UNIWERSYTET EKONOMICZNY W KRAKOWIE

THE ANALYSIS OF THE IMPACT OF RESEARCH ACTIVITY ON SOCIETY AND ECONOMY WITH THE USE OF TOPIC MODELLING

PAWEŁ LULA, MAGDALENA TALAGA

KRAKÓW UNIVERSITY OF ECONOMICS, POLAND

IT FOR PRACTICE 2022, VSB - TECHNICAL UNIVERSITY OF OSTRAVA, OCTOBER 13TH 2022

The outline

- Evaluation of the quality of research activity in Poland
- Latent Dirichlet Allocation and Correlated Topic Modelling as potential tools for topic identification
- The analysis of main topics identified in description of cases of impact
- Conclusions

Evaluation of the quality of scientific activity at Polish universities

Historical remarks

- 1991 – first evaluation of the quality of scientific activity
- …
- 2022 – the first evaluation of the quality of scientific activity according to the new regulations (2018), covers the period 2017-2021

Evaluation criteria

1. scientific or artistic level of the conducted activity (publications)
   a. monographs
   b. peer-reviewed papers published in journals listed in the international databases
   c. peer-reviewed materials from international conferences

2. financial results of scientific research and development work
   a. projects financed through a competitive procedure by EU, EFTA or other international organizations,
   b. projects financed by NCN, NCBiR, Foundation for Polish Science
   c. commercialization of research results,
   d. research services for non-HEI

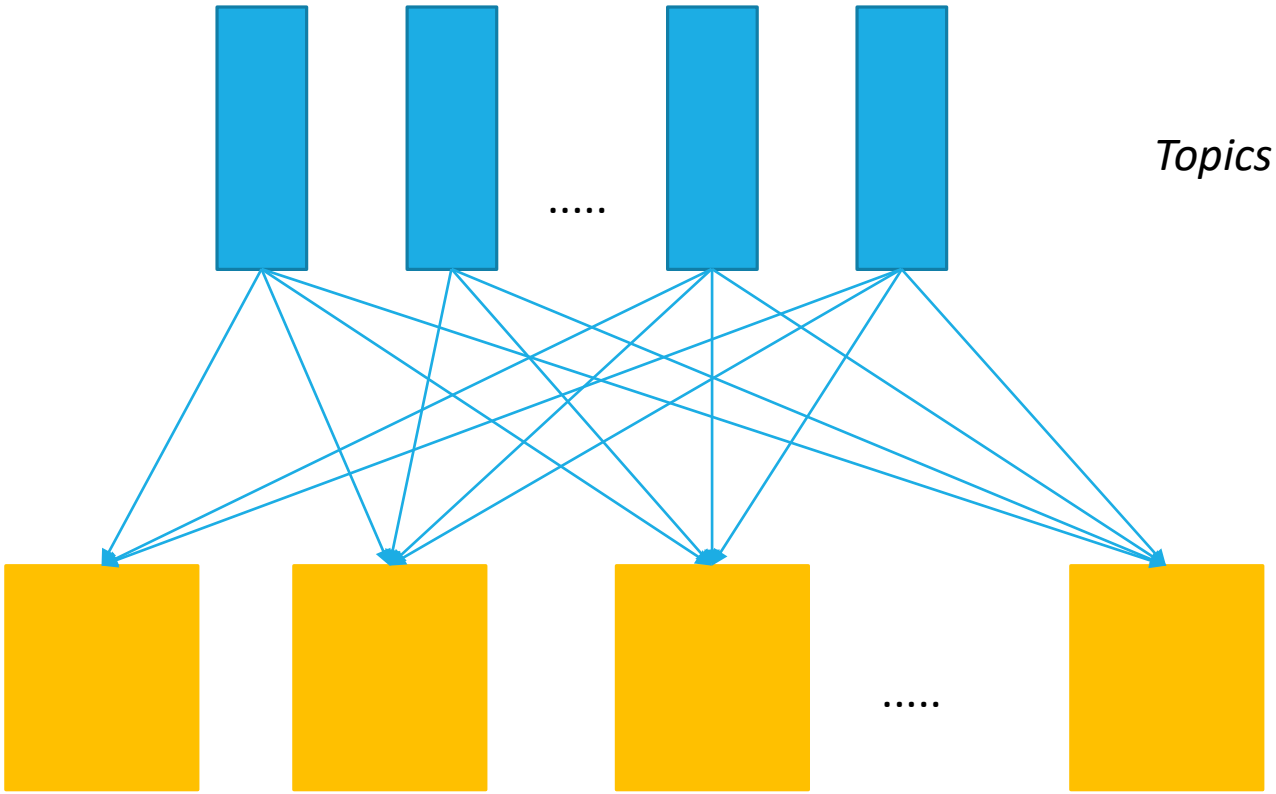
3. impact of scientific activity on:
   a. economy,
   b. health care system,
   c. functioning of public administration,
   d. culture,
   e. the arts,
   f. environmental protection,
   g. national security and defense

Evaluation process

- Reporting by universities
- Analysis by the Commission for the evaluation of Science (KEN)
- Granting of scientific categories by the Minister – based on the resolution of the KEN (A+, A, B+, B, C). The category achieved has an impact on:
  - the right to create and conduct study programs
  - the right to run doctoral schools
  - the right to confer academic degrees,
  - the amount of subsidy received from the state budget

Topic modelling as a tool for identification of main issues mentioned in cases of impact

