



# **Вікіджерела — вільна електронна бібліотека, яку можуть наповнювати всі охочі**

**5-а міжнародна науково-практична конференція  
«Наукова комунікація в цифрову епоху»  
29-30 березня 2017 року, Київ**





# Проект Wikisource:

- Заснований у **2003** році
- **65** мовних розділів
  - **8** розділів (de, en, fr, es, pl, ru, zh, he) **>100 тис.** текстів





# Українські Вікіджерела:

- Засновані у **2005** році  
(окремий домен з **2006**)
- **8,3 тис.** текстів  
українською — **30-те** місце  
(відповідає місцю мови в  
світі по кількості native  
speakers)



## Відвідуваність:

- Wikisource:
  - Головна: **160 тис.** (місяць)
  - Всі сторінки: **3790 тис.** (місяць)
- Вікіджерела:
  - Головна: **3,2 тис.** – **37-те** місце
  - Всі сторінки: **99 тис.**





## Score проекту:

- Вільні тексти українською мовою





# Вільні:

- Не є об'єктами авторського права:
  - закони
  - документи
- Звільнені автором:
  - ліцензовані під Creative Commons
  - або передані в суспільне надбання
- Термін дії авторського права закінчився





## Тексти:

- Вичитані тексти з прив'язкою до сканованої копії:
  - не самі лише скановані копії і не суто тексти
  - не машинні, а вичитані людиною після OCR й перевірені іншою
- Такі тексти:
  - можна перевірити на відповідність оригіналу
  - здійснити пошук по тексту
  - тексти індексуються пошуковими системами ( $\frac{2}{3}$  відвідувачів — з пошуку)







## Українською мовою:

- Тільки тексти та переклади українською мовою
- Також переклади здійснені редакторами (користувачами) проекту
- Тексти іншими мовами — в інших мовних розділах





# Що можна знайти на Вікіджерелах?

- Біблія (Огієнка)
- Українська класика (Енеїда 1798 та 1922 років видань, Кобзар тощо)
- Сучасні наукові статті опубліковані під вільною ліцензією
- Історичні документи та їх переклади
- Закони, підзаконні акти в т. ч. й історичні
- Художня література та переклади
- Вірші
- ....



## Нещодавно додані (приклад):

- Kobzar' Tarasa Szewczenka польською латинкою, Львів, 1914
- Статті Вячеслава Липинського
- Переклади промови й наказу Трампа
- П. Дяків, Товариство «Міщанське брацтво» в Тернополі, з нагоди 40-ліття, 1931
- Олександр Пивень, Веселымъ людямъ на втиху!, Москва, 1906 — етнографічна збірка, ярижкою
- Переклад кількох звітів ЦРУ
- ...



# Найпопулярніші тексти

— наразі спеціалізовані тексти цікаві окремим згуртованим групам:

- Твори шкільної програми (сезонно), особливо короткі
- Закони: українські, міжнародні, Універсали Центральної Ради тощо
- Біблія, молитви
- Художня література





## В планах:

- Енциклопедії, словники:
  - Енциклопедія Історії України
  - Енциклопедія Українознавства
  - Українські кобзарі, бандуристи, лірники.  
Енциклопедичний довідник
- Документи, закони, архіви:
  - Акти, що стосуються природно-заповідного фонду
  - Архів Буковського, документи з електронної читальні ЦРУ тощо
  - Історичне законодавство
- Українська література ХІХ ст., а також бібліографії, історія літератури, антології тощо
- Шкільна програма

# Кількість українських книжок

За даними „Бібліографічного покажчика нової української літератури. (1798—1883 р.)“ Михайла Комарова, Київ, 1883 р.

