

Institute of History, Polish Academy of Sciences  
Marek Słoń, Mikołaj Słoń, Adam Zapala

# Indexing of Tax Registers from 16th Century

TUC, Wien, 8-9.11.2018

# Historical Atlas of Poland - history of the project

1880 - general idea: reconstruction of the settlement network and administrative structures in the 16th cent.

1953 - founding of the atlas team and start of the series

2020 - project and series completion





|   |                         |          |
|---|-------------------------|----------|
| Miedzichod                                      | Manfis Cmt' 3½ per      | 30.      |
| Manfis defecto                                  |                         | 10       |
| Ab Hortulamis nymloeskowicz 2                   | per                     | 7        |
| Ventili Annuerfario                             |                         | 5 + 8    |
| Kadzín  | Manfis Cmtom' 3         | per 30   |
| Ab Hortulamis nymloeskowicz 4                   | per                     | 7. 3 10  |
| Trambinko                                       | Manfis Cmt' 4½ per      | 20. 4 15 |
| Lubiatowo idr                                   | Ex sorte x. Erasmi Lin: |          |
| biatowsky & medys Manfis 3                      | per                     | 10.      |
| Ab Hortulamis nymloeskowicz 2                   | per                     | 4.       |
| Ab Fratru Colom                                 |                         | 30       |
| Ventili Annuerfario                             |                         | 5 2 28   |
| Lubiatowo ndr                                   | Ex sorte x. Hedinger de |          |
| Miskono wilkowsky & Manfis vno                  |                         | 30       |
| Ab Hortulamis nymloeskowicz 3                   | per                     | 7.       |
| Ab Fratru Colom                                 |                         | 30       |
| Ventili Annuerfario                             |                         | 5 2 17   |
| Item Ex sorte x. Mathie Linbiatowsky choczyski: |                         |          |
| sky diti. Ab Fratru Colom                       |                         | 30       |
| Ab Hortulano vno nymloeskowicz                  |                         | 4. 1 4   |
| Ostrowieczno idr                                | Ex sorte x. Pe:         |          |
| bi Garzeczky Ab Fratru Curie                    |                         | 15       |
| Ab Hortulamis nymloeskowicz 2                   | +                       | 4.       |
| Ab Inguilino vno paupero                        |                         | 2.       |
| Item Ex sorte x. Georgy Kulegowsky & medio      |                         | 25       |
| Manfo vno                                       |                         | 15       |
| Ab Hortulano nymloeskowicz                      |                         | 4.       |
| Ostrowieczno ndr                                | Manfis scull            | 19       |
| Trubno Cmt' Manfo vno                           | per                     | 30.      |
| Manfis defecto 2                                | per                     | 10.      |
| Ab Inguilino paupero 3                          | per                     | 2. 5 5   |

79 18 3 41  
Manfi. Defect. Anaxial. Curie. Hort. nym. Hort. libe. Anaxial.

25 18

[illegible][illegible]

Central Archives of Historical  
Records, Archives of State  
Treasury, division I, vol. 6, p. 484v,  
vol. 13, p. 139v, vol. 11, p. 1516

# Data from tax registers

The **character** of settlement, its name, taxed categories

|   |       |   |    |
|---|-------|---|----|
| Villa Bronikowo. De sorte nobilis Alberti Bronikowski |       |   |    |
| X tribus mansis agri per . . . . .                    | gr 30 |   |    |
| X rator a duobus Aratriis per . . . . .               | gr 30 |   |    |
| X tribus hortulanis agri no habent per . . . . .      | gr 4  |   |    |
| X molendino ventili deserto de quo nihil solt         |       |   |    |
| Opilio a triginta quinqs ouibus . . . . .             | gr 7  | 5 | 19 |
| Scti solt die 4 maij                                  |       |   |    |



# Data from tax registers

The character of settlement, its **name**, taxed categories

|   |       |      |
|---|-------|------|
| Villa <b>Bronikowo</b> De sorte nobilis Alberti Bronikowski |       |      |
| X tribus mansis agri per . . . . .                          | gr 30 |      |
| X rator a duobus Aratris per . . . . .                      | gr 30 |      |
| X tribus hortulanis agri no habent per . . . . .            | gr 4  |      |
| X molendino ventili deserto de quo nihil solt               |       |      |
| Opilio a triginta quinqs ouibus . . . . .                   | gr 7  |      |
| Sic solt die 4 maij   |       |      |
|   |       | 5 19 |