Topics models

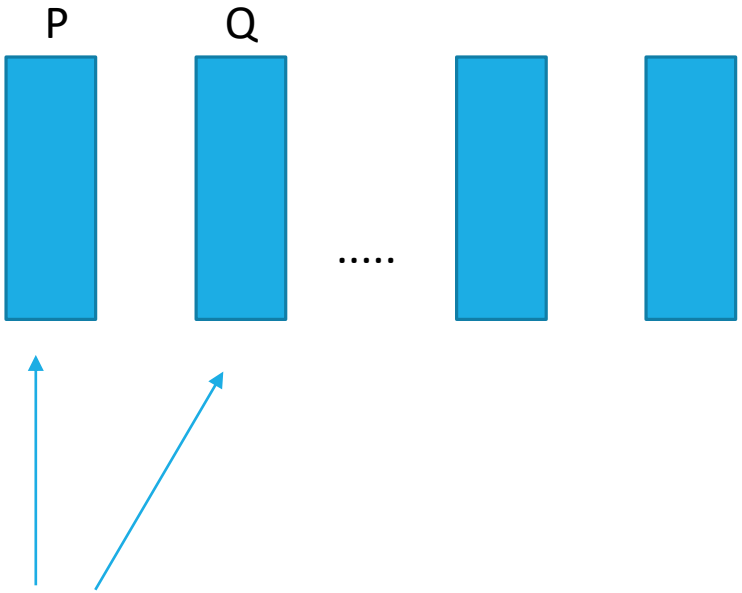


Topics

Term	Prob.
W_1	P_1
W_2	P_2
W_3	P_3
...	...
...	...
...	...
W_N	P_N

Documents

Quality of topics – topics dissimilarity



Topics

Kullback-Leibler Divergence

$$D_{\mathrm{KL}}(P \parallel Q) = \sum_{x \in \mathcal{X}} P(x) \log\left(\frac{P(x)}{Q(x)}\right)$$

Term	Prob.
W_1	P_1
W_2	P_2
W_3	P_3
...	...
...	...
...	...
W_N	P_N

Quality of topics – uniqueness of words

Latent Dirichlet Allocation $\sum_{j=1}^{3} p_{ij} - \max_j p_{ij}$ *Correlated Topic models*

	LDA					CTM			
	Topic_1	Topic_2	Topic_3			Topic_1	Topic_2	Topic_3	
term_1	0,1612	0,1405	0,0006	0,1410		0,0546	0,1129	0,0049	0,0596
term_2	0,0168	0,1595	0,1789	0,1764		0,0961	0,1509	0,0477	0,1438
term_3	0,0206	0,0338	0,1926	0,0544		0,1661	0,0286	0,0293	0,0579
term_4	0,1707	0,0658	0,0345	0,1003		0,0177	0,1031	0,2007	0,1209
term_5	0,0569	0,0837	0,1942	0,1407		0,0957	0,1449	0,0701	0,1658
term_6	0,2356	0,1775	0,1878	0,3653		0,1652	0,0939	0,1854	0,2592
term_7	0,0886	0,0558	0,0128	0,0687		0,0380	0,1586	0,0428	0,0808
term_8	0,0300	0,2018	0,0570	0,0870		0,0740	0,0026	0,1275	0,0766
term_9	0,1464	0,0810	0,0464	0,1274		0,1524	0,0232	0,1163	0,1395
term_10	0,0731	0,0005	0,0951	0,0736		0,1401	0,1813	0,1753	0,3154
SUM	1,0000	1,0000	1,0000	**1,6652**		1,0000	1,0000	1,0000	**1,5806**
				44,49%					**47,31%**

$3 - \sum_{i-1}^{10} s_i$

$\frac{\sum_{i-1}^{10} s_i}{3}$

smaller value indicates the better model

Comparison of LDA and CTM – discipline: political sciences and administration

Dissimilarity between topics:

```
> KL(prob_lda)
Metric: 'kullback-leibler' using unit: 'log2'; comparing: 5 vectors.
         v1       v2       v3       v4       v5
v1 0.000000 5.944242 5.758093 5.602890 5.534179
v2 5.944242 0.000000 5.582333 5.912793 6.068053
v3 5.758093 5.582333 0.000000 5.811606 5.531025
v4 5.602890 5.912793 5.811606 0.000000 4.940401
v5 5.534179 6.068053 5.531025 4.940401 0.000000
> KL(prob_ctm)
Metric: 'kullback-leibler' using unit: 'log2'; comparing: 5 vectors.
          v1        v2        v3        v4       v5
v1  0.000000 13.749837 13.496332 13.605192 5.358096
v2 13.749837  0.000000 14.488934 15.851105 6.334454
v3 13.496332 14.488934  0.000000 14.950687 6.832479
v4 13.605192 15.851105 14.950687  0.000000 6.727585
v5  5.358096  6.334454  6.832479  6.727585 0.000000
>
```

Uniqueness of words:

