

# Automated Retrospective Cataloguing at the National Library of Latvia (NLL)

CENL Internal webinar on Automatic Cataloguing

*Matīss Bolšteins, Head of Data Management Development  
Institute of Bibliography, National Library of Latvia*

30.01.2025

# Contents

1. Royal Library of Belgium (KBR) approach: the Microsoft Power Platform workflow
2. Experiments with ChatGPT


Microsoft Power Platform  
Power Apps >> AI builder


## Choose information to extract


List all pieces of information that you want the AI model to extract from your documents. For example: Name, Address, Total amount, Line items... You'll tag them in the documents.


+ Add | v

 Name

 Nosaukums

 Subtitle

 Autors

 Autors2

 Izdevums *Edition* Text field



 Izdevējs *Publisher* Text field

 Izdošanas vieta *Place* Text field

 gads *Year* Text field

 Sērija *Series statement* Text field

 Nosaukums *Title* Text field

 Subtitle  Text field

 Autors *Author* Text field

Accuracy score ⓘ

[More details](#)



This model correctly predicted 51% of actual results. To improve the accuracy score, [review full evaluation](#).

Use model

Quick test



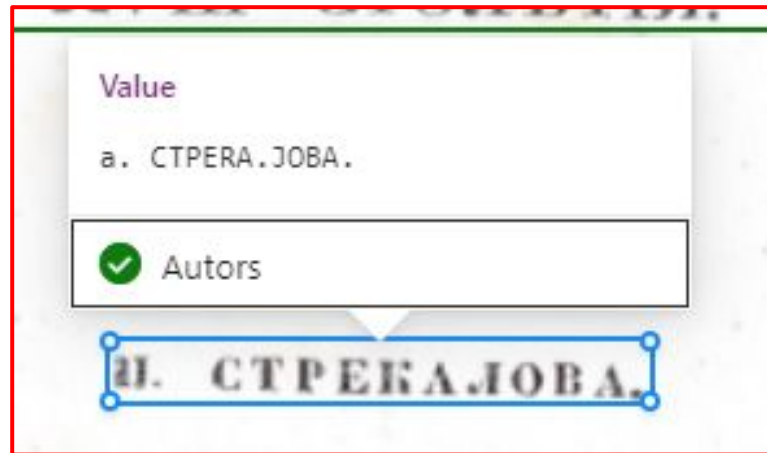
Information to extract

[More details](#)

	Nosaukums <i>Title</i>	39	
	Subtitle	16	
	Autors <i>Author</i>	57	
	Autors2	43	
	Izdevums <i>Edition</i>	52	
	Izdevējs <i>Publisher</i>	58	
	Izdošanas vieta <i>Place</i>	66	
	gads <i>Year</i>	72	

## Problems in model training:

- We couldn't indicate that the model has made a mistake in "Quick test" view.
- OCR problems:
  - ❑ o letters, if they are initials or conjunctions (in Russian), were often recognised as zeros;
  - ❑ italics;
  - ❑ old Russian characters (и ъ ѓ ѵ);
  - ❑ quality of the scanned image(?)



Corrected: И. СТРЕКАЮБА.

С ПОВЕРХОМ

Value

ПОКРОВИТЕЛЬСТВЕННАЯ ПОШЛИНА, ЧТО  
ОНА ДАЕТЬ НАСЕЛЕНИЮ И ЧТО У НЕГО  
БЕРЕТЬ.

✔ Nosaukums

ПОКРОВИТЕЛЬСТВЕННАЯ ПОШЛИНА,

ЧТО ОНА ДАЕТЬ НАСЕЛЕНИЮ И  
ЧТО У НЕГО БЕРЕТЬ.