# Data from tax registers

The character of settlement, its name, **taxed categories**

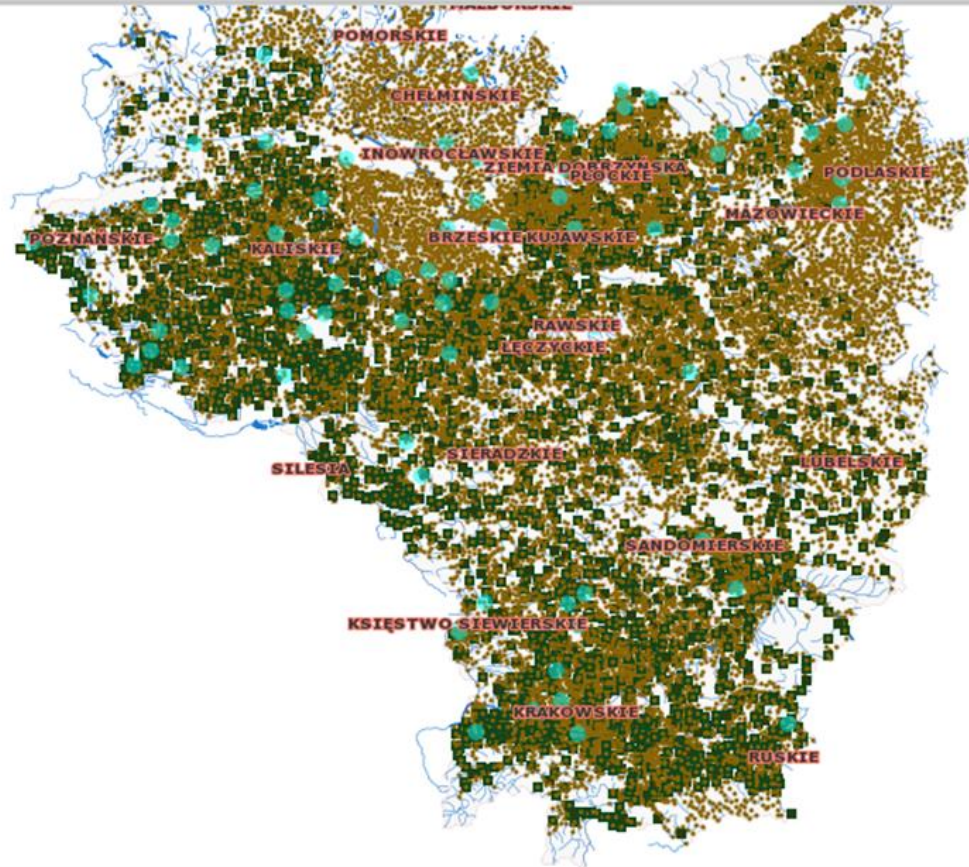
Villa Bronikowo. De sorte nobilis Alberti Bronikowski

|  |       |
|--|-------|
| X tribus mansis agri per . . . . .               | gr 30 |
| X rator a duobus Aratriis per . . . . .          | gr 30 |
| X tribus hortulanis agri no habent per . . . . . | gr 4  |
| X molendino ventili deserto de quo nihil solt    |       |
| Opilio a triginta quinqs ouibus . . . . .        | gr 7  |