- LDA: 21,84%
- CTM: 44,48%

Results for selected disciplines

Disciplines under evaluation in the Kraków University of Economics

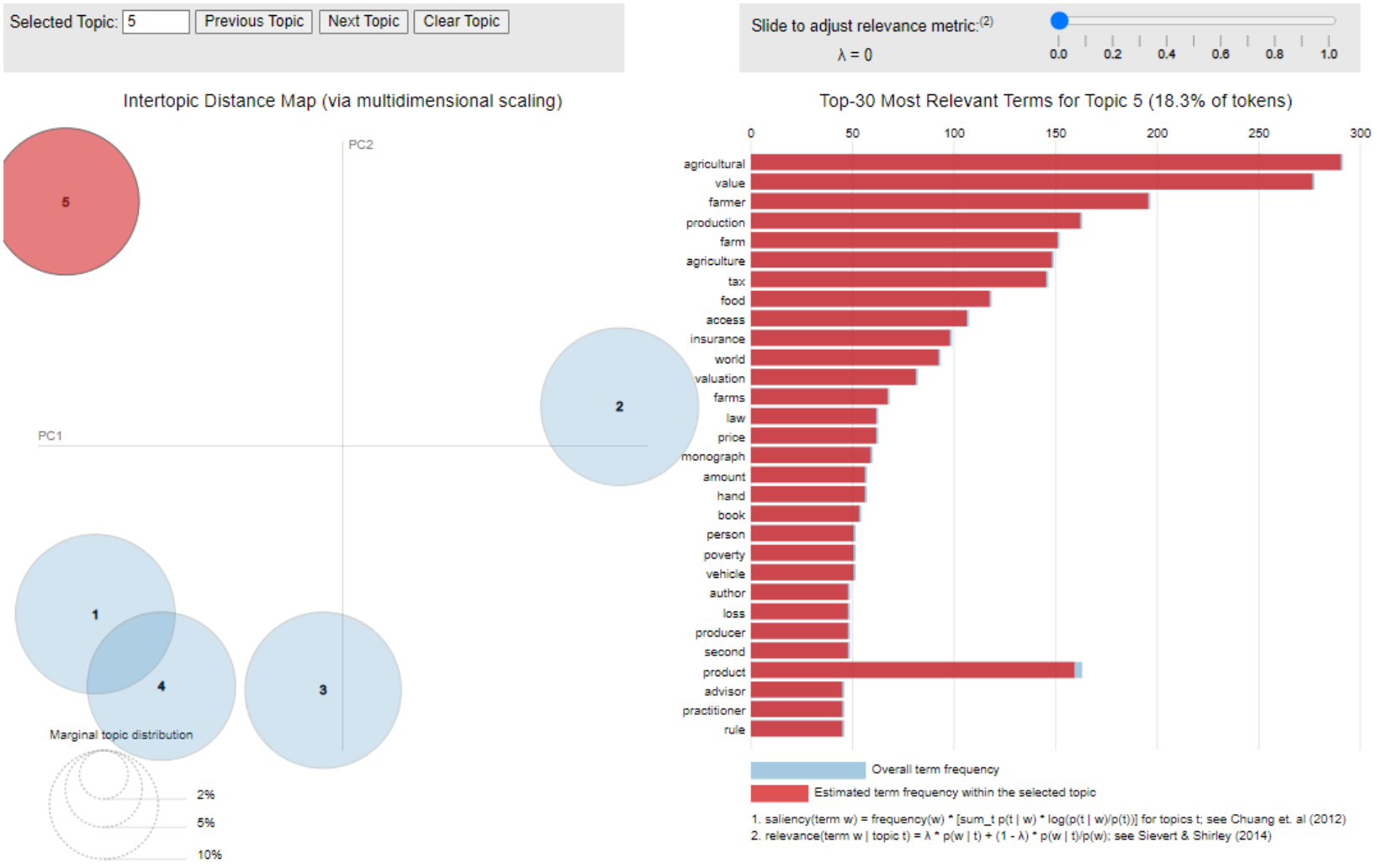
- economics and finance,
- socio-economic geography and spatial economy,
- political science and administration,
- legal sciences,
- management and quality sciences.