Здебільшого дані по підросійській Україні, бо до галицьких видань автор мав обмежений доступ

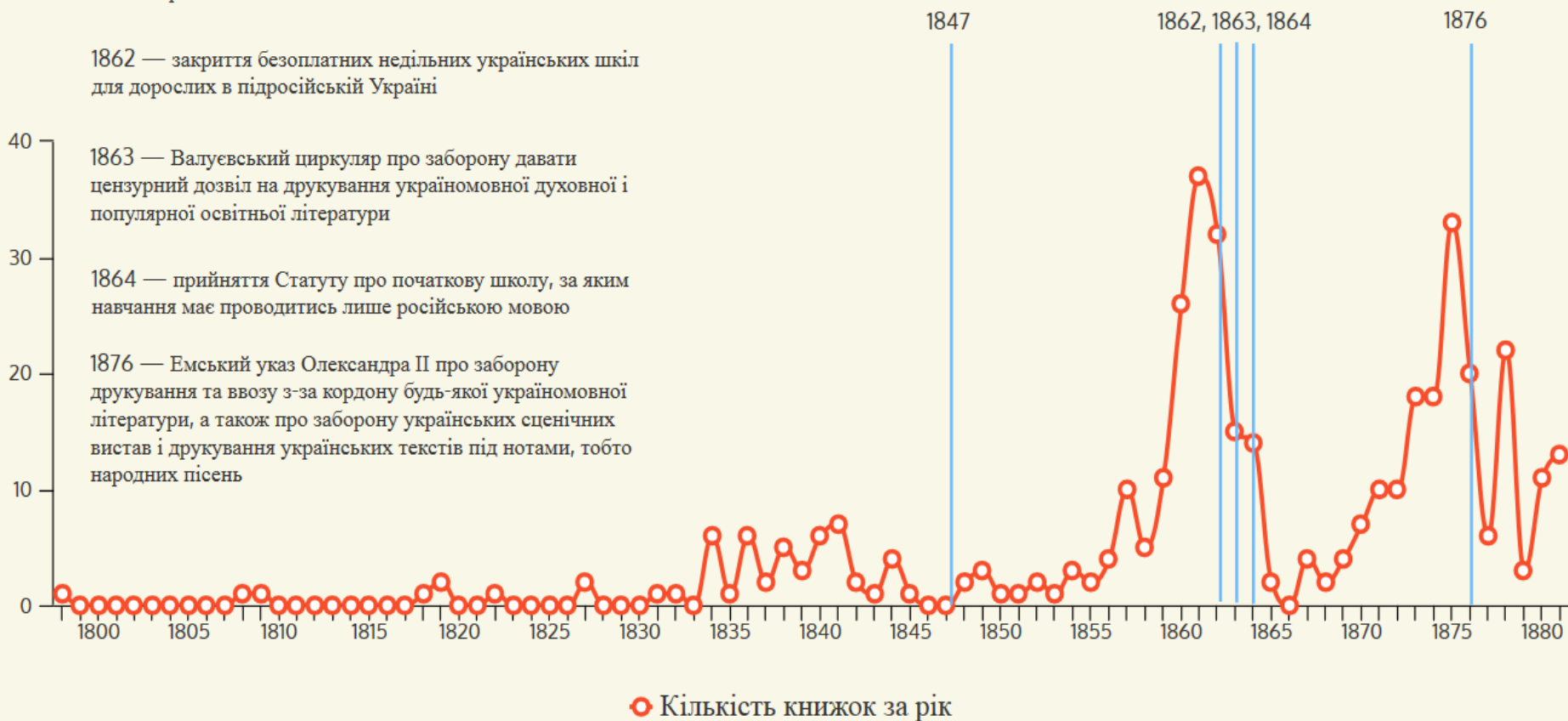
1847 — розгром Кирило-Мефодіївського товариства й посилення жорстокого переслідування української мови та культури, заборона найкращих творів Шевченка, Куліша, Костомарова та інших

1862 — закриття безоплатних недільних українських шкіл для дорослих в підросійській Україні

1863 — Валуєвський циркуляр про заборону давати цензурний дозвіл на друкування україномовної духовної і популярної освітньої літератури

