

ІДЕНТИФІКАЦІЯ РІВНЯ СПОРІДНЕНОСТІ НАУКОВИХ СПЕЦІАЛЬНОСТЕЙ НА ОСНОВІ ДАНИХ СИСТЕМИ DIMENSIONS

Ідентифіковано рівні спорідненості наукових спеціальностей у межах Австралійсько–Новозеландської стандартної класифікації наук ANZCRC-2020. Ідентифікація здійснена з використанням інформаційної системи Dimensions шляхом аналізу 33.8 млн публікацій за 2019–2023 рр. Рівень спорідненості оцінено за індексом Жаккара. Встановлено, що із 14535 можливих пар спеціальностей, лише 131 пара має значущу спорідненість, за індексом Жаккара, що перевищує 0.05. З них для 20 пар спеціальностей рівень спорідненості є високим, а для 61 пари – середнім.

Ключові слова: ідентифікація, класифікація наук, спорідненість спеціальностей, аналіз даних, індекс Жаккара, наукові публікації, підбір рецензентів, наукометрія, Dimensions, ANZSRC-2020.

Вступ

Управління науковою діяльністю здійснюється в рамках деякої системи класифікації наук. В Україні одночасно діє дві системи класифікації наук: трирівнева та дворівнева. Трирівнева система формалізована «Переліком наукових спеціальностей», який складається з 27 галузей та 488 спеціальностей. Галузь утворюють від 3 до 133 спеціальностей. В деяких галузях спеціальності об'єднані у групи, наприклад, в галузі «Технічні науки» утворено 20 груп спеціальностей. З 2015 р. діє і дворівнева система класифікації освітньо-наукової діяльності, яка формалізована «Переліком галузей знань і спеціальностей, за якими здійснюється підготовка здобувачів вищої освіти». Відповідно до його поточної версії є 28 галузей знань, кожна з яких містить від 1 до 9 спеціальностей. Загальна кількість спеціальностей дорівнює 121. Є багато інших систем класифікації наук, як національного рівня, так і міжнародного. Поміж національних систем виділимо трирівневу Австралійсько–Новозеландську класифікацію ANZCRC-2020 [1], яка включає 22 галузі та 171 спеціальність та класифікацію Організації економічного співробітництва та розвитку [2], яка включає 6 галузей та 42 спеціальності.

Вдала система класифікації наук дозволяє краще зрозуміти особливості філософських процесів пізнання у різних науках, дослідити історичний розвиток та взаємодію різних галузей знань, підвищити ефективність пошуку документів і нау-

кової інформації та удосконалити адміністрування та управління дослідженнями [3].

Метою статі є ідентифікація поточного рівня спорідненості наукових спеціальностей. Оцінки рівня спорідненості спеціальностей необхідні для ухвалення рішень під час вирішення таких задач як:

- удосконалення системи класифікації наук, як-от, об'єднання спеціальностей або зміна їхньої галузевої належності;
- виявлення схожих наукових та освітніх установ для налагодження нових кооперативних зв'язків або проведення їх реорганізації;
- автоматизація підбору рецензентів для експертизи дисертацій, рукописів статей, заявок на гранти тощо;
- формування міждисциплінарних наукових досліджень та освітніх програм.

Огляд літератури та ідея дослідження

Найчастіше ідентифікація спорідненості наукових спеціальностей здійснюється шляхом аналізу цитування [4]. Ідея оцінювання спорідненості за цитованістю полягає в тому, що спорідненість між спеціальностями збільшується, якщо в статті за однією спеціальністю є посилання на статтю з іншої спеціальності. Також використовуються і методи статистичного аналізу тексту, зокрема, в [5] запропоновано

лінгвістичний підхід для дослідження організації та еволюції наукових галузей у Web of Science. За лінгвістичним підходом спорідненість визначається частотами появи слів у контексті конкретних спеціальностей. У роботі [5] порівнюються 3 підходи до визначення спорідненості спеціальностей: на основі експертної класифікації; на основі цитувань; на основі лінгвістичної схожості. Спорідненість оцінюється за метрикою розбіжності (*dissimilarity*) – чим менше значення метрики, тим більша спорідненість і навпаки. Експерименти здійснено для статей з Web of Science, що розмічені за трирівневою системою класифікації наук. Виявлено, що підходи на основі цитувань та лінгвістичного аналізу дають подібні значення спорідненості, водночас вони значно відрізняються від експертних оцінок. Експертна оцінка дає ідеалізоване уявлення про спорідненість спеціальностей, в той час як аналіз цитувань дозволяє виявити соціальну спорідненість спеціальностей, а лінгвістичний підхід дозволяