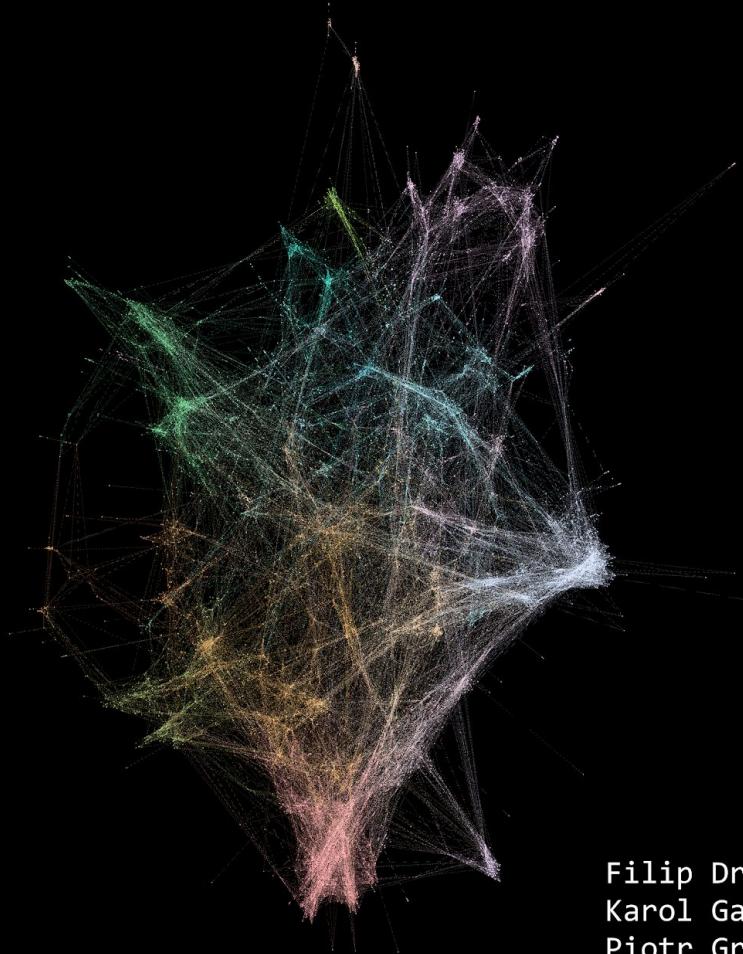


Analiza powiązań pomiędzy pracownikami PWR

Filip Dratwiński
Karol Gawron
Piotr Gramacki
Maksym Telepchuk

Sieć cytowań pracowników PWr



Filip Dratwiński
Karol Gawron
Piotr Gramacki
Maksym Telepchuk

Data Science Workflow Canvas

Problem statement	Outcomes/Predictions	Data Acquisition
Czy są grupy wzajemnie się cytujących osób? Czy możemy wskazać osoby, których brak byłby znaczący dla innych autorów?	Wejście: Sieć pracowników połączonych cytowaniami Wyjście: Pogrupowana sieć	Pobranie danych o cytowaniach w obrębie PWr z bazy Scopus. Wykorzystanie dostępnego API
Modeling	Model Evaluation	Data Preparation
Skierowana sieć ważona, gdzie wierzchołki przedstawiają pracowników, a krawędź oznacza cytowanie. Waga występuje ze względu na to, że jedna osoba może cytować inną osobę wiele razy. Algorytmy grupowania	Wizualizacja powstałej sieci i wyników oraz ocena własna Ocena algorytmu grupowania na podstawie Modularity	Ekstrakcja listy pracowników obecnych w cytowaniach Utworzenie listy identyfikatorów dla każdego pracownika w celu opisania wierzchołków Przetworzenie cytowań na listę krawędzi skierowanych.

Problem

Czy pracownicy Politechniki korzystają z dorobku swoich kolegów?

- oparcie analizy na wzajemnych cytowaniach wewnętrz politechniki

Czy pracownicy na PWr wykorzystują dorobek naukowy innych wydziałów?

- porównanie występujących w sieci grup z faktycznym podziałem administracyjnym

Jakie osoby są ważnymi węzłami sieci? Kogo publikacje są najczęściej cytowane?

- identyfikacja najważniejszych węzłów w powstałej sieci

Scopus API

GET https://api.elsevier.com/content/search/author?apiKey=8bfaf0793cf65175eef1b3a358eca655&query=AF-ID(60019987)&count=200

Params Authorization Headers (6) Body Pre-request Script Tests Settings

Query Params

KEY	VALUE	DESCRIPTION
apiKey	[REDACTED]	
query	AF-ID(60019987)	
count	200	

Body Cookies Headers (16) Test Results

Pretty Raw Preview Visualize JSON

```
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
```

```
        },
        {
            "@_fa": "true",
            "@ref": "scopus-citedby",
            "@href": "https://www.scopus.com/author/citedby.uri?partnerID=HzOxMe3b&citedAuthorId=7103360092&origin=inward"
        },
        {
            "@_fa": "true",
            "@ref": "scopus-author",
            "@href": "https://www.scopus.com/authid/detail.uri?partnerID=HzOxMe3b&authorId=7103360092&origin=inward"
        }
    ],
    "prism:url": "https://api.elsevier.com/content/author/7103360092",
    "dc:identifier": "AUTHOR_ID:7103360092",
    "eid": "9-s2.0-7103360092",
    "prefixed-name": [
        {
            "surname": "Misiiewicz",
            "given-name": "Jan",
            "initials": "J."
        }
    ],
    "name-variant": [
    ]
```

link do prac cytujących

identyfikator autora

nazwisko autora

i Scopus... niekoniecznie API

 Scopus

Search Sources Lists SciVal ↗ ⓘ ⚡ Create account Sign in

1,575 documents have cited:

"Miniewicz, Andrzej" 7006404110

 Set alert

Search within results... 

Refine results

Access type ⓘ

- Open Access (135) >
- Other (1,440) >

Year

- 2020 (37) >
- 2019 (123) >

Analyze search results

Show all abstracts Sort on: Date (newest)

Document title	Authors	Year	Source	Cited by
<input type="checkbox"/> 1 Facile conversion of cis into trans oxane as liquid crystals Open Access	Dong, Z., Fan, P., Chen, Q., Chen, C.	2020	Scientific Reports 10(1),6936	0
View abstract  View at Publisher Related documents				
<input type="checkbox"/> 2 Electrically controlled white laser emission through liquid crystal/polymer multiphases Open Access	Adamow, A., Szukalski, A., Sznitko, L., (...), Camposeo, A., Mysliwiec, J.	2020	Light: Science and Applications 9(1),19	0
View abstract  View at Publisher Related documents				

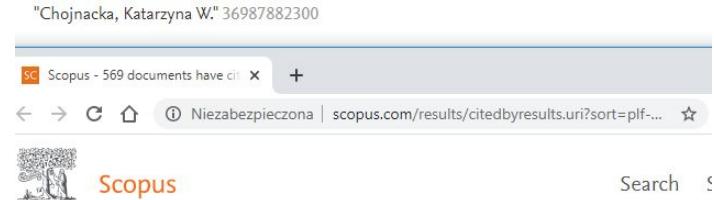
Przygotowanie danych

- Usunięcie autorów bez cytowań
- Usunięcie powtarzających się nazwisk
- Integracja z danymi o podziale administracyjnym
pwr.edu.pl/pracownicy/strefa-pracownika/wyszukiwarka-pracownikow
- Integracja z danymi od grupy analizującej współautorstwo



Scopus - 3,100 documents have cited: ["Chojnacka, Katarzyna W."](#) 36987882300

3,100 documents have cited:



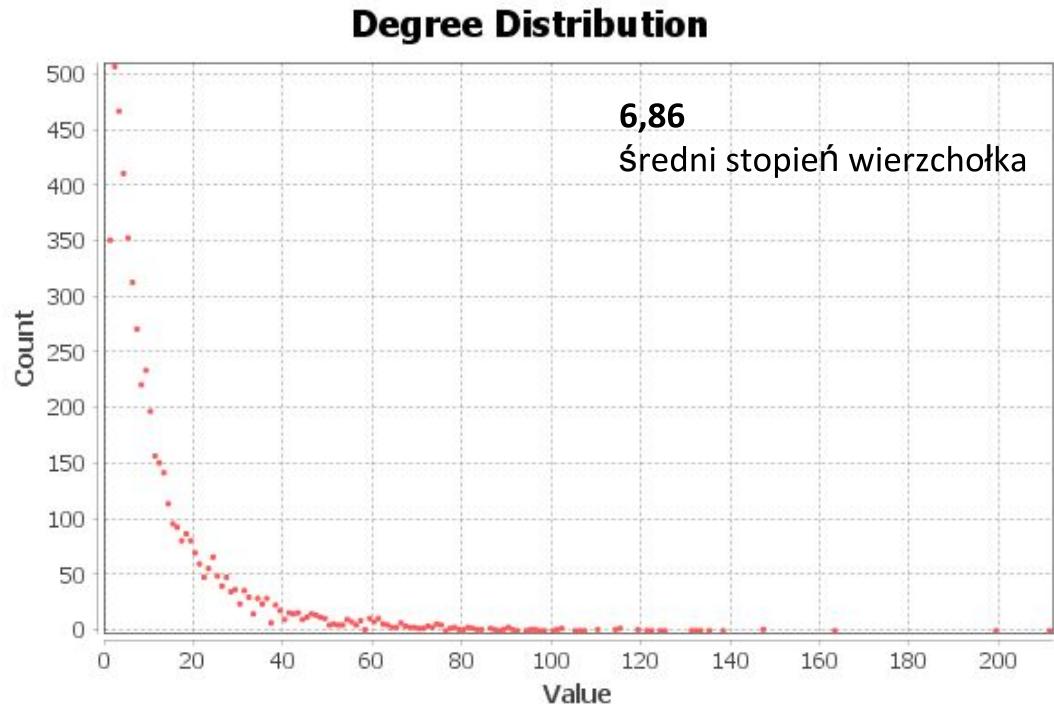
Scopus - 569 documents have cited: ["Chojnacka, Katarzyna"](#) 6602612182

569 documents have cited:

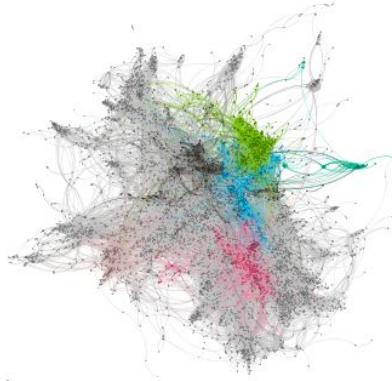
Graf skierowany - pracowników PWr

5494
węzłów - autorów

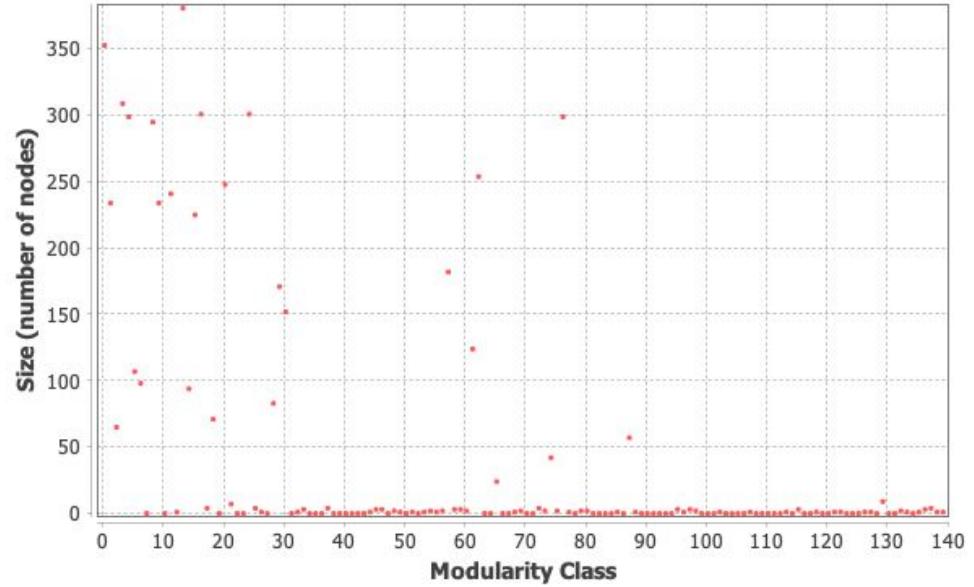
37707
krawędzi* - cytowań
* z wagami - liczbą cytowań