1864 — прийняття Статуту про початкову школу, за яким навчання має проводитись лише російською мовою

1876 — Емський указ Олександра II про заборону друкування та ввозу з-за кордону будь-якої україномовної літератури, а також про заборону українських сценічних вистав і друкування українських текстів під нотами, тобто народних пісень





# В якому вигляді текст зберігається

- Джерело (скан, зазвичай, у форматі PDF чи DjVu з OCR чи без нього)
- Окремі вичитані сторінки, одна за одною, як вони фізично розташовані в скані
- Книга (твір, стаття), яка акцептує текст зі сторінок, причому:
  - текст з різних сторінок об'єднується
  - і далі розбивається на глави, розділи



# Як ведеться робота над текстом

1. Пошук тексту для вичитки
2. Отримання дозволу на текст  
(якщо текст не є вільним)
3. Сканування (якщо сканована копія відсутня або поганої якості)
4. Заливка скану на сайт проекту







# Як ведеться робота над текстом

5. Вичитка тексту на кожній сторінці:
  - a. Автоматична, OCR
  - b. Ручна перевірка, напів автоматична  
заміна регулярних помилок
6. Створення чи відновлення структури: глави, розділи
7. Перевірка вчитаного тексту іншим користувачем







# Робота над текстом. Індекс:



ВІКІДЖЕРЕЛА

Головна сторінка  
Нові редагування  
Нові сторінки  
Автори  
Випадкова робота  
Випадковий автор  
Випадкове оцифрування

Вигляд II  
Учась  
Скрипторій  
Довідка  
Пожертування

Інструменти  
Послання суди  
Порівняні редагування  
Завантажити файл  
Спеціальні сторінки  
Постійне посилання  
Інформація про сторінку  
Цитувати сторінку  
Статистика відвідувань  
Послання за ID  
Підсторінки

Мови  
Додати послання

Artem.komisarenko Обговорення Налаштування Бета Список спостереження Внесок Вийти

Індекс Обговорення

Читати Редагувати Переглянути історію Ще Пошук у Вікіджерелах

## Індекс:Kobzar'. Wybier dejakych najkрасzczych poezyj Tarasa Szewczenka. Lwiw, 1914.pdf



**Назва** Kobzar'  
**Друга назва** Wybier dejakych najkрасzczych poezyj Tarasa Szewczenka  
**Автор** Taras Szewczenko  
**Редактор** Wołodumyr Ochrymowycz  
**Рік** 1914  
**Видавець** Z drukarni „Dila” Rynek 10. Dim „Proświty”  
**Місце видання** Lwów  
**Сканування** pdf  
**Прогрес** До вичитки

Zmist (de szczo je):

Зверніть увагу на інформацію щодо форматування, вказану на:  
Обговорення індексу<sup>1</sup> на сторінці: Wybier dejakych najkрасzczych poezyj Tarasa Szewczenka. Lwów, 1914.pdf

Сторінки (клич до статусів сторінок)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96

Chto buw Taras Szewczenko	III
POEZIJA TARASA SZEWCZENKA	
1. Iwan Pidkowa	1
2. Do Osnowlanenka	3
3. Topola	7
4. Chustyna	14
5. Rozryta mohyla	16
6. Zapowit	18
7. Dumka syroty (Oj odna ja)	19
8. Weczir	20
9. Dumka (Sonce zachodyt)	20
10. Dumka dwiczny (Jakby meni czerewyky)	21
11. Dumka matery (Oj luli, luli)	22
12. Dumka (Oj czoho ty poczornilo)	23
13. Spiwanka (Litoptala slezczuku)	24
14. Pokynena dwiczyna (Na horodi)	25
15. Nad Dniprowuju sahoju	25
16. Najmyczka	27
17. Kateryna	46
18. Nowolnyk	70

Категорії (\*): Індекс - Не вичитано | Індеси | Індеси 1914-ого | Індеси книг | (+)  
Прихована категорія: Індеси українською мовою

Цю сторінку востаннє змінено: 23:56, 18 березня 2017.