Sete solt die 4 maij

5 19

Search for



Map

GIS  
Database

Footnotes



# Atlas

Tabela haseł Tabela zapisek Skany rękopisu Mapa

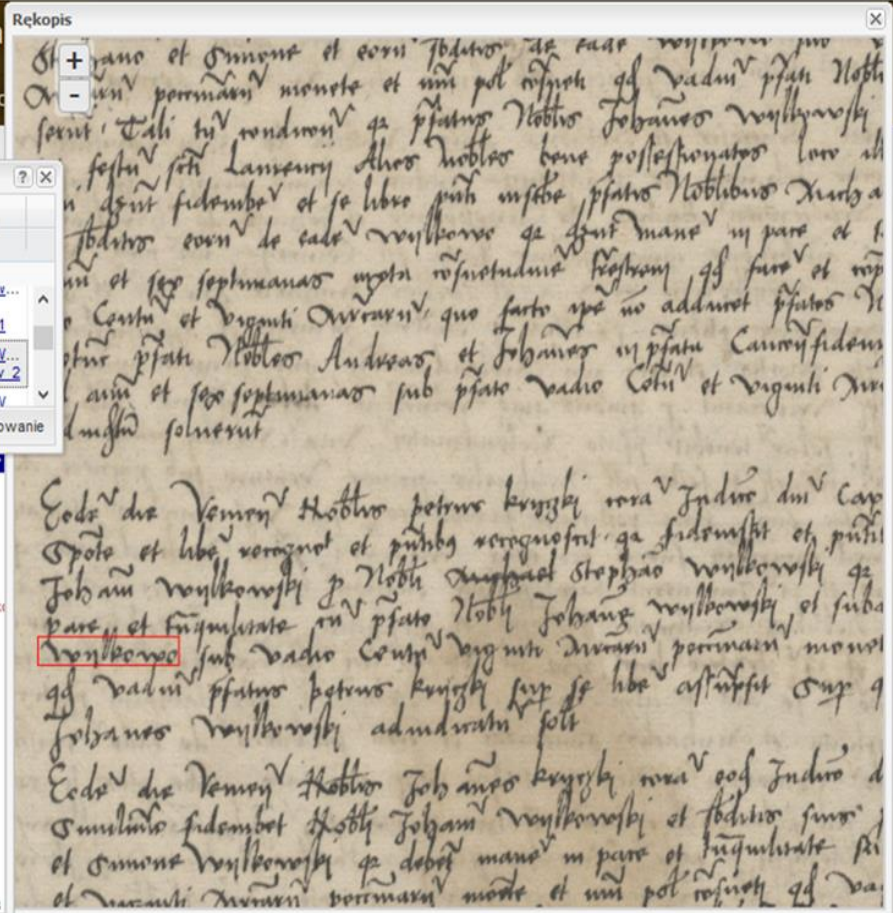
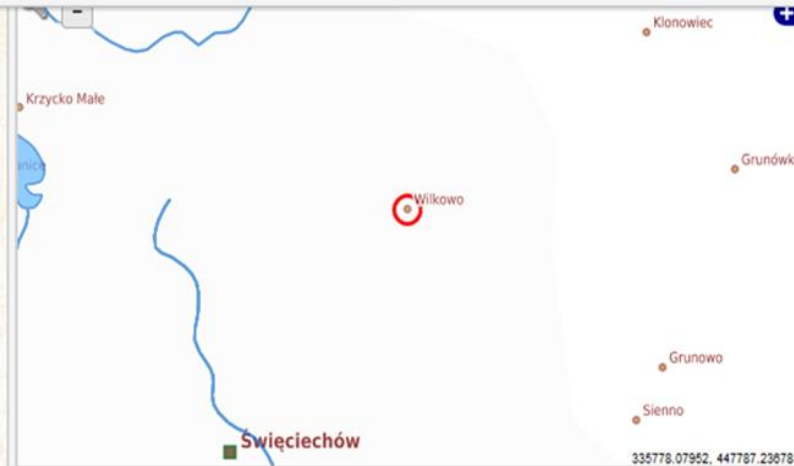
## Wschowska księga gro...

Kliknięcie powiększa (z klawiszem Ctrl tylko przesuwa) skan rękopisu

| Id  | Osoba - źródło                   | Osoba - informacje | Osoba                       | Osoba - Nazwisko     | Osada   | Osada - źródło | Hasło rzecz... | Uwagi | SGH | Zapiska                  |
|-----|----------------------------------|--------------------|-----------------------------|----------------------|---------|----------------|----------------|-------|-----|--------------------------|
| 1.3 | nub. Casper Rotemburk Zbarzewski |                    | Casper Rotenberg Zbarzew... | Kutemurzy Zbarzew... |         |                |                |       |     | APPell W... Gr. 1 100v 1 |
| 8.1 |                                  |                    |                             | Wilkowo              | Wilkowo |                |                |       |     | APPzn W... Gr. 1 73v 2   |
| 9.1 | honest                           | maż                | ładwina                     | Ilklich              |         |                | nazwisko       |       |     | APPzn W                  |