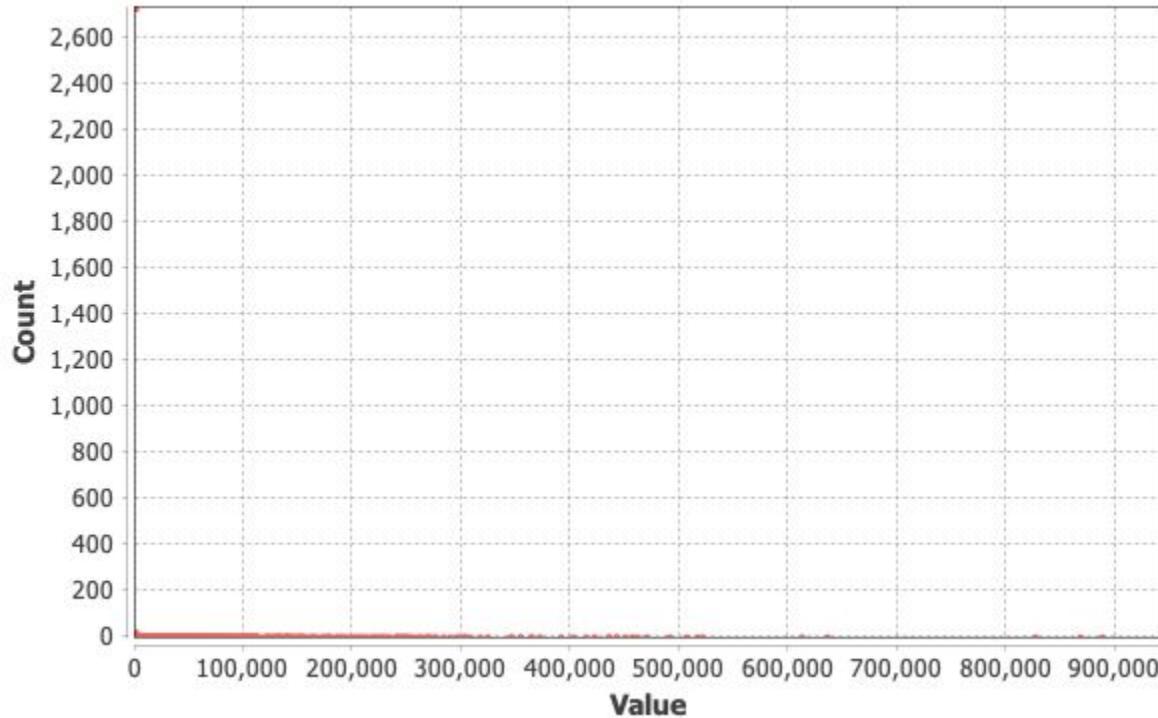
Modularity



Size Distribution



Betweenness Centrality Distribution



Graf skierowany - pracowników PWr

113

spójnych składowych*

* ang. connected components

5394

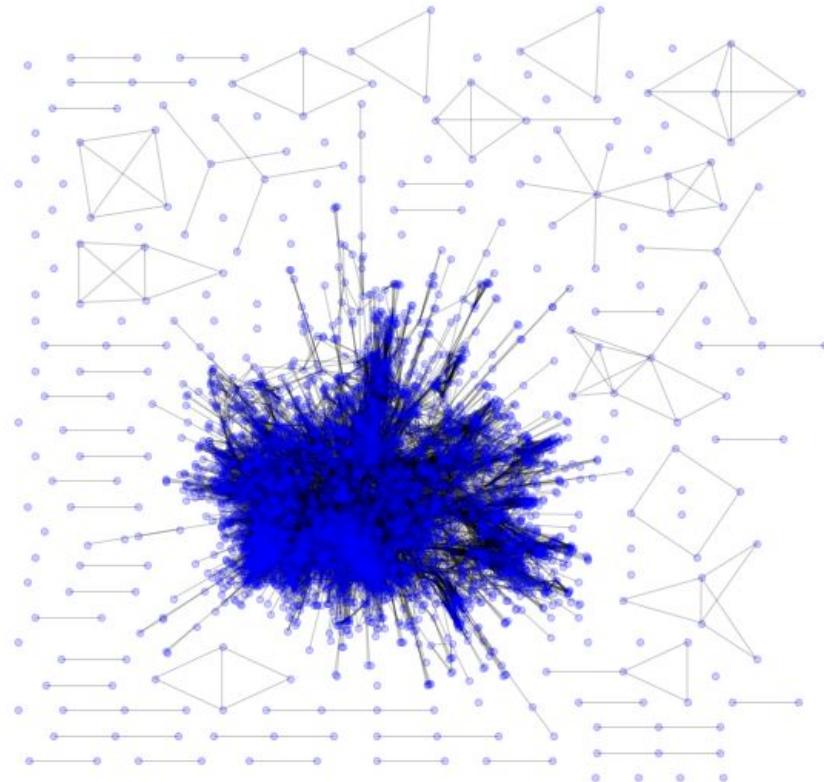
silnie spójnych składowych*

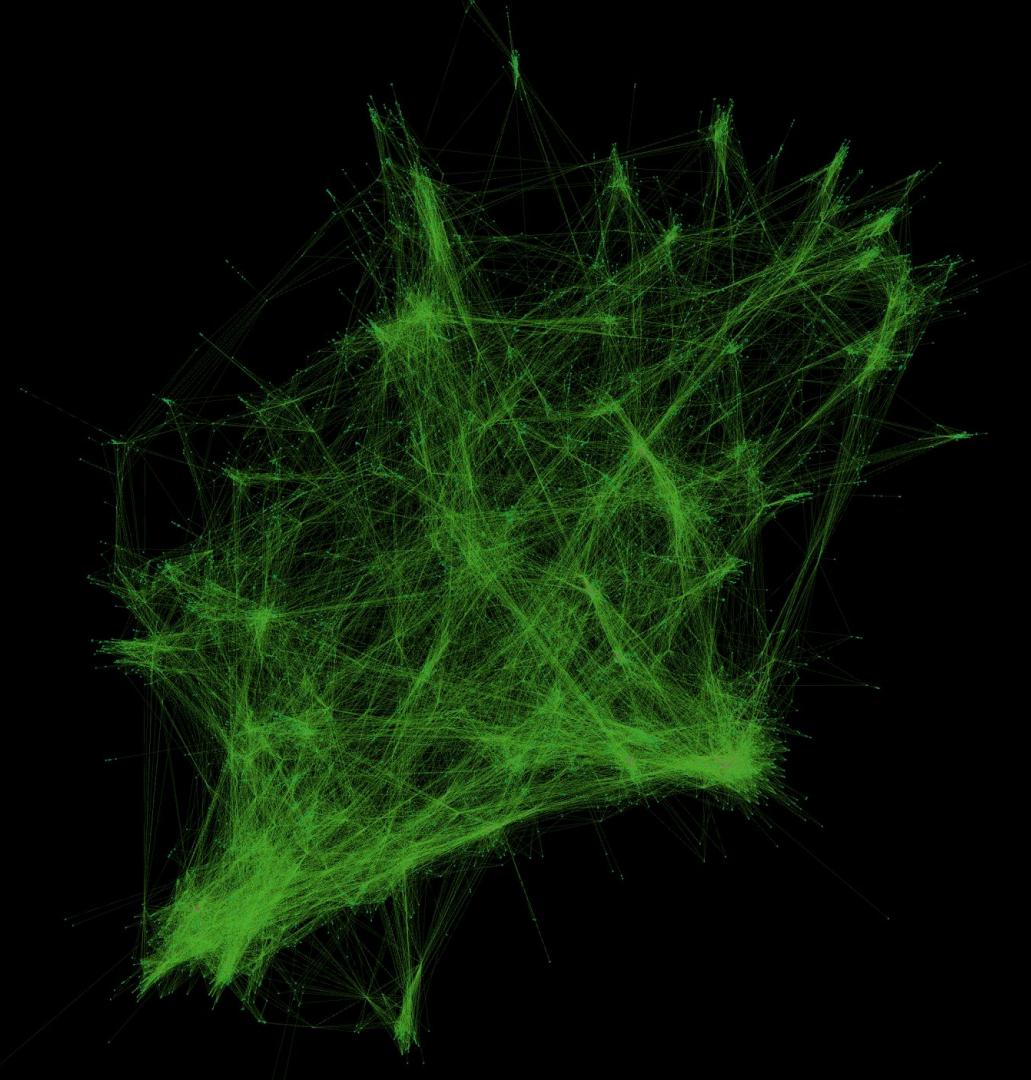
* ang. strongly connected components

5218

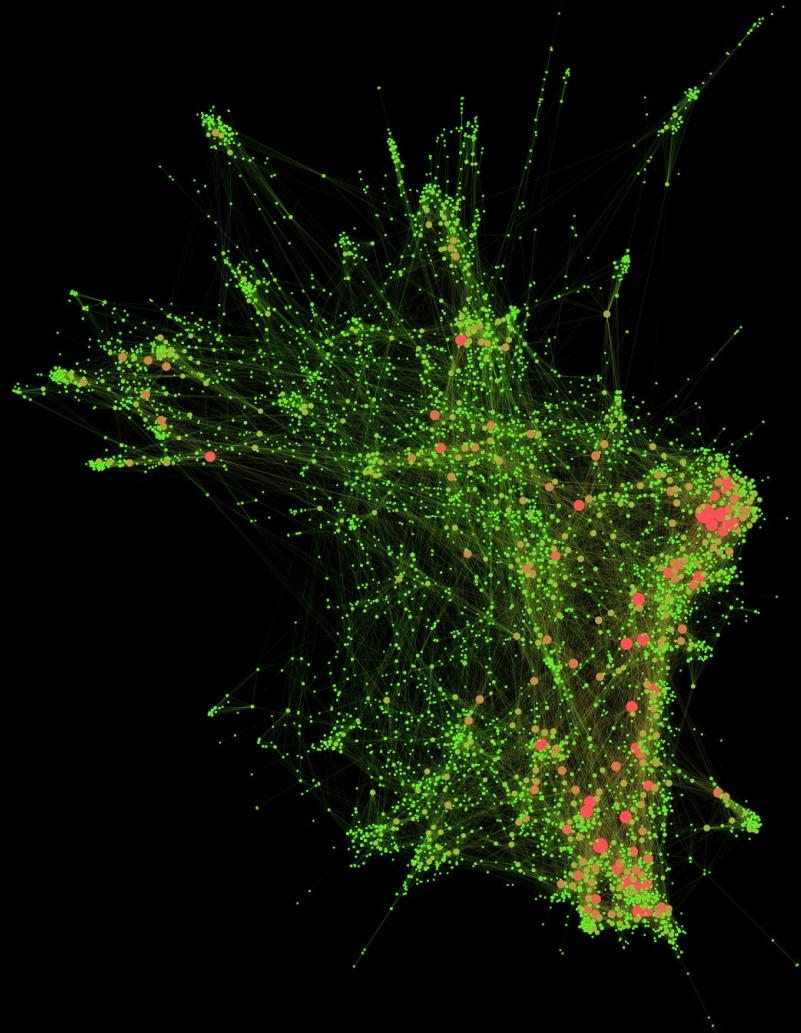
węzłów w największym komponencie*