Текст доступний на умовах ліцензії Creative Commons Attribution/Share-Alike, також можуть діяти додаткові умови. Детальніше див. Умови використання.

Політика конфіденційності Про Вікіджерела Відмова від відповідальності Розробники Кука Мобільний вигляд





# Робота над текстом. Сторінка:

 **ВІКІДЖЕРЕЛА**


- Головна сторінка
- Нові редагування
- Нові сторінки
- Автори
- Випадкова робота
- Випадковий автор
- Випадкове оцифрування



Участь


- Скринторій
- Довідка
- Пожертування

Інструменти





























- Посилання сюди
- Пов'язані редагування
- Завантажити файл
- Спеціальні сторінки
- Інформація про сторінку
- Статистика відвідувань
- Посилання за ID
- Підсторінки



























Мови 

Artem.komisarenko   Обговорення Налаштування Бета Список спостереження Внесок [Вийти](#)

< > Сторінка **Обговорення** зображення ^ Читати Редагувати Переглянути історію  Ще

## Редагування Сторінка:Kobzar'. Wybir dejakych najkraszczych poezyj Tarasa Szewczenka. Lwiw, 1914.pdf/5

**A A**                             

Заголовки  Формат                          

Заголовок (не включаться):

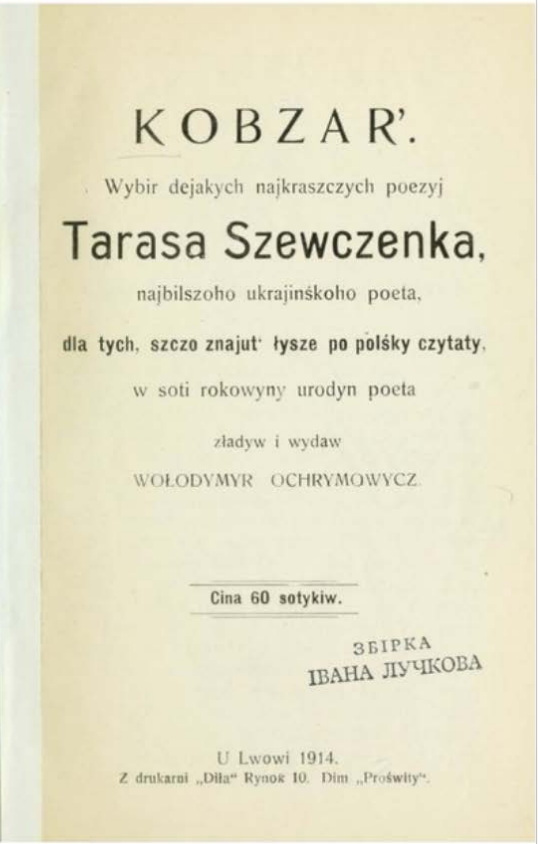
Типо сторінки (буде включатися):

```

{{c|[[Кобзар|{{xx-larger|{{uc|''Kobzar' .''}}}}]]}}
{{c|{{larger|Wybir dejakych najkraszczych poezyj}}}}
{{c|[[Автор:Тарас Шевченко|{{xx-larger|''Tarasa Szewczenka, ''}}]]}}
{{c|{{larger|najbilszoho ukrajins'koho poeta,}}}}
{{c|{{larger|''dla tych, szczo znajut' lysze po pol'sky czytaty, ''}}}}
{{c|{{larger|w soti rokowyny urodyu poeta}}}}
{{c|zładyw i wydaw}}
{{c|[[Автор:Володимир Охримович|{{uc|Wołodymyr Ochrymowycz.}}]]}}

```

Нижній колонтитул (не включаться):