Discipline: economics and finance

Discipline: economics and finance

- Number of institutions: 47
- Number of cases of impact (descriptions): 113

Discipline: economics and finance

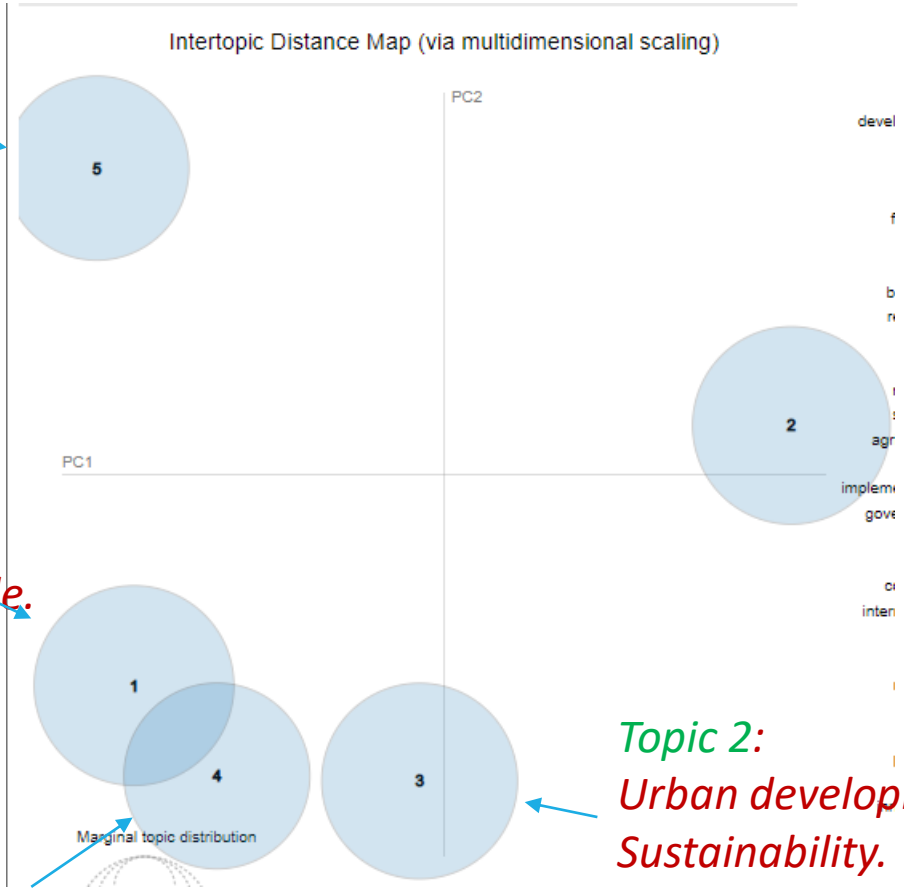


Discipline: economics and finance

Topic 1:
Agriculture.
Food market

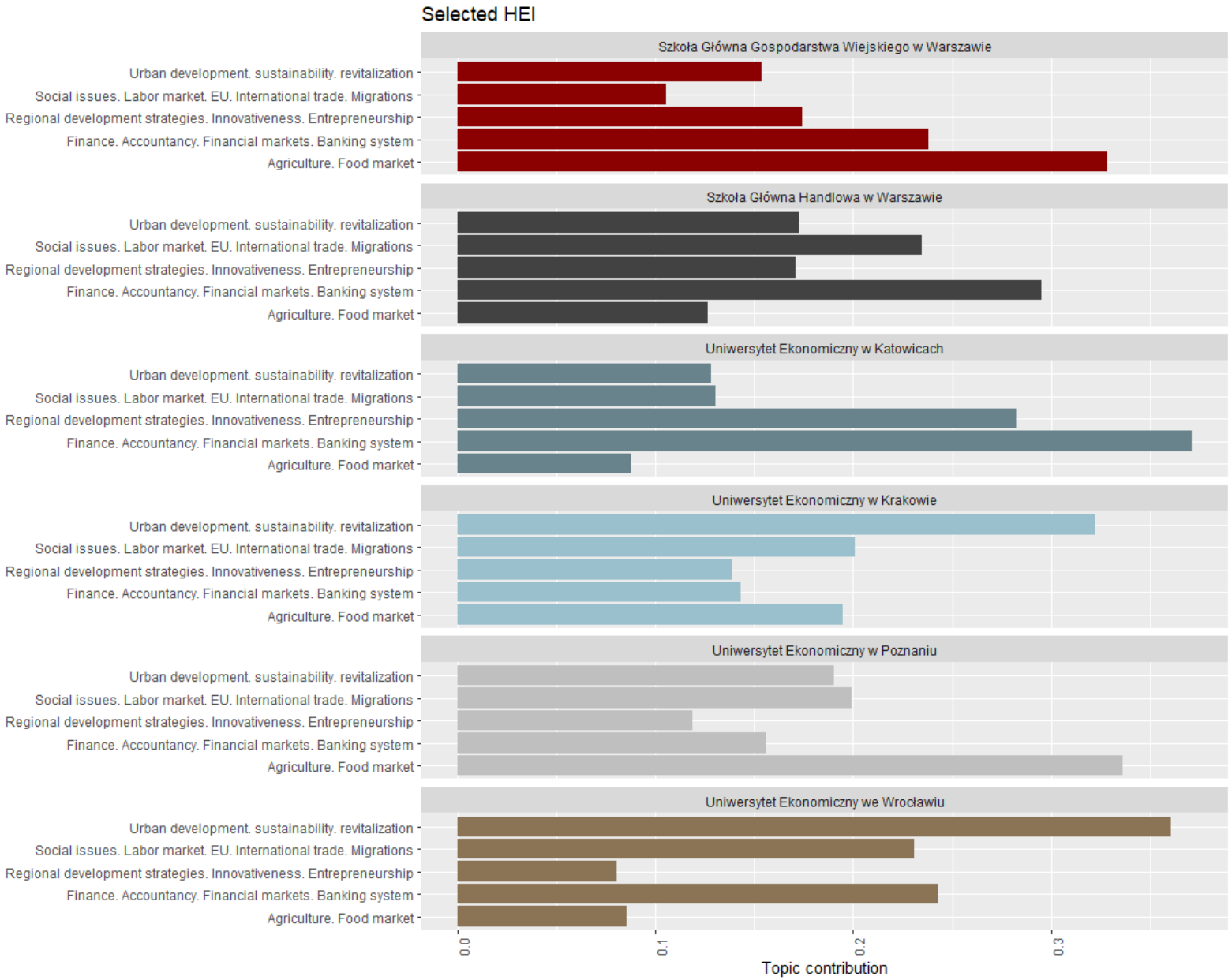
Topic 3:
Social issues.
Labor market.
EU.
International trade.
Migrations

Topic 4:
Finance.
Accountancy.
Financial markets.
Banking system.