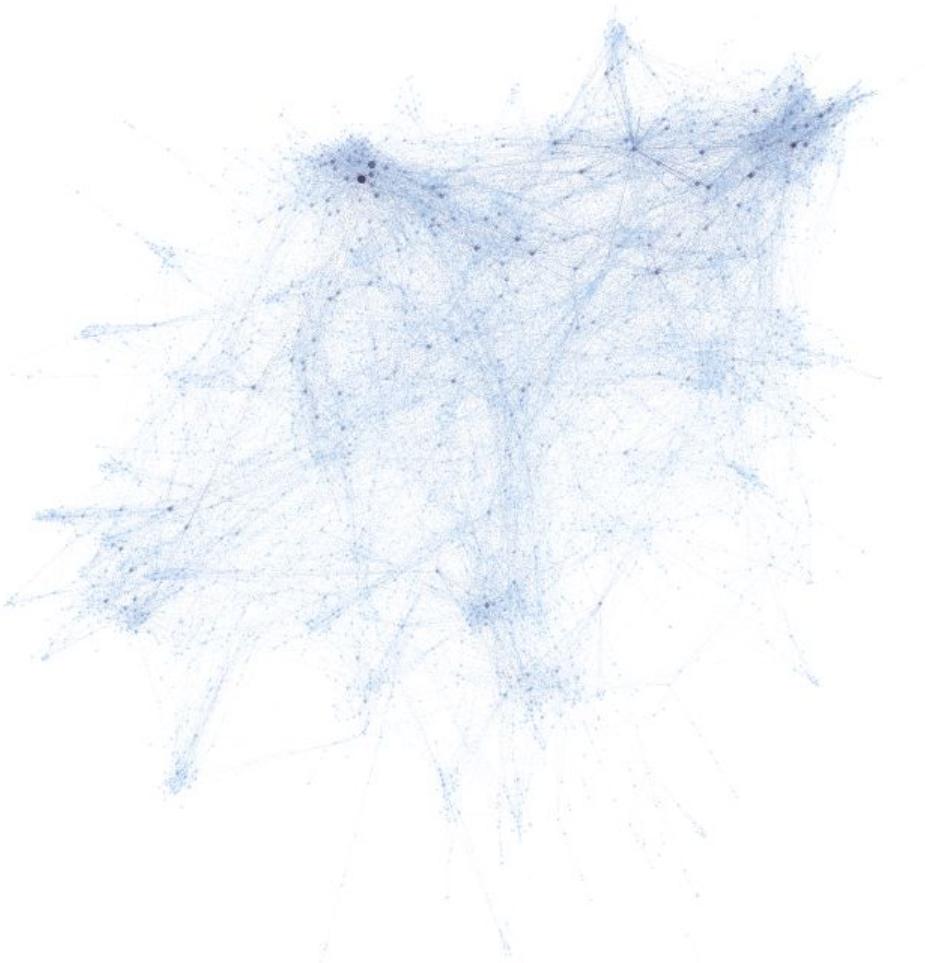
* ang. nodes in giant component







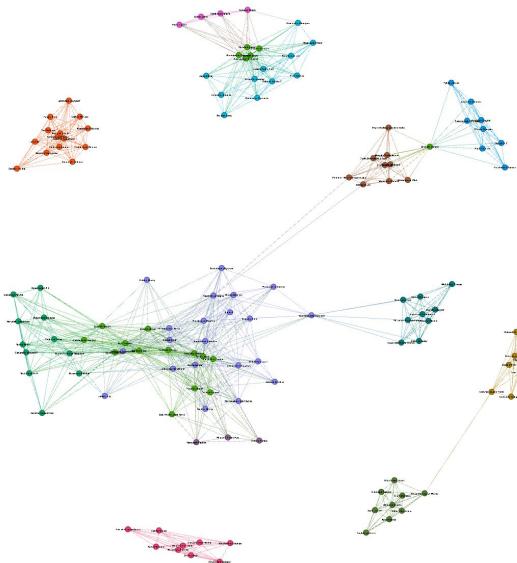




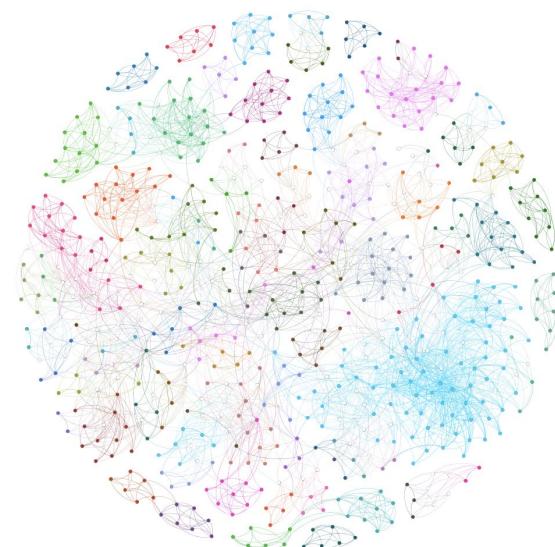
CFinder

- Jest algorytmem typu **Clique Percolation**
- Są wymagane krawędzi nieskierowane
- Nie skierowana krawędź występuje przy wzajemnym cytowaniu
- Jest w stanie znaleźć klastry nakładające się
- Wymaga parametru wejściowego K
- Wyszukuje silnie połączone grupy