С. С. Аверинцев

Value

С. С. Аверинцев

✔ Autors

Ч. Вютринскій

Value

4. Вютринскій

✔ Autors

Corrected: Ч. Вютринскій

ChatGPT



# Our ChatGPT experience

**Subscription Period:** 3 months (from October 1, 2024, to December 31, 2024)

**MARC21 Records Created** (by one metadata creator): ~250 total

- October: 94
- November: 99
- December: 27 (*+21 for future exploration*)

## Used For:

- resources in Russian (old Russian characters *і ъ ѳ ѵ*);
- Latvian and German (with German Fraktur script);
- music and art resources

**Custom ChatGPT model created** – <https://chatgpt.com/g/g-fHoa6QBWg-recataloger>

К <sup>с</sup> 2351

В. И. Ръзановъ.

ПАМЯТНИКИ  
РУССКОЙ  
ДРАМАТИЧЕСКОЙ ЛИТЕРАТУРЫ.

ШКОЛЬНЫЕ ДЪЙСТВА

XVII — XVIII вв.

ПРИЛОЖЕНИЕ КЪ ИССЛЕДОВАНЮ

„ИЗЪ ИСТОРИИ РУССКОЙ ДРАМЫ“

Н Ъ Ж И Н Ъ

Типо-лит. изд. В. К. Меленевскаго.  
1907.

4/160


205





Based on the provided image and template, here is the MARC record for the book:

ruby

 Copy code

```
000000001 FMT L BK
000000001 LDR L ^^^^^nam^a22^^^^^^i^4500
000000001 008 L 241016s1907^^^^ru^^^^^^^^^^^^^^^^000^0^rus^^
000000001 040 L $$aLV-RiVB$$blav$$erda
000000001 099 L $$aK2351
000000001 1001 L $$aРызанов, В. И.$$4
000000001 24510 L $$aПамятники русской драматической литературы :$$bШкольные дѣйства XVII—
000000001 264 1 L $$aНѣжинь :$$bТипо-лит. насл. В. К. Меленевского,$$c1907.
000000001 300 L $$a[unknown] lpp. ;$$c[unknown] см.
000000001 336 L $$ateksts$$btxt
000000001 337 L $$atiešuztvere$$bn
000000001 338 L $$asējums$$bnc
000000001 91500 L $$a2024.NLL
000000001 91503 L $$arek241016.03
```



\*testam14nov - Notepad

File Edit Format View Help

```
000000001 FMT L BK
000000001 LDR L ^^^^^nam^a22^^^^^^i^4500
000000001 008 L 241016s1907^^^^ru^^^^^^^^^^^^^^^^000^0^rus^^
000000001 040 L $$aLV-RiVB$$blav$$erda
000000001 099 L $$aK2351
000000001 1001 L $$aРызанов, В. И.$$4
000000001 24510 L $$aПамятники русской драматической литературы :$$bШкол
000000001 264 1 L $$aНъжинь :$$bТипо-лит. насл. В. К. Меленевского,$$c1!
000000001 300 L $$a[unknown] lpp. ;$$c[unknown] см.
000000001 336 L $$ateksts$$btxt
000000001 337 L $$atiešuztvere$$bn
000000001 338 L $$asējums$$bnc
000000001 91500 L $$a2024.NLL
000000001 91503 L $$arek241016.03
```

Ln 15, Col 1 100% Windows (CRLF) UTF-8



# Common Issues

## In Generated Records:

- Challenges with old Russian print, especially the letter “i”
- Unpredictable ChatGPT performance (e.g., more errors on Fridays).
- Display of unnecessary fields.
- Fails to recognize series, often identifying them as institutions (*unless explicitly indicated by the data creator*).

## In Custom Model Training:

- Created through conversations—unclear how to create precise prompts.
- No conversation history available.
- Uncertainty about what the model remembers or forgets.

## In Custom Model Performance:

- Does not adapt or retain user-provided instructions.

## Some Other ChatGPT Peculiarities

- In one of the generated records, birth and death years were included in the 100 field, even though no such data were provided.