Topic 5:
Regional development strategies.
Innovativeness.
Entrepreneurship

Topic 2:
Urban development.
Sustainability.
Revitalization


Selected HEI
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Szkoła Główna Handlowa w Warszawie
Uniwersytet Ekonomiczny w Katowicach
Uniwersytet Ekonomiczny w Krakowie
Uniwersytet Ekonomiczny w Poznaniu
Uniwersytet Ekonomiczny we Wrocławiu
Urban development. sustainability. revitalization
Social issues. Labor market. EU. International trade. Migrations
Regional development strategies. Innovativeness. Entrepreneurship
Finance. Accountancy. Financial markets. Banking system
Agriculture. Food market
0.0
0.1
0.2
0.3
Topic contribution

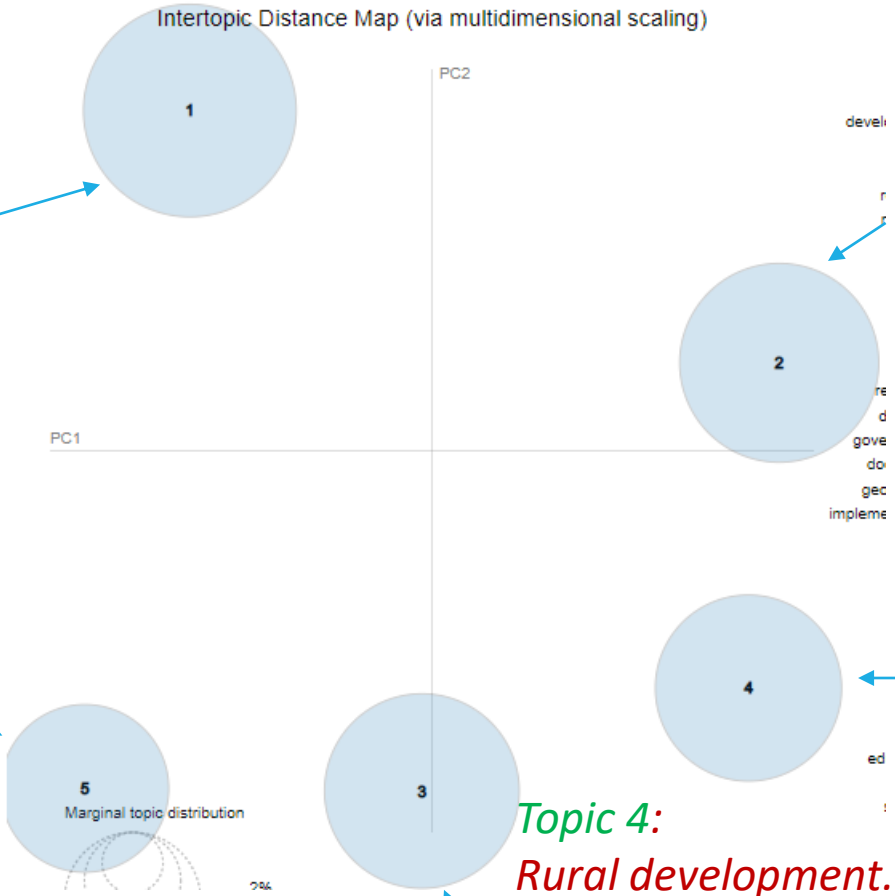
Discipline: Socio-economic geography and spatial economy

Dyscyplina: Socio-economic geography and spatial economy

- Number of institutions: 17
- Number of cases of impact (descriptions): 38

Dyscyplina: Socio-economic geography and spatial economy

Topic 5:
Regional development. Spatial management. Urbanization.