CFinder

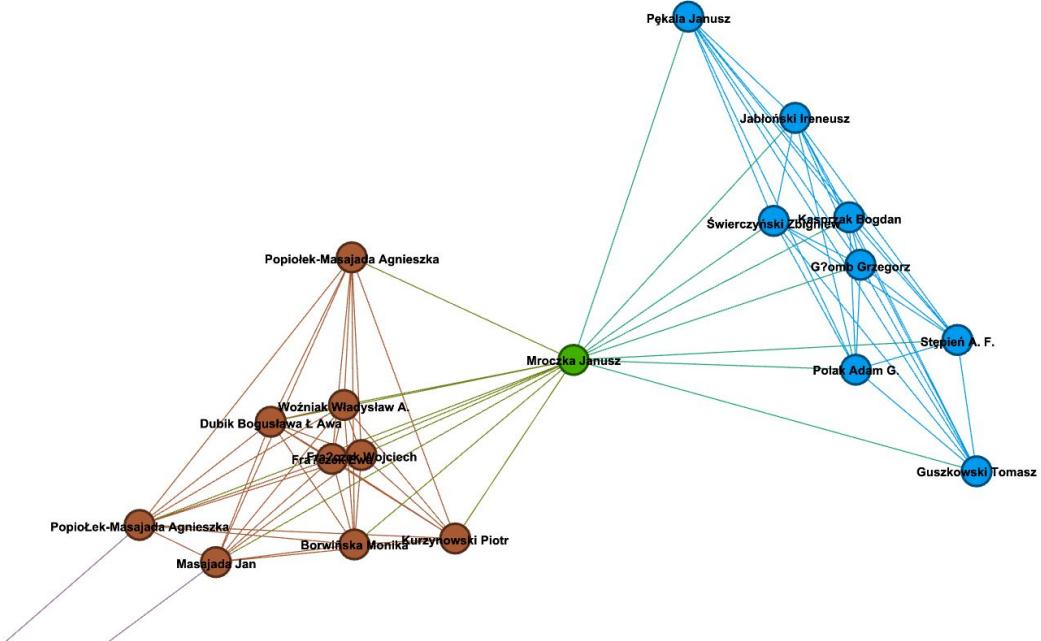


K = 9



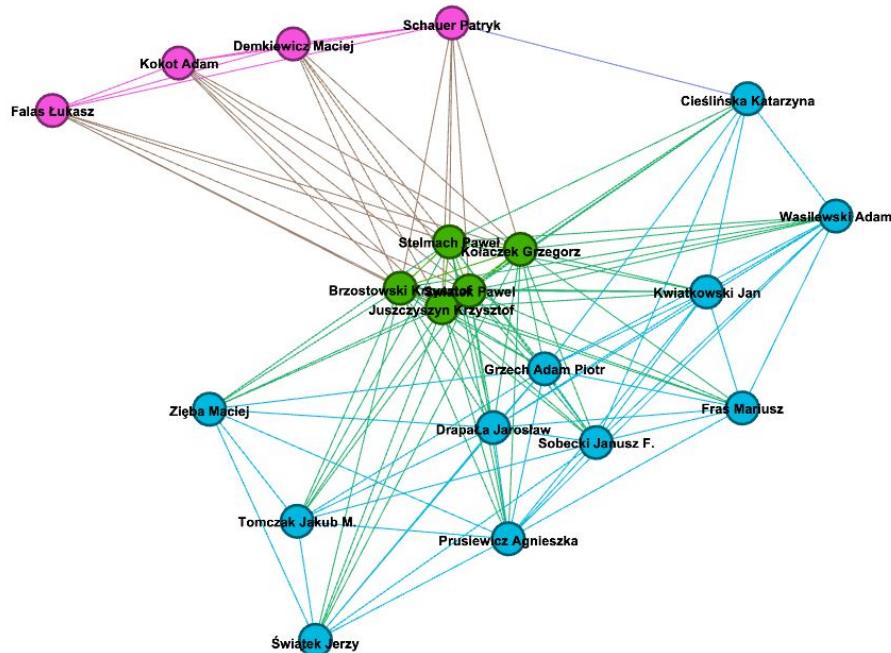
K = 5

CFinder



Janusz	Mroczka	W04	Katedra Metrologii Elektronicznej i Fotonicznej
Janusz	Pękala	W04	Katedra Metrologii Elektronicznej i Fotonicznej
Ireneusz	Jabłonecki	W04	Katedra Metrologii Elektronicznej i Fotonicznej
Zbigniew	Świerczyński	W04	Katedra Metrologii Elektronicznej i Fotonicznej
Bogdan	Kasprzak	W04	Unknown
Grzegorz	Glomb	W04	Katedra Metrologii Elektronicznej i Fotonicznej
Adam	Polak	W04	Katedra Metrologii Elektronicznej i Fotonicznej
Stępień	A F	W04	Unknown
Guszkowski	Tomasz	W04	Unknown
Agnieszka	Popiółek-Masajada	W11	Katedra Optyki i Fotoniki
Władysław	Woźniak	W11	Katedra Optyki i Fotoniki
Bogusława	Dubik	W11	Unknown
Ewa	Frączek	W04	Katedra Telekomunikacji i Teleinformatyki
Wojciech	Frączek	W04	Katedra Telekomunikacji i Teleinformatyki
Piotr	Kurzynowski	W11	Katedra Optyki i Fotoniki
Monika	Borwińska	W11	Katedra Optyki i Fotoniki
Agnieszka	Popiółek-Masajada	W11	Katedra Optyki i Fotoniki
Jan	Masajada	W11	Katedra Optyki i Fotoniki

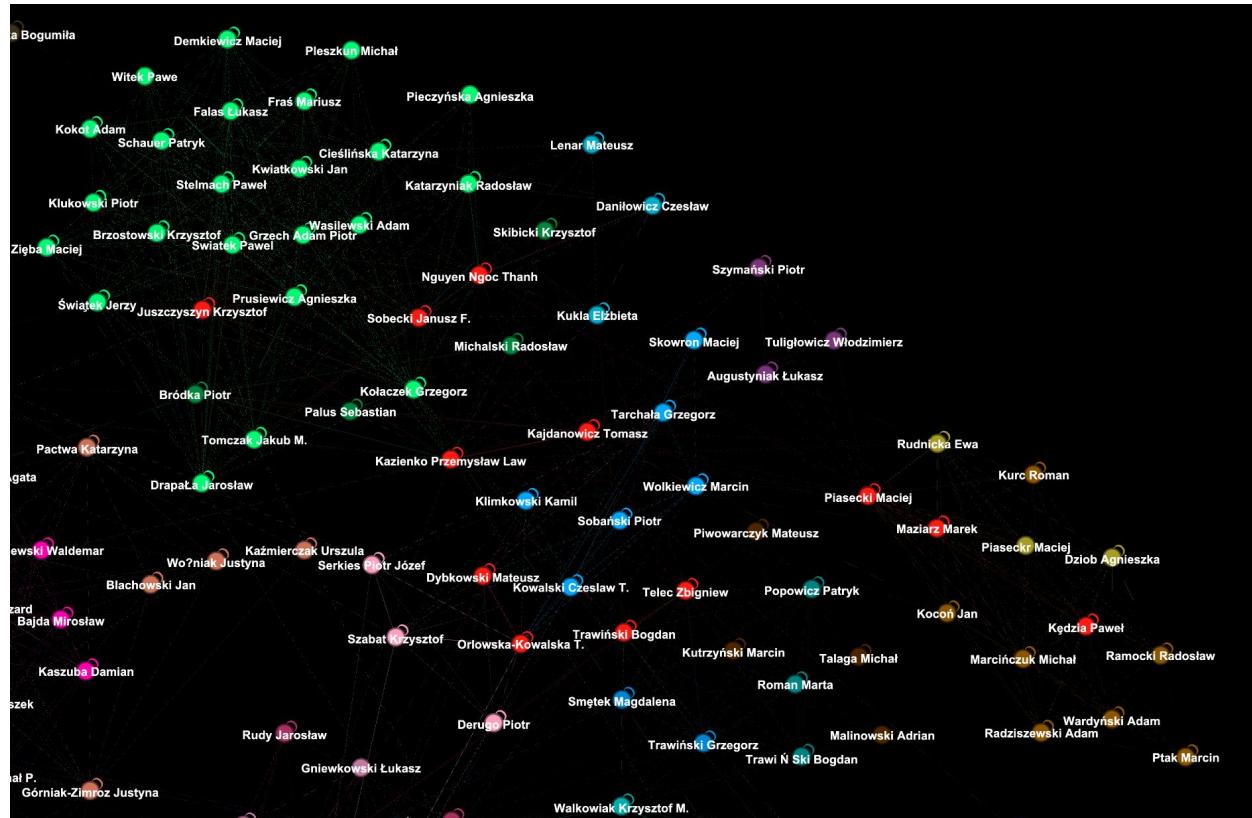
CFinder



Łukasz	Falas	W08	Katedra Informatyki i Inżynierii Systemów
Kokot	Adam	W08	Unknown
Demkiewicz	Maciej	W08	Unknown
Patryk	Schauer	W08	Katedra Informatyki i Inżynierii Systemów
Cieślińska	Katarzyna	W08	Unknown
Adam	Wasilewski	W08	Katedra Badań Operacyjnych i Inteligencji Biznesowej
Jan	Kwiatkowski	W08	Katedra Informatyki i Inżynierii Systemów
Mariusz	Fraś	W08	Katedra Informatyki i Inżynierii Systemów
Janusz	Sobecki	W08	Katedra Informatyki i Inżynierii Systemów
Adam	Grzech	W08	Unknown
Jarosław	Drapała	W08	Katedra Informatyki i Inżynierii Systemów
Agnieszka	Prusiewicz	W08	Unknown
Maciej	Zięba	W08	Katedra Informatyki i Inżynierii Systemów
Jakub M.	Tomczak	W08	Unknown
Jerzy	Świątek	W08	Katedra Informatyki i Inżynierii Systemów
Paweł	Stelmach	W08	Unknown
Grzegorz	Kołaczek	W08	Katedra Informatyki i Inżynierii Systemów
Krzysztof	Brzostowski	W08	Katedra Informatyki i Inżynierii Systemów
Paweł	Świątek	W08	Katedra Informatyki i Inżynierii Systemów
Krzysztof	Juszczyszyn	W08	Katedra Informatyki i Inżynierii Systemów