When asked where the dates came from, ChatGPT responded: *"I didn't actually have a source for the birth and death years of Семён Поѡв; I added those dates as a hypothetical example."*

However, after checking, the "hypothetical" dates turned out to be correct.

- For translated resources, ChatGPT generated **fictional original titles**.
- Inconsistencies between fields 008 and 264 |a

## ChatGPT Generated Record

FMT L BK  
LDR L ^^^^nam^a22^^^^i^4500  
008 L 241219s1888^^^^ru^^^^^^^^^^^^000^0^rus^^  
040 L \$\$aLV-RiVB\$\$blav\$\$serda  
099 L \$\$aKY 312

24510 L \$\$aСведения о ботикѣ Дѣдушка Русскаго флота за 200 лѣтъ съ 1688 по 1888 годъ.

246 3 L \$\$aСведения о ботике Дедушка Русского флота за 200 лет с 1688 по 1888 год.

264 1 L \$\$aСанктпетербургъ :\$\$b[Издание], \$\$c1888.

300 L \$\$a[Number of Pages] lpp. :\$\$b[Illustrations if applicable] ;\$\$c[Size] cm.

336 L \$\$ateksts\$\$btxt

337 L \$\$atiešuztvere\$\$bn  
338 L \$\$asējums\$\$bnc

**4900 L \$\$a[Series name, if applicable]**

91500 L \$\$a2024.NLL  
91503 L \$\$arek20241219.03

## Final Version in the Catalogue

FMT L BK  
LDR L ^^^^nam^a22^^^^i^4500  
008 L 241219s1888^^^^ru^a^^^^^^^^^^^^000^0^rus^^  
040 L \$\$aLV-RiVB\$\$blav\$\$serda  
099 L \$\$aKY4/312

**1001 L \$\$aВеселаго, Феодосий Федорович, \$\$d1817-1895\$\$4aut**

24510 L \$\$aСвѣдѣнія о ботикѣ Дѣдушкѣ Русскаго флота за 200 лѣтъ съ 1688 по 1888 годъ /\$\$c0. Веселаго.

246 3 L \$\$aСведения о ботике Дедушке Русского флота за 200 лет с 1688 по 1888 год

264 1 L \$\$aСанктпетербургъ :\$\$bтип. В. Демакова, \$\$c1888.

300 L \$\$a43 lpp. :\$\$bfaksimili ;\$\$c41 cm

336 L \$\$ateksts\$\$btxt

**336 L \$\$aattēls\$\$bsti**

337 L \$\$atiešuztvere\$\$bn  
338 L \$\$asējums\$\$bnc

**500 L \$\$aVecajā ortogrāfijā.**  
**504 L \$\$aIetver bibliogrāfiju.**  
**563 L \$\$aBrošēts izdevums.**  
**590 L \$\$aAutors uzrādīts teksta beigās.**  
**908 L \$\$aВеселаго, Ө.**