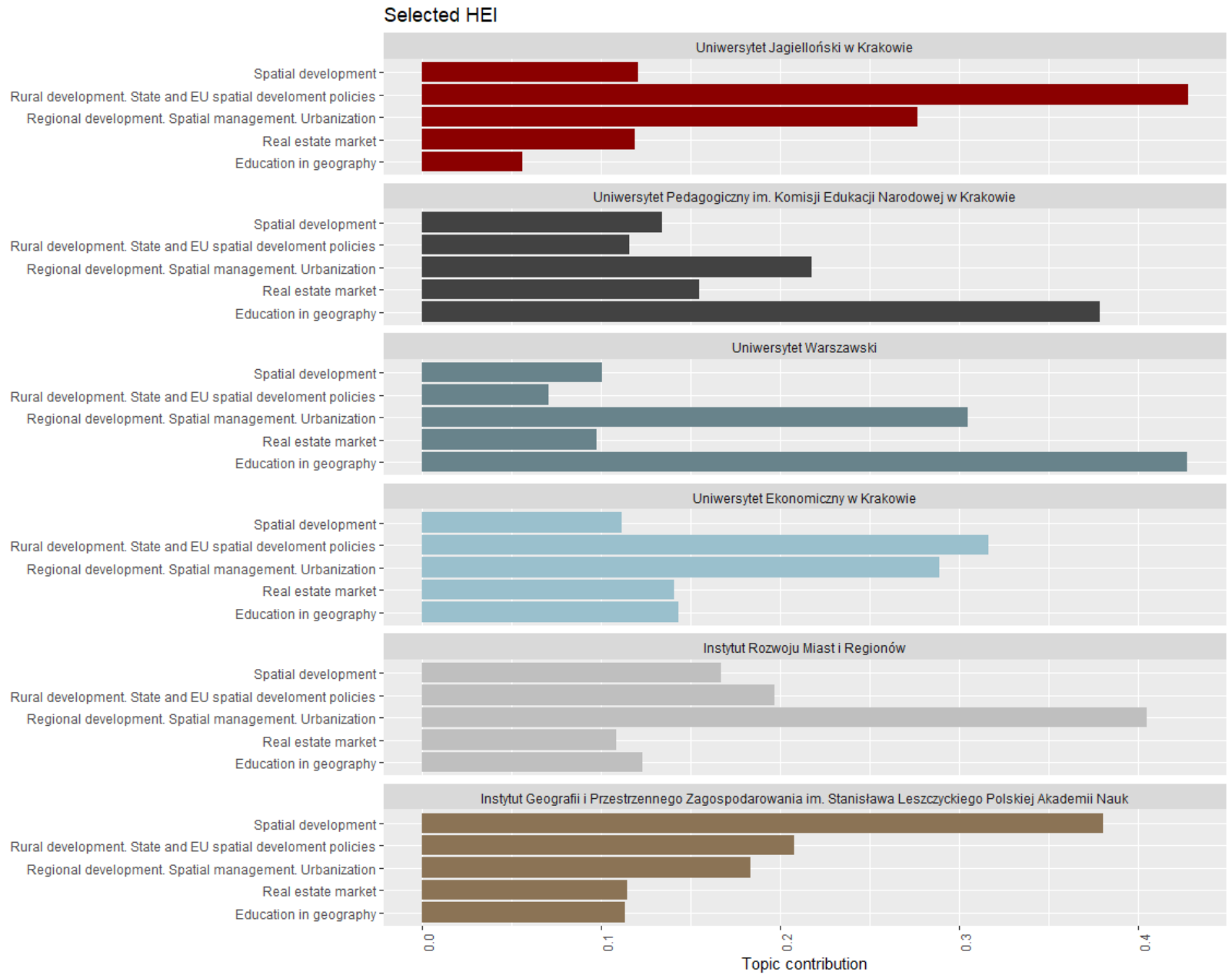


Topic 2:
Spatial development.

Topic 3:
Education in geography

Topic 1:
Real estate market.

Topic 4:
Rural development. State and EU spatial develoment policies.

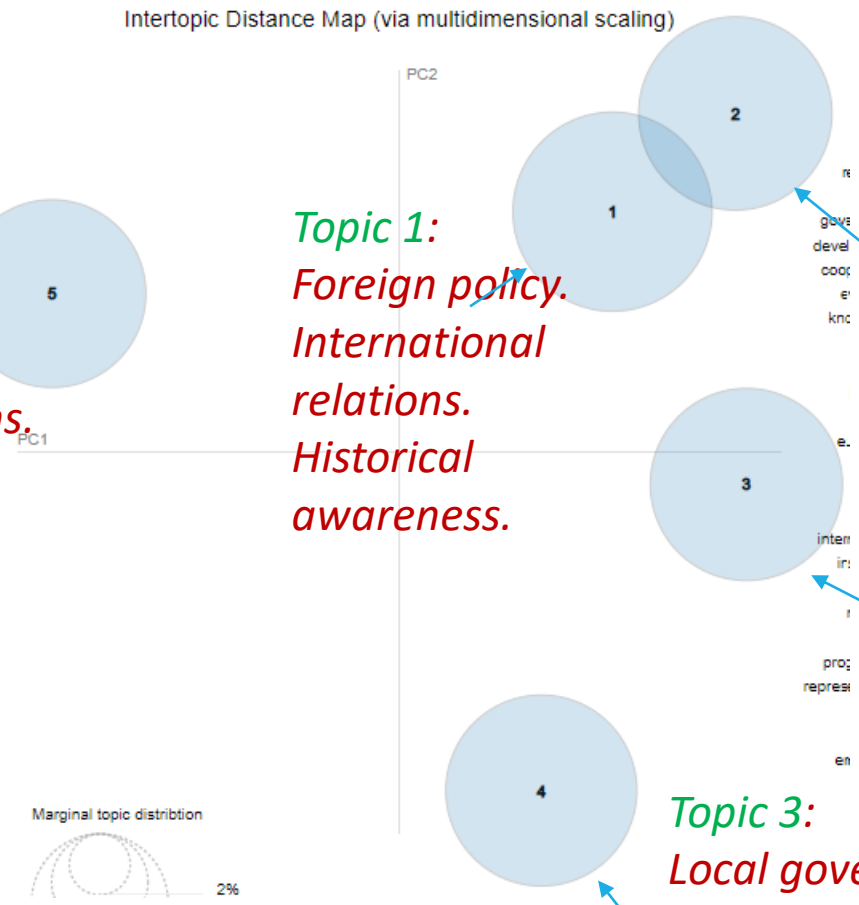

Selected HEI
Uniwersytet Jagielloński w Krakowie
Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej w Krakowie
Uniwersytet Warszawski
Uniwersytet Ekonomiczny w Krakowie
Instytut Rozwoju Miast i Regionów
Instytut Geografii i Przestrzennego Zagospodarowania im. Stanisława Leszczyckiego Polskiej Akademii Nauk
Spatial development
Rural development. State and EU spatial develoment policies
Regional development. Spatial management. Urbanization
Real estate market
Education in geography
0.0
0.1
0.2
0.3
0.4
Topic contribution