CFinder

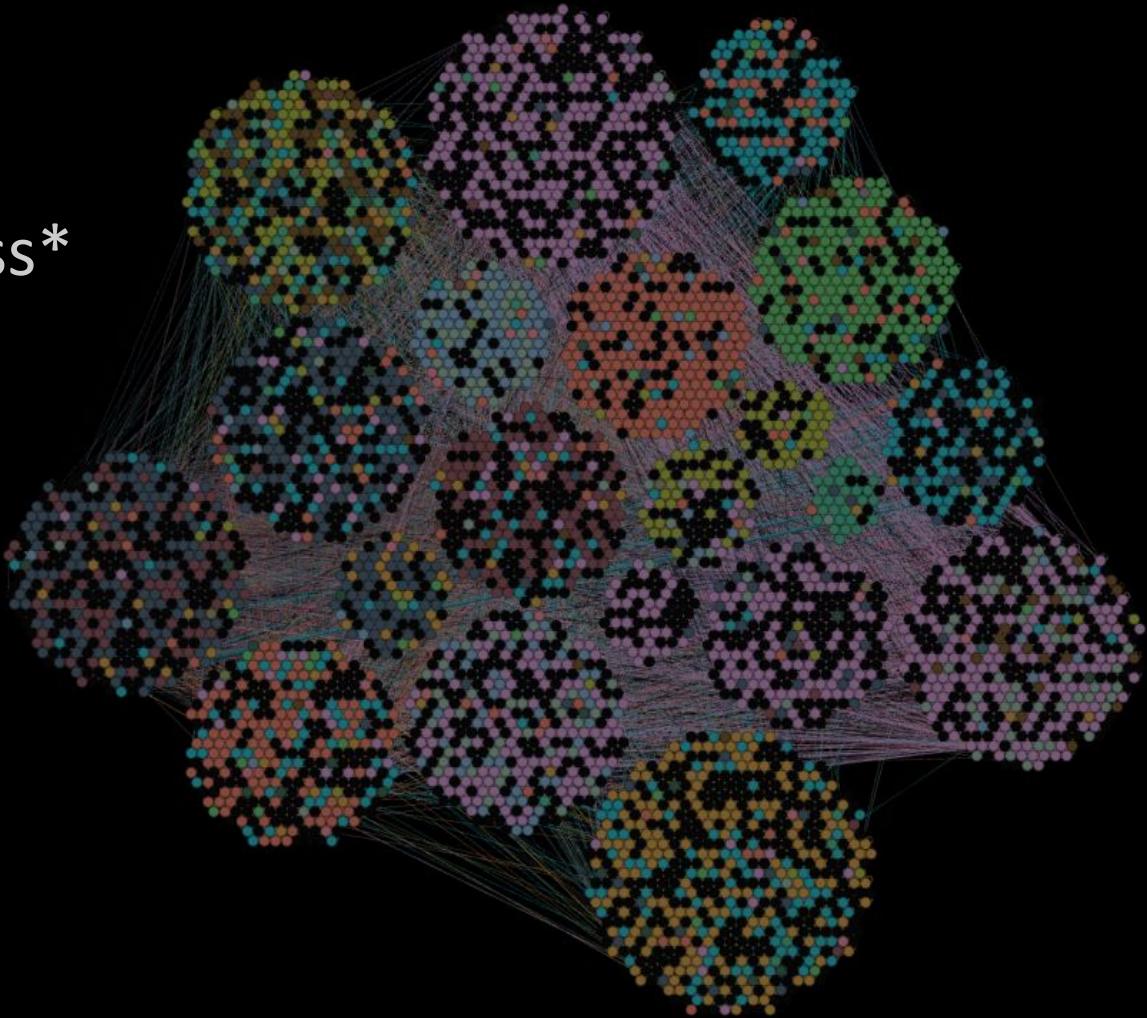
- K = 5
 - kolorem czerwonym
oznaczono osoby
należące do kilku
grup