Page 1 of 103

Usun filtrowanie



c.a. 25 000 settlements

c.a. 70 000 pages of manuscripts

hundreds of different hands and writing styles

c.a. 250 000 headlines with settlements names

# Material used in the model

| Paro: Zagorcz Regalis         |   |    |  |
|-------------------------------|---|----|--|
| od Zagrodnikow 10 na rolinach | 1 | 10 |  |
| od komornikow 4               |   | 8  |  |
| od Ribitwa 1.                 |   | 8  |  |
| Skarżym dorozym przewoznej    |   | 12 |  |
| Zemlinn od keta 1.            |   | 12 |  |

| Gorky                            |   |    |   |
|----------------------------------|---|----|---|
| Paro kysicel pazanowks Zlanen 7. | 5 | 5  | 5 |
| od komornikow 2. Zbielen         |   | 8  |   |
| od komornikow 3. przez bielen    |   | 6  |   |
| Skarżym przewoznej               |   | 12 |   |

| Kartki vtrague. Sous dmi Zarduchky    |   |     |  |
|---------------------------------------|---|-----|--|
| ab fortulans ano in agro              | 4 |     |  |
| A. O. Andree od 3. mofsi de fis vniy3 | 3 | 10. |  |
| aihts Zarduchky de lancis gving8.     |   | 12  |  |
| Zarduchky 3. in agro                  |   |     |  |

| Sulkowycze Magnifiki domi        |   |   |  |
|----------------------------------|---|---|--|
| Andree Zbornuski A. Colong 12 in | 3 |   |  |
| tab. 3.                          |   |   |  |
| Jugutinski al penta 1.           |   | 8 |  |
| Jugutinski panga 2.              |   | 4 |  |

| Skotmiky parua Gusi domi      |   |    |   |
|-------------------------------|---|----|---|
| Nicolaus Lubomski A. Colong 4 | 2 |    |   |
| in landis 2.                  |   |    |   |
| Hortulan 3 in agro            |   | 18 |   |
| Taberna Condurub              |   | 12 |   |
| De Agro emed ranga            |   | 2  | 9 |
| Artificis 2.                  |   | 8  |   |
| Jugutinski 5. panga           |   | 10 |   |
| Propmans GemaV                | 1 | 8  |   |
| Molunda 3. Rotarum 3. Alia 1  |   | 8  |   |
| Skomp 3.                      | 1 | 15 |   |

| Zakrzow x Nicolaus Hfold                  |   |       |  |
|---|---|-------|--|
| Sors emist de Lancis 1 1/2                |   |       |  |
| ide ex medio lanceo deserto               | 1 |       |  |
| Ex sorte Stai et Petri omieginisk         |   | 5     |  |
| de medio lanceo 3 media quanta ag         |   | 1 1/2 |  |
| Sors joannis Lesmionski de medio lanceo 3 |   | 1 1/2 |  |
| media quanta                              |   |       |  |

| Kostki maior                                  |  |   |  |
|---|--|---|--|
| Ex sorte domi Stai Zbikalski de medio ppris 6 |  |   |  |
| Ex sorte 8 misfachs Glosigki                  |  | 6 |  |
| de medio lan ppris Cult                       |  |   |  |
| Ex 8 Adami pazanuski                          |  | 3 |  |
| de ppris Cult                                 |  |   |  |
| Ex sorte 8 Georgi Glosigki                    |  | 6 |  |
| de medio lan ppris Cult                       |  |   |  |
| ab fortulans 2 in fery                        |  | 4 |  |
| Sum. 0/25                                     |  |   |  |



# Characteristics

Changing of the hands in one registers volume

Variation of the language (polish, latin, polish/latin)

Irregular abbreviations

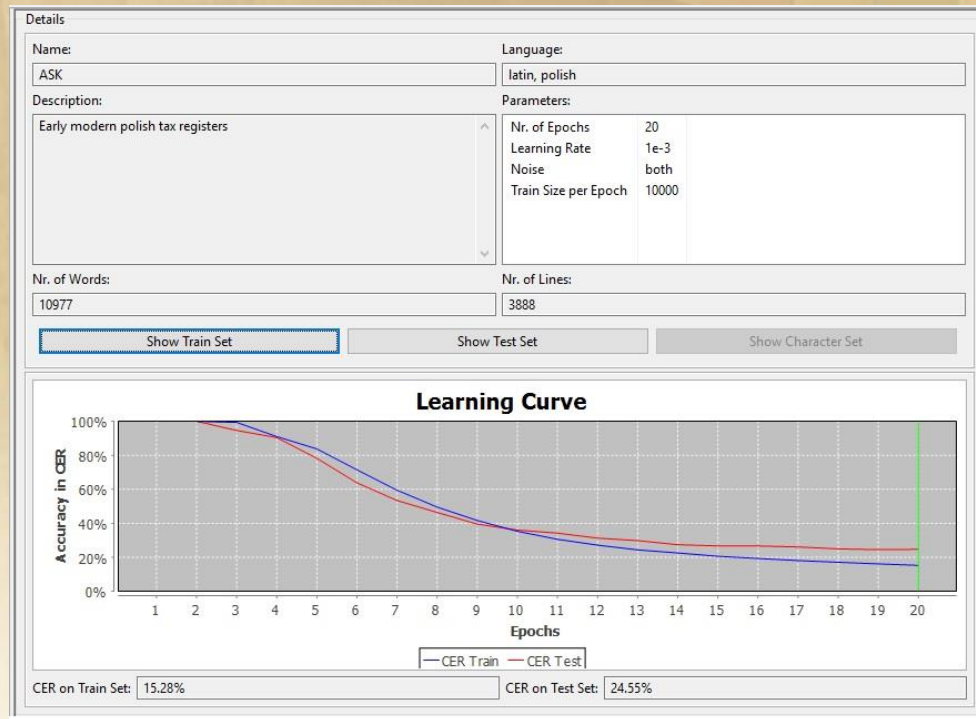
Different styles of writing in the headlines and in the main text