Discipline: Political science and administration

Discipline: Political science and administration

- Number of institutions: 34
- Number of cases of impact (descriptions): 76

Discipline: Political science and administration



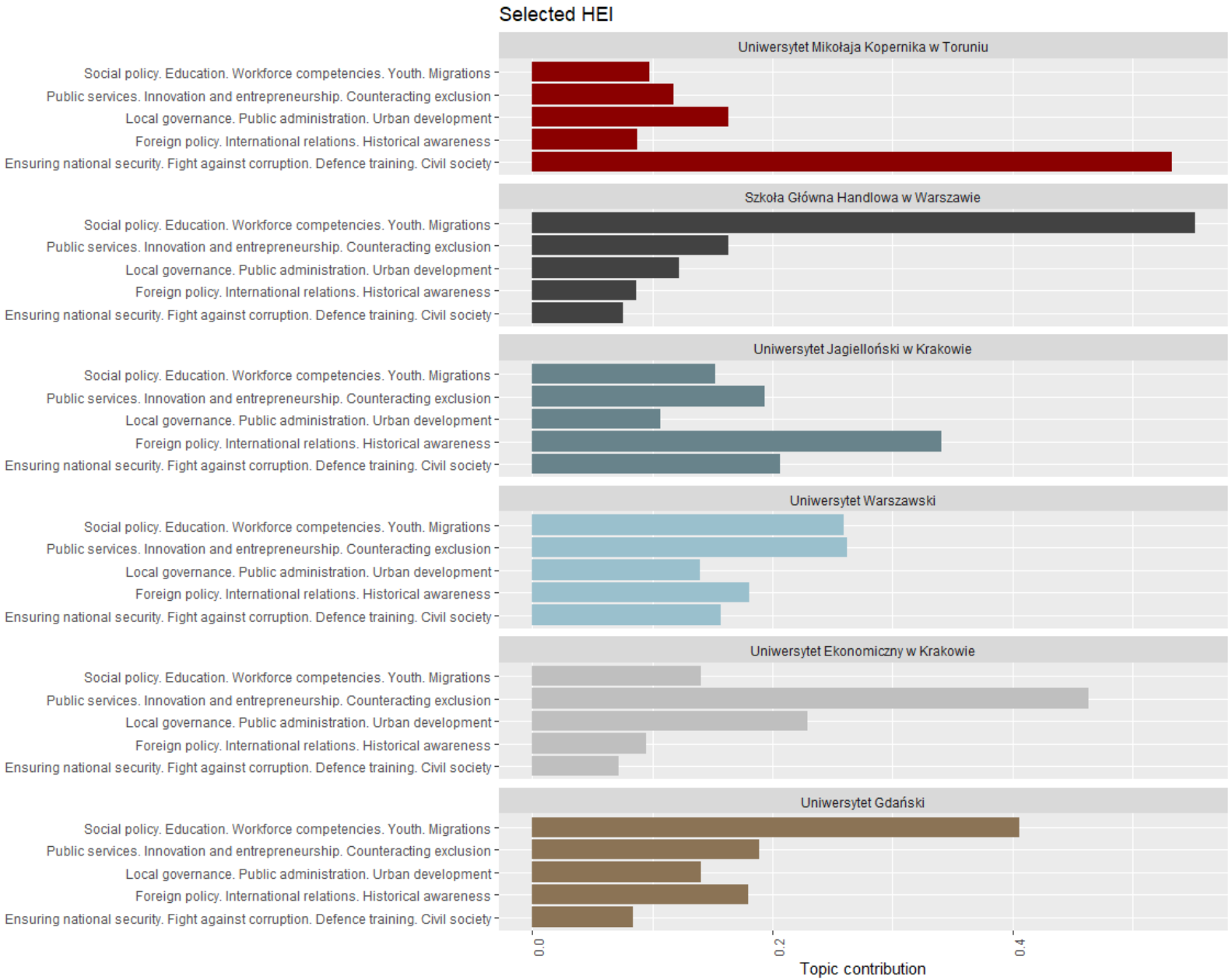
Topic 2:
Social policy. Education. Workforce competencies. Youth. Migrations.

Topic 1:
Foreign policy. International relations. Historical awareness.

Topic 4:
Public services. Innovation and entrepreneurship. Counteracting exclusion.

Topic 5:
Ensuring national security. Fight against corruption. Defence training. Civil society.

Topic 3:
Local governance. Public administration. Urban development.


Selected HEI
Uniwersytet Mikołaja Kopernika w Toruniu
Szkoła Główna Handlowa w Warszawie
Uniwersytet Jagielloński w Krakowie
Uniwersytet Warszawski
Uniwersytet Ekonomiczny w Krakowie
Uniwersytet Gdański
Social policy. Education. Workforce competencies. Youth. Migrations
Public services. Innovation and entrepreneurship. Counteracting exclusion
Local governance. Public administration. Urban development
Foreign policy. International relations. Historical awareness
Ensuring national security. Fight against corruption. Defence training. Civil society
0.0
0.2
0.4
Topic contribution