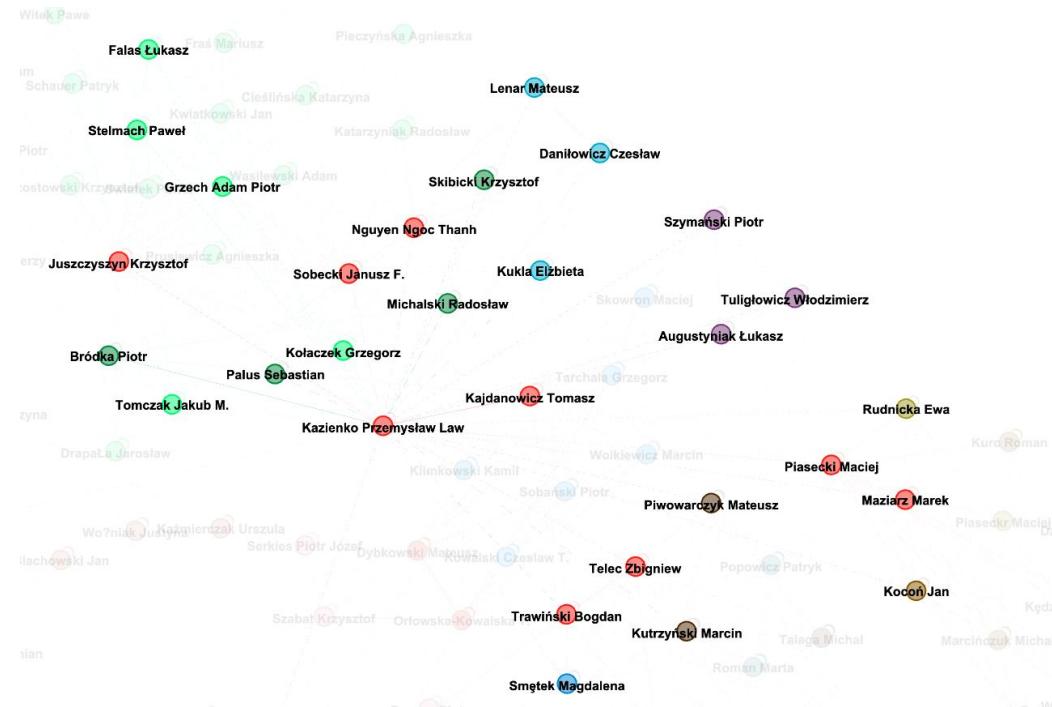
Wydziały, a Modularity class*

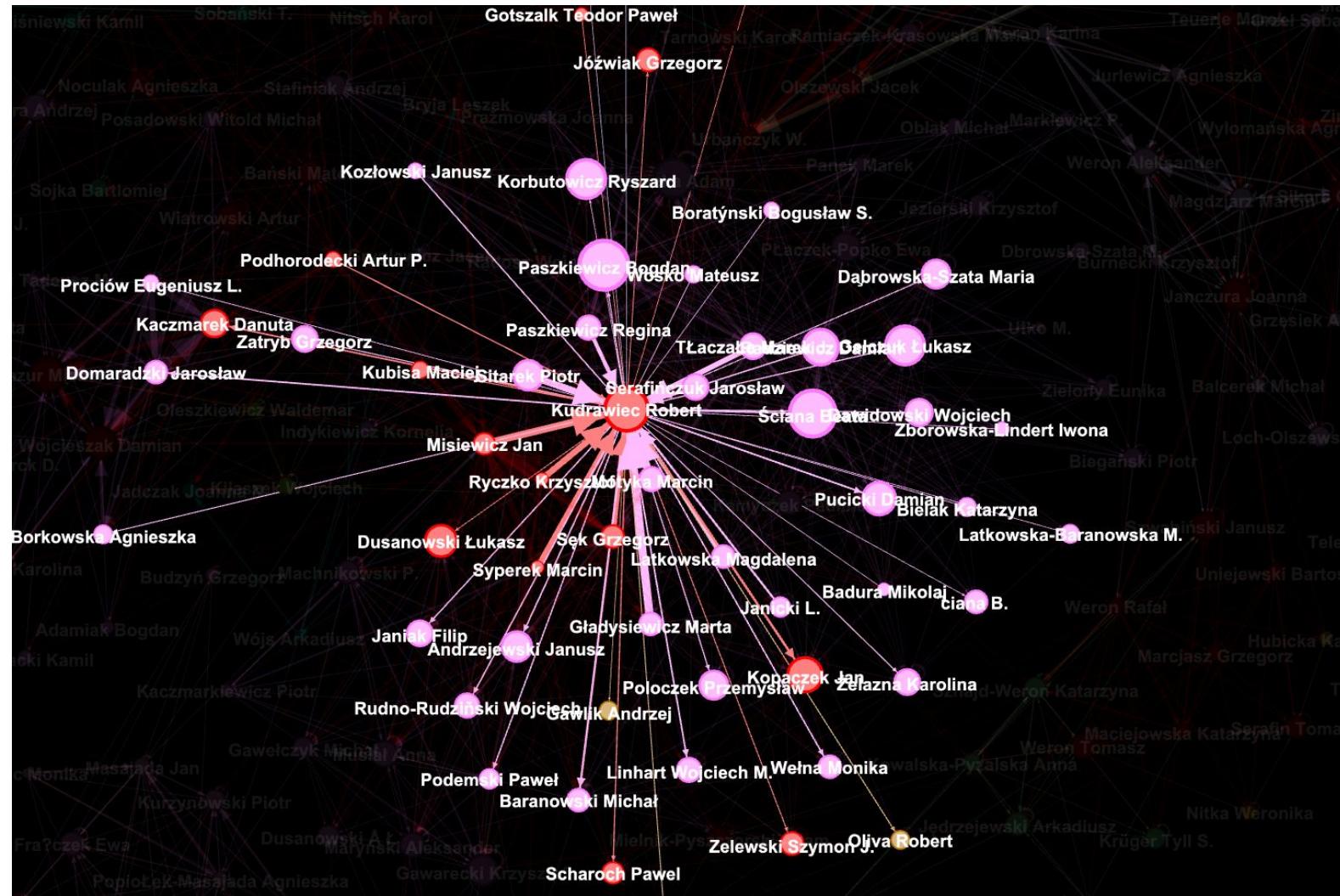
*Louvain

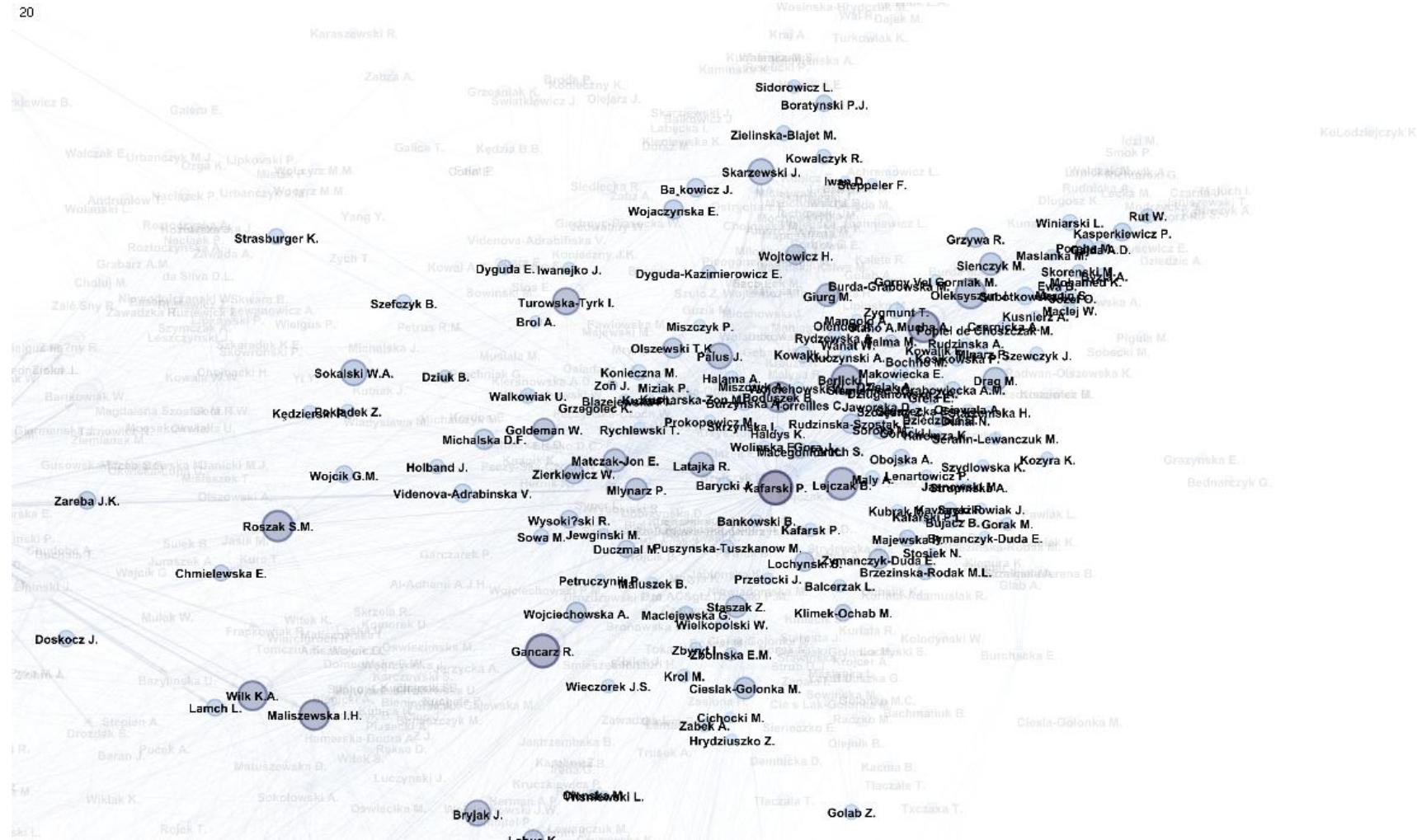
- Unknown
- W03
- W04
- W10
- W11
- W08
- W02
- W12
- W07
- W05
- W09
- W06
- W13
- PRD
- W01
- AC
- PRN
- W16
- W15
- null