# HTR(CITlab) Model

Nr. of Words: 10 977

CER on Train Set: 15.28%

CER on Test Set: 24.55%

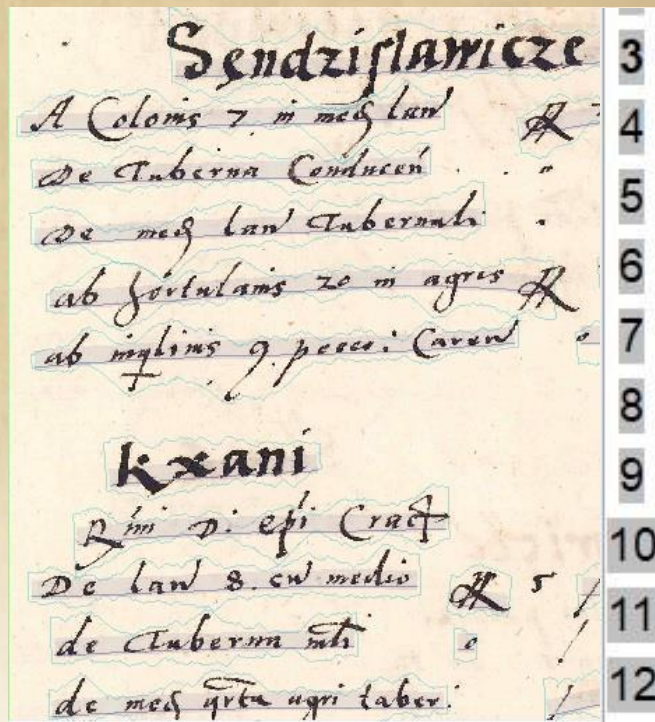


# An example of the automatic text recognition

Proper recognition of frequently used words and abbreviations

Worse recognition of rarely used words and abbreviations

Very bad recognition of headlines written with different style.