Discipline: Legal sciences

Discipline: Legal sciences

- Number of institutions: 41
- Number of cases of impact (descriptions): 96

Discipline: Legal sciences

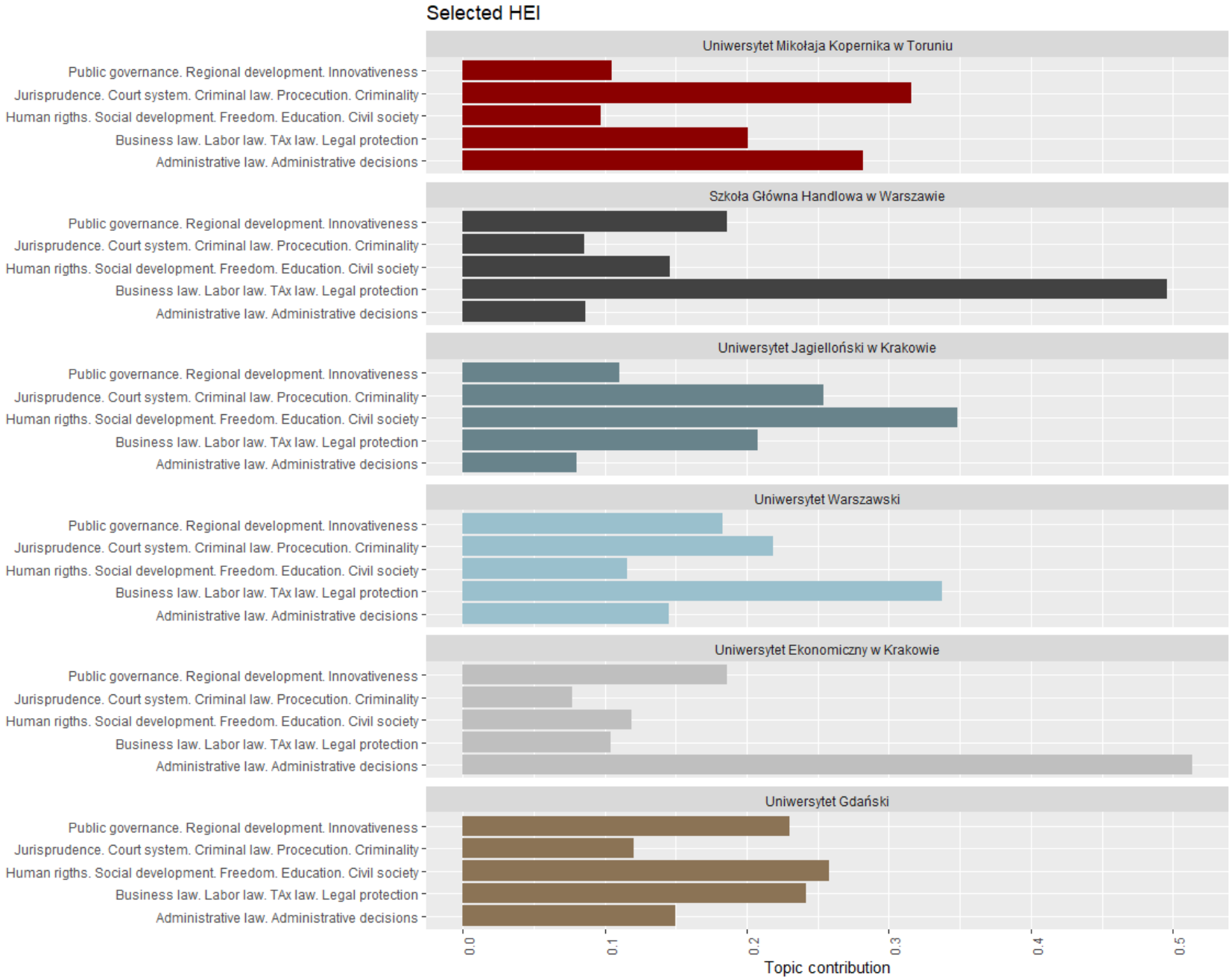
Topic 1:
Human rigths. Social development. Freedom. Education. Civil socjety.

Topic 4:
Administrative law. Administrative decisions.

Topic 3:
Business law. Labor law. Tax law. Legal protection

Topic 5:
Public governance. Regional development. Innovativeness.

Topic 2:
Jurisprudence. Court system. Criminal law. Procecution. Criminality.


Selected HEI
Uniwersytet Mikołaja Kopernika w Toruniu
Szkoła Główna Handlowa w Warszawie
Uniwersytet Jagielloński w Krakowie
Uniwersytet Warszawski
Uniwersytet Ekonomiczny w Krakowie
Uniwersytet Gdański
Public governance. Regional development. Innovativeness
Jurisprudence. Court system. Criminal law. Procecution. Criminality
Human rigths. Social development. Freedom. Education. Civil society
Business law. Labor law. TAx law. Legal protection
Administrative law. Administrative decisions
0.0
0.1
0.2
0.3
0.4
0.5
Topic contribution

Discipline: Management and quality sciences

Discipline: Management and quality sciences

- Number of institutions: 57
- Number of cases of impact (descriptions): 129

Discipline: Management and quality sciences

Topic 2:
Tourism. Cultural heritage. Urban governance.

Topic 3:
Regional development. Transportation. Entrepreneurship & innovation.

Topic 4:
Public services. Health care. Crisis management.

Topic 1:
Product management. Energy saving. Ecological production. Environmental protection

Topic 5:
Education. Competence development. People with disabilities.


Selected HEI
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Szkoła Główna Handlowa w Warszawie
Uniwersytet Ekonomiczny w Katowicach
Uniwersytet Ekonomiczny w Krakowie
Uniwersytet Ekonomiczny w Poznaniu
Uniwersytet Ekonomiczny we Wrocławiu
Tourism. Cultural heritage. Urban governance
Regional development. Transportation. Entrepreneurship & innovation
Public services. Health care. Crisis management
Product management. Energy saving. Ecological production. Environmental protection
Education. Competence development. People with disabilities
0.0
0.1
0.2
0.3
0.4
0.5
Topic contribution

Conclusions

Conclusions

- Topic modelling methods allowed to identify main issues reported by Polish universities as achievements evaluated as cases of impact on society or economy.
- Comparison of LDA and CTM methods shows:
  - CTM is better in terms of dissimilarity of topics,
  - LDA is better in terms of uniqueness of words.
- Achievements of Polish universities in selected social sciences are diversified.

UNIWERSYTET EKONOMICZNY W KRAKOWIE

THANK YOU FOR YOUR ATTENTION!

PAWEŁ LULA, MAGDALENA TALAGA