# Робота над текстом. Результат:



ВІКІДЖЕРЕЛА

Головна сторінка  
Нові редагування  
Нові сторінки  
Автори  
Випадкова робота  
Випадковий автор  
Випадкове оцифрування

Опції

Вигляд II  
Сторінки показані  
Сторінки окремо

Участь

Скрипторій  
Довідка  
Пожертвування

Друк/експорт

Створення книги  
Завантажити як PDF  
Версія до друку

Інструменти

Посилання сюди  
Пов'язані редагування  
Завантажити файл  
Спеціальні сторінки  
Постійне посилання  
Інформація про сторінку  
Цитувати сторінку  
Посилання за ID  
Статистика відвідувань  
Підсторінки

Мови

Додати посилання

Artem.komisarenko Обговорення Напаштування Бета Список спостереження Внесок Вийти

Стаття Джерело Обговорення Читати Редагувати Переглянути історію Ще Пошук у Вікіджерелах

## Kobzar'

**Kobzar'** (1914)

*Taras Szewczenko, під ред. Wołodumyr Ochrymowycz*

Завантажити у

Цей текст написаний *латинкою на основі польської абетки*. • Існують інші версії цієї роботи: **Кобзар**

## KOBZAR'.

Wybir dejakych najkraszczyczych poezyj

**Tarasa Szewczenka,**

najbilszoho ukrajinszkoho poeta,

**dla tych, *szezo znajut' lysze po polsky czytaty,***

w soti rokowyny urodyn poeta

zładyw i wydaw

WOŁDUMYR OCHRYMOWYCZ.

**Cina 60 sotykiw.**

U Lwowi 1914.

Z drukarni „Dila” Rynek 10. Oim „Proświty”.

сторінка



WIKIMEDIA  
UKRAİNA



# Проблеми проекту:

1. Вичитка тексту вимагає великого обсягу роботи
2. Волонтерів як і завжди, як і всюди не вистачає
3. OCR допомагає, але частково:
  - a. скани часто поганої якості
  - b. різноманітні правописи ускладнюють перевірку текстів, пошук помилок OCR
4. Авторське право:
  - a. Розстріляне відродження, всі репресовані й вбиті взагалі, — під копірайтом
  - b. Використання orphan works та відумерлої спадщини — не унормоване
  - c. Покинуті радянські т. зв. службові твори (УРЕ, літописи, зібрання творів тощо)
  - d. Статус документів УНР, Гетьманату, ОУН, УПА невизначений, для держави існує тільки УРСР





# Переваги Вікіджерел:

- Як сховища:
  - стабільне: проект Фонду Вікімедіа, сестринський проект Вікіпедії
  - вільне: безкоштовне, без реклами
  - містке: умовно необмежений диск та трафік
  - надійне: розташоване за межами України, що забезпечує від долі ресурсів Криму та Донеччини
- Як ресурсу:
  - міжнародний, добре відомий проект
  - інтеграція з Вікіпедією → висока позиція у видачі пошуковиків
- Також:
  - текст краще індексується, краще шукається, простіше читається ніж традиційні PDF
  - можливість скористатись досвідом та технічними напрацюваннями більш розвинутих ком'юніті



## Недоліки Вікіджерел:

- Робота з ~~дилетантами~~ аматорами
- Відсутність формального та неформального контролю за матеріалами
- Обмежений вибір ліцензій, суворі вимоги щодо авторського права
- Вичитка тексту — працемістка робота







## На останок:

- Культурна спадщина — вразлива і в цифровому вигляді в т. ч.
- Не кладіть всі яйця в один кошик
- Якщо є готові майданчики — ними варто користуватись





**Дякую за увагу!  
Долучайтесь до проекту!**

**[uk.wikisource.org](http://uk.wikisource.org)**

**Користувач: artem.komisarenko**

Член Правління ГО «Вікімедіа Україна»,  
відповідальний за напрямок за напрямком БоГеМА  
(бібліотеки, галереї, музеї, архіви)