Sendzicawicze Regalis

A Colonis 7 in medo lan

De Taberna Conducen

De med lan Taberali

b hortulanis 20 in agri

ab inqulinis peccco Curen

Keai

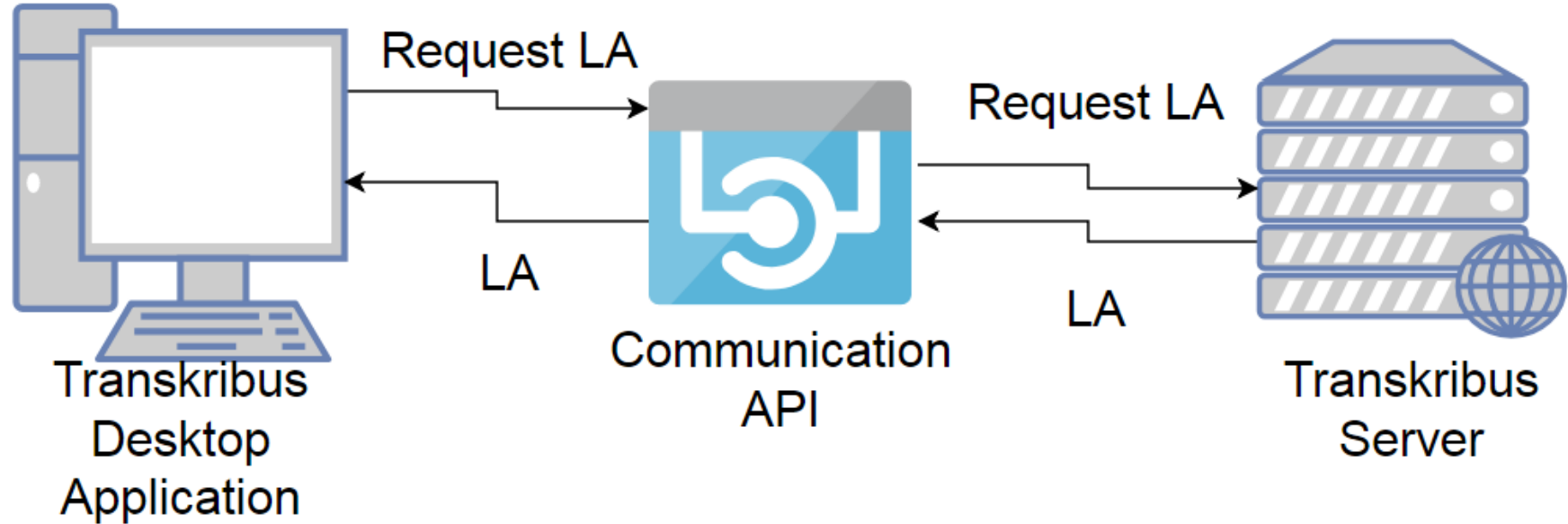
Reisii: psopii Crac

De lan 8 cum medio

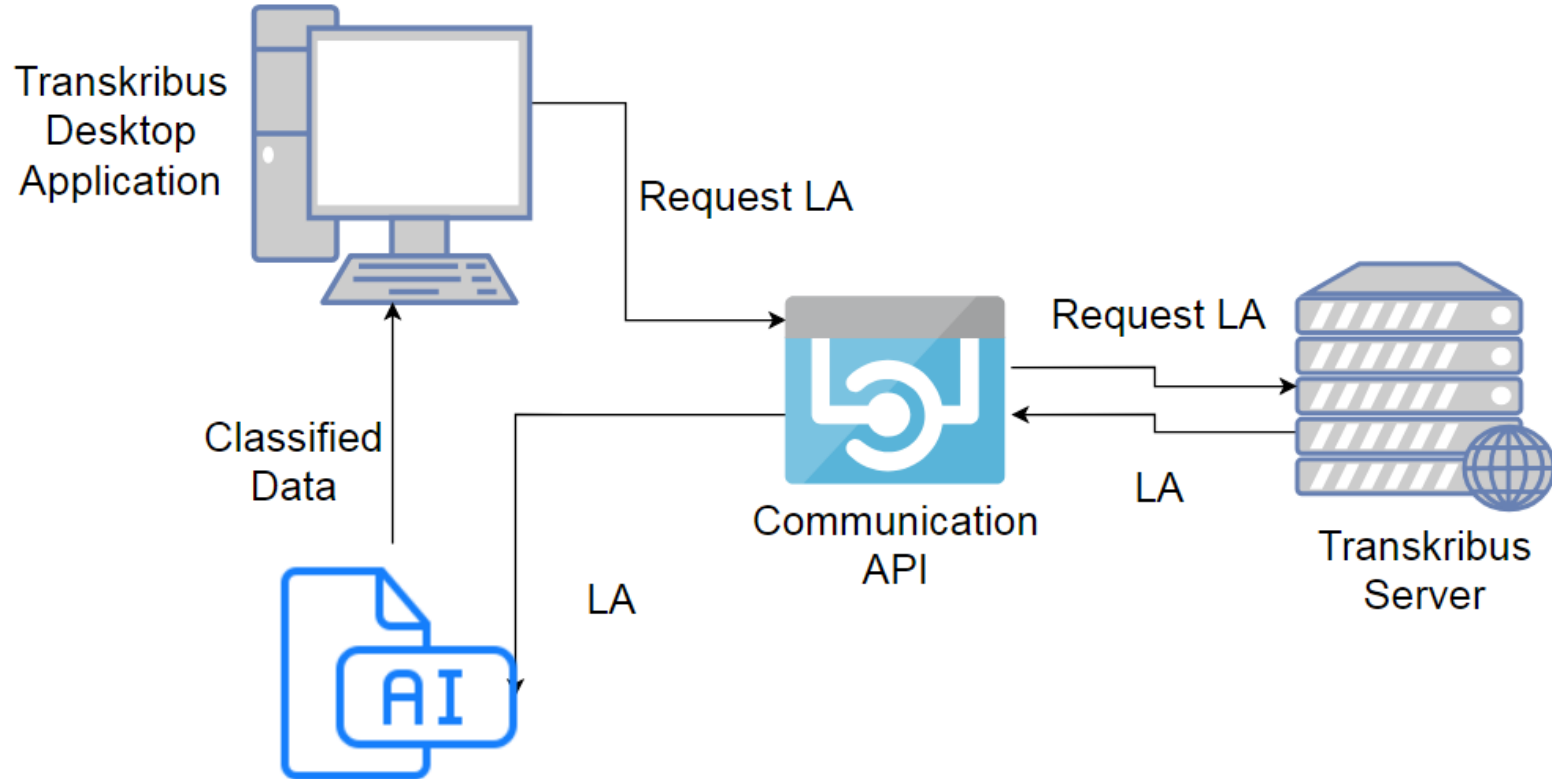
de Taberna nul



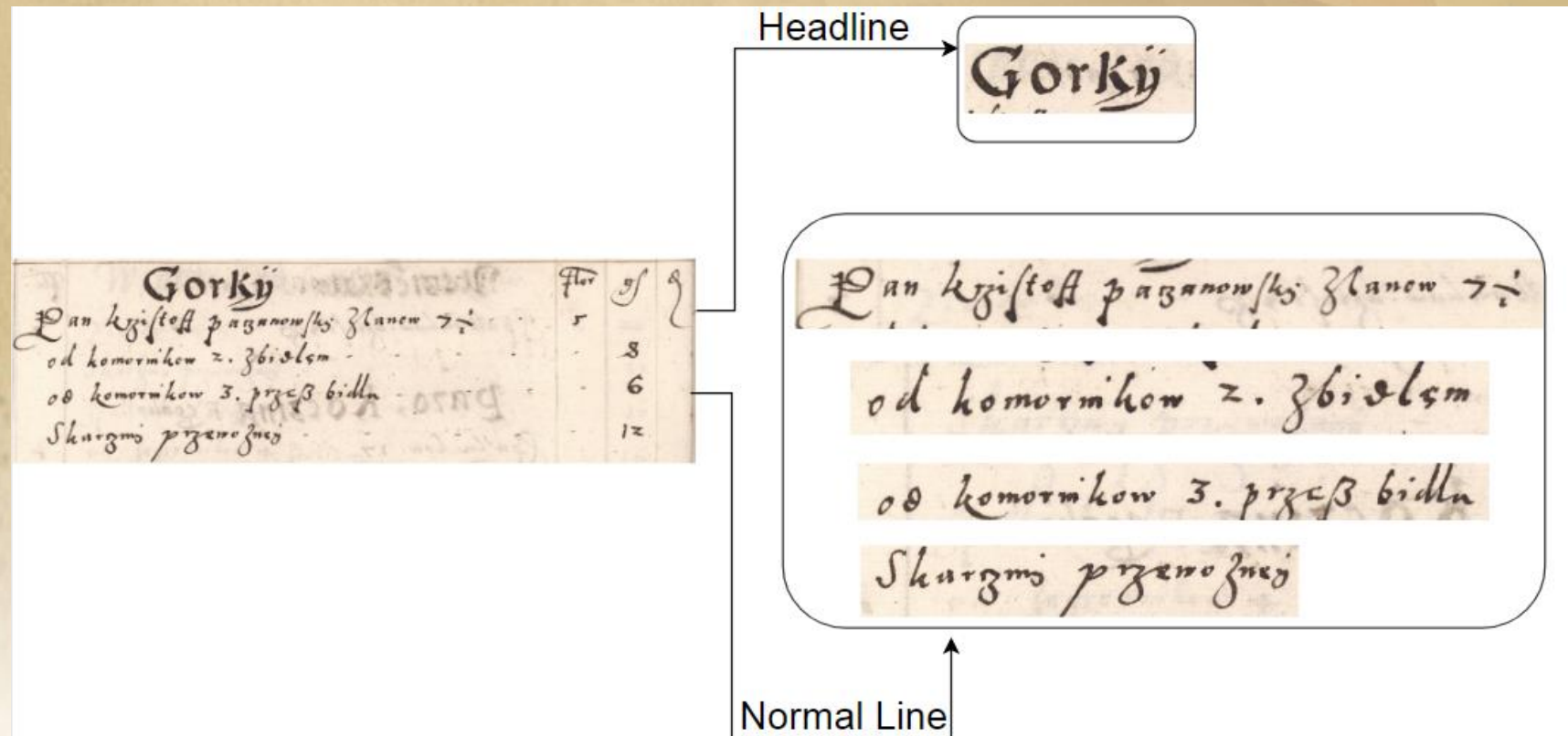