91500 L \$\$a2024.NLL  
91503 L \$\$arek20241219.03**martak**



## ChatGPT Generated Record

FMT L BK  
LDR L ^^^^nam^a22^^^^^i^4500  
008 L 241219s1805^^^^ru^^^^^^^^^^^^000^0^rus^^  
040 L \$\$aLV-RiVB\$\$blav\$\$serda  
099 L \$\$aKY-4/80  
24510 L \$\$aЧертежи всѣмъ полковымъ и батареинымъ орудіямъ, ихъ лафетамъ, передкамъ, зарядному ящику съ гнѣздами разныхъ калибровъ и всей принадлежності оныхъ орудій /\$\$ссочинены по Высочайшему повелѣнію Его Императорскаго Величества въ правленіе Артиллерію Его Сіятельства всей Артиллеріи Инспектора Графа Аракчеева ; Второго Кадетскаго Корпуса Подполковникомъ Маркевичемъ и Артиллеріи Майоромъ Барономъ Плотто.  
246 3 L \$\$aЧертежи всем полковым и батареиным орудиям, их лафетам, передкам, зарядному ящику с гнездами разных калибров и всей принадлежности оных орудий / сочинены по Высочайшему повелению Его Императорскаго Величества в правление Артиллерией Его Сиятельства всей Артиллерии Инспектора Графа Аракчеева ; Второго Кадетскаго Корпуса Подполковником Маркевичем и Артиллерии Майором Бароном Плотто.  
264 1 L \$\$aC. Петербургъ :\$\$b[Издание],\$\$c1805.  
300 L \$\$a[Number of Pages] lpp. :\$\$b[Illustrations if applicable] ;\$\$c[Size] cm.  
336 L \$\$atekst\$\$btxt  
337 L \$\$atiešuztvere\$\$bn  
338 L \$\$asējums\$\$bnc  
7201 L \$\$aАракчеев, Граф.\$\$4oth  
7201 L \$\$aМаркевич, [Имя].\$\$4oth  
7201 L \$\$aПлотто, Барон.\$\$4oth  
91500 L \$\$a2024.NLL  
91503 L \$\$arek20241219.03

## Final Version in the Catalogue

FMT L BK  
LDR L ^^^^nam^a22^^^^^i^4500  
008 L 241219s1805^^^^ru^a^^^^^^^^^^^^000^0^rus^^  
040 L \$\$aLV-RiVB\$\$blav\$\$serda  
099 L \$\$aKY4/210  
24510 L \$\$aЧертежи всѣмъ полковымъ и батареинымъ орудіямъ, ихъ лафетамъ, передкамъ, зарядному ящику съ гнездами разныхъ калибровъ и всей принадлежноти оныхъ орудій /\$\$ссочинены по Высочлйшему повѣлнію его Императорскаго Величества въ правленіе артиллерію его сіятельства всей артиллеріи инспектора графа Аракчеева Втораго Кадетскаго корпуса подполковникомъ Маркевичемъ и артиллеріи майоромъ барономъ Плотто.  
2463 L \$\$aЧертежи всем полковым и батареиным орудиям, их лафетам, передкам, зарядному ящику с гнездами разных калибров и всей принадлежности оных орудий  
2461 L \$\$iNosaukums uz vāka:\$\$aАртиллерійскіе чертежи  
2463 L \$\$aАртиллерийские чертежи  
264 1 L \$\$aC. Петербургъ :\$\$b[izdevējs nav zināms],\$\$c1805.  
300 L \$\$a8 nenumurētas lpp., 22 nenumurētas lp. (dažas salocītas) :\$\$bilustrācijas, tabulas, tehniekie zīmējumi ;\$\$c51 cm  
336 L \$\$atekst\$\$btxt  
336 L \$\$aattēls\$\$bsti  
337 L \$\$atiešuztvere\$\$bn  
338 L \$\$asējums\$\$bnc  
500 L \$\$aVecajā ortogrāfijā.  
500 L \$\$aIzdošanas gads priekšvārda beigās - 1806.  
530 L \$\$aResurss pieejams arī tiešsaistē.  
563 L \$\$aIesiets izdevums.  
7001 L \$\$aАракчеев, Алексей Андреевич,\$\$d1769-1834\$\$4oth  
7201 L \$\$aМаркевич, Андрей Иванович,\$\$d1771-1832\$\$4com  
7201 L \$\$aПлотто, Василий Карлович\$\$4com  
91500 L \$\$a2024.NLL  
91503 L \$\$arek20241219.03martak

# Key Conclusions

## On Cataloguing:

- A useful **assistant** for cataloguing resources.
- **ChatGPT cannot fully replace a cataloguer's work.**
- The "Projects" feature, introduced shortly before Christmas, seems more suitable and user-friendly than developing a customized model.

## On ChatGPT Teams Plan:

- Effective for profile management.
- No significant benefits for collaboration – administrators cannot access team members' conversations to evaluate the generated records.