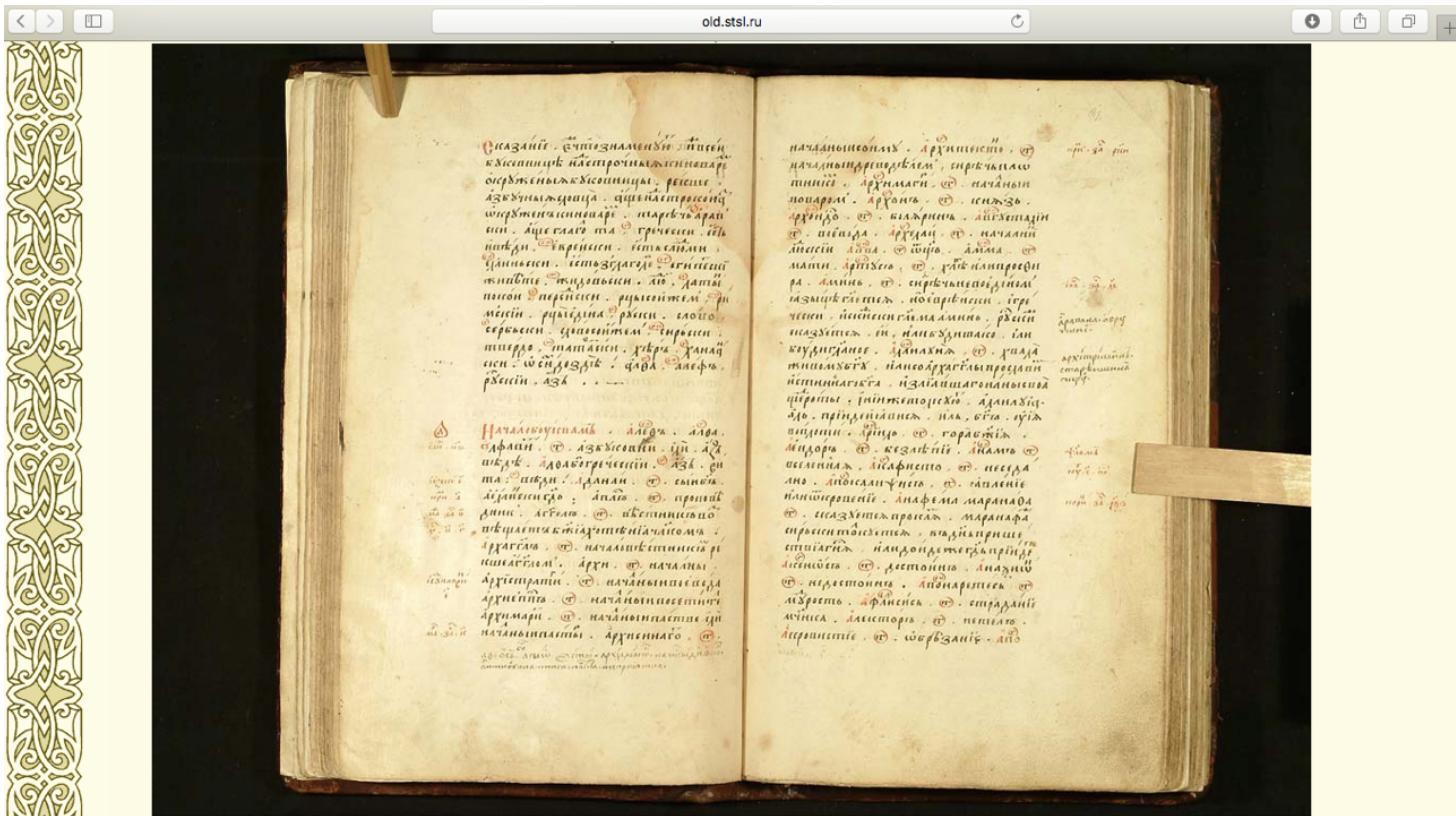
## general information

- ❖ 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



# 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>)

# JOINT PROJECTS OF THE ACDH AND ILS



- ❖ 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

Search bar: 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

- OpenJPL (296)
- Athena (4)

INSTITUTION

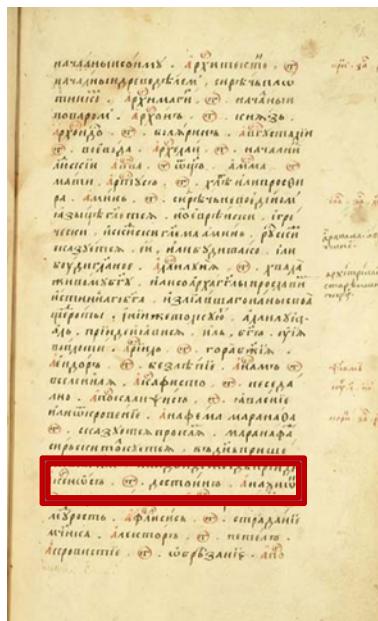
1 - 12 of 300 results

Per page: 12 GRID LIST

# INTERNATIONAL RESOURCES



## EUROPEANA



### Technical metadata

|                 |  |
|-----------------|--|
| Title           | Taraxacum ceratophorum (Lebed.) DC.  |
| Classifications | Type: Specimen , <a href="http://dbpedia.org/resource/Specimen">http://dbpedia.org/resource/Specimen</a><br>Subject: Одуванчик рогоносный, Одуванчик рогатый, Одуванчик алтайский, Одуванчик хамарский, Одуванчик короткогоркий, Одуванчик Шамиссо, Одуванчик Хульстена, Одуванчик ключевской, Одуванчик сахалинский, Одуванчик Малеза, Одуванчик короткогоратый, Одуванчик карагинский, <a href="http://openup.nhm-wien.ac.at/commonNames/references/scientificName/141536">http://openup.nhm-wien.ac.at/commonNames/references/scientificName/141536</a><br>Has Type: <a href="http://s.idwg.org/dwc/dwctype/Specimen">http://s.idwg.org/dwc/dwctype/Specimen</a><br>Concept: Одуванчик рогоносный, Одуванчик рогатый, Одуванчик алтайский, Одуванчик хамарский, Одуванчик короткогоркий, Одуванчик Шамиссо, Одуванчик Хульстена, Одуванчик ключевской, Одуванчик сахалинский, Одуванчик Малеза, Одуванчик короткогоратый, Одуванчик карагинский, Specimen |
| Time            | Creation Date:   |

### FIND OUT MORE

Europeana is displaying this item via the Biologische Zentren der Oberösterreichischen 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!