# Classification tool for headlines



# Classification tool for headlines



# Classification tool for headlines





# Tool selection - Convolutional Neural Network

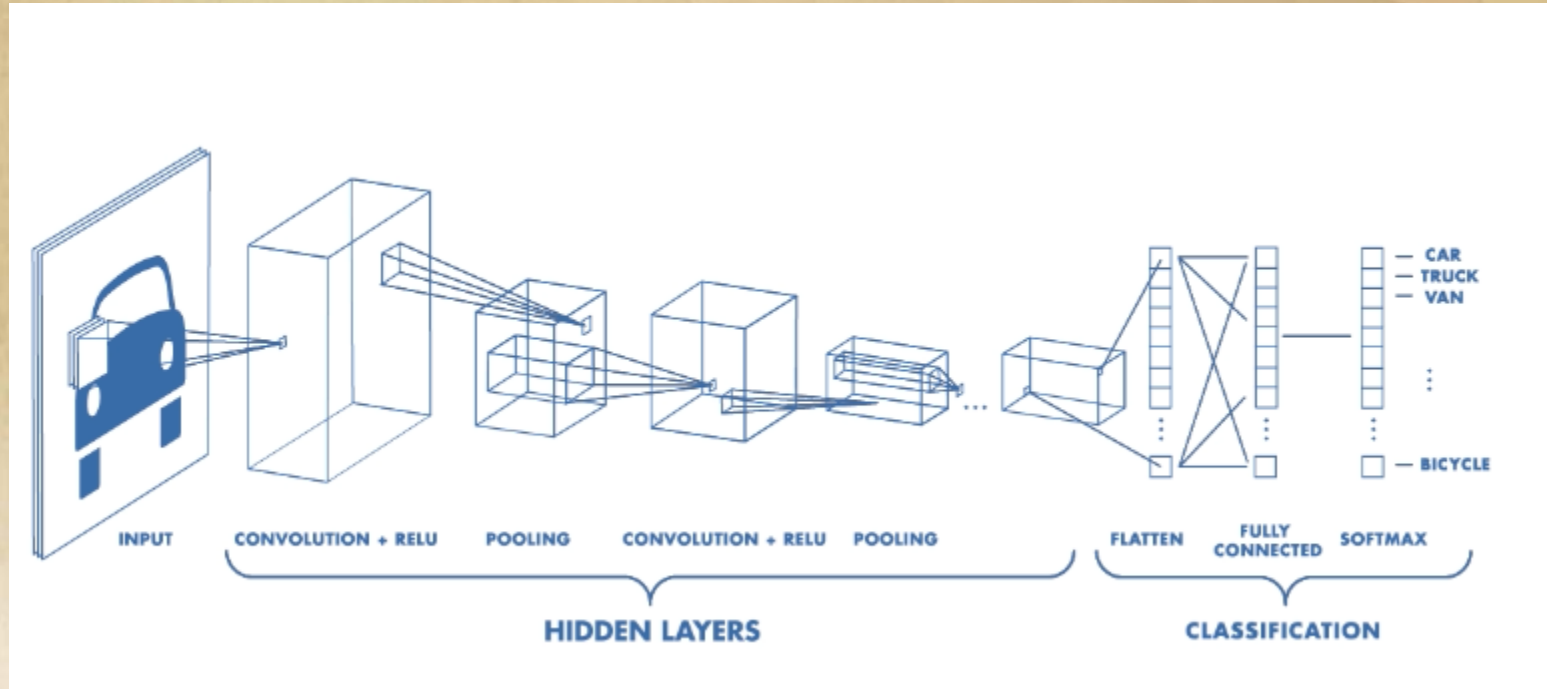


Image source: <https://medium.freecodecamp.org>

## Challenges - Image normalization

Gorky

Pan Krzysztof Paganowski znanow 7 1/2

Skargom przemożnej

Institute of History, Polish Academy of Sciences  
Marek Słoń, Mikołaj Słoń, Adam Zapala

## Indexing of Tax Registers from 16th Century

[www.ihpan.edu.pl](http://www.ihpan.edu.pl)

[www.atlasfontium.pl](http://www.atlasfontium.pl)

TUC, Wien, 8-9.11.2018