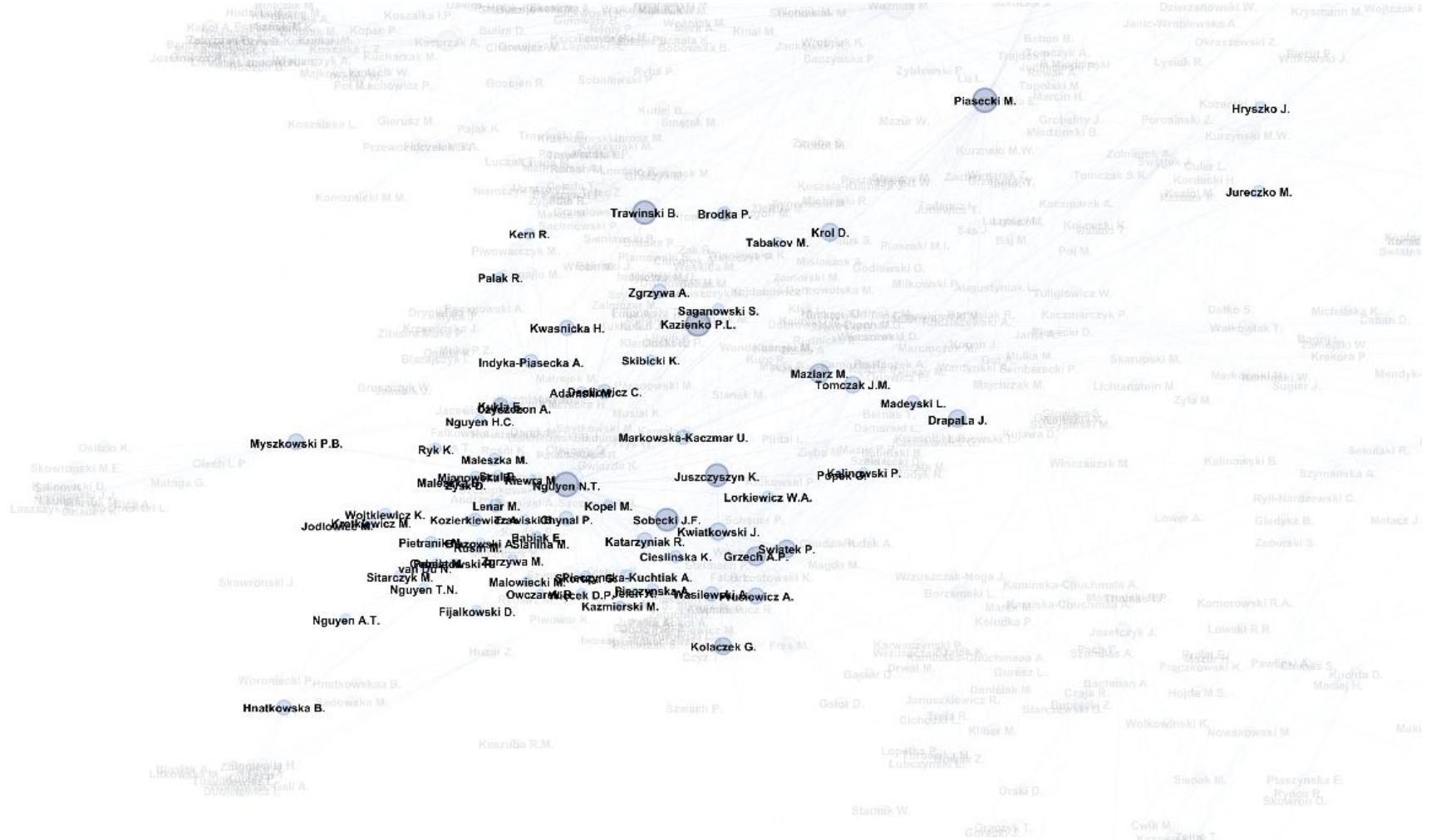


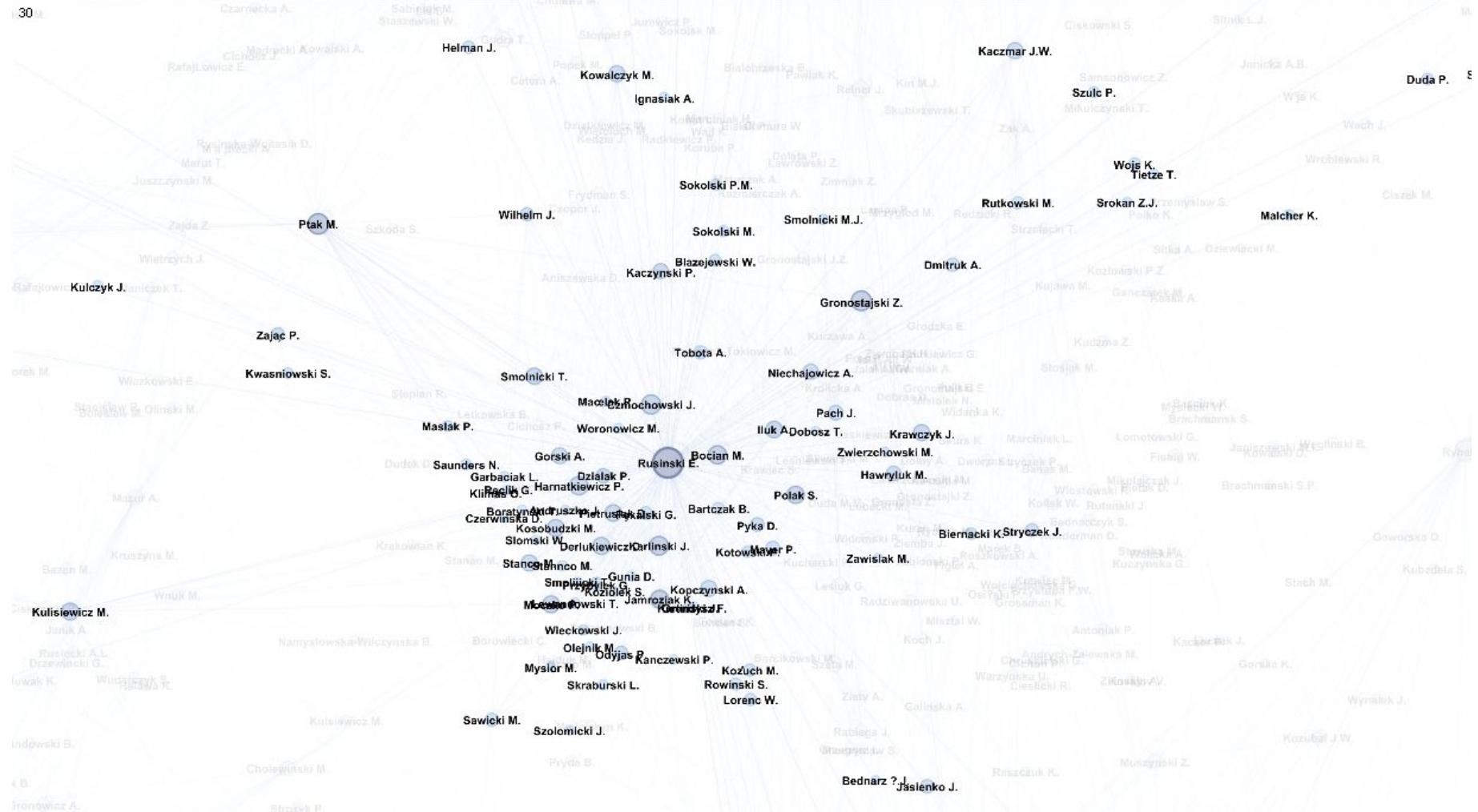
CFinder











Prototyp/Prezentacja wyników

- <https://frugile.github.io/onos-pwr/>
- Strona internetowa utworzona z wykorzystaniem GitHub Pages z Jekyll theme
 - Podstrony z interaktywnymi grafami wygenerowane w Gephi SigmaJS
 - Integracja z wynikami grupy analizującej sieć współautorstwa
- Umożliwia intuicyjne przeglądanie wygenerowanych sieci

Podsumowanie

- Zastosowanie ONoS we własnym projekcie na rzeczywistych danych pozwoliło lepiej poznać zagadnienia kursu i pokazało szerokie zastosowania tej dziedziny
- Algorytmy grupowania umożliwiły wydobycie bardzo wartościowej wiedzy z powstałej sieci
- Ważną częścią projektu był dobry wybór źródła danych i odpowiednie przetworzenie tych danych w celu uzyskania dokładnych analiz
- Interaktywna strona może być przydatnym narzędziem dla studentów i pracowników do wyszukiwania zależności pomiędzy pracownikami
- Głównym tematem tego projektu były cytowania na PWr, ale jak to się przekłada na cytowania na zewnątrz uczelni?



Dziękujemy za uwagę!

Podsumowanie projektu

Celem projektu “PWr cite” było zbadanie sieci cytowań pomiędzy pracownikami PWr. Sieć powstała na podstawie danych o publikacjach z bazy Scopus. Powstała sieć pozwoliła na przeanalizowanie, w jaki sposób dorobek naukowy pracowników jest wykorzystywany w ramach wydziałów i pomiędzy nimi. Wyniki zostały udostępnione na stronie internetowej frugile.github.io/onos-pwr/cite, gdzie dostępne są interaktywne grafy.

The goal of the “PWr cite” project was to examine the network of citations between WUST employees. Data required for network construction was obtained from Scopus database. The results allowed to analyze if and how employees’ work is used within departments and between them. The results are shared in the form of interactive graphs on this web page: frugile.github.io/onos-pwr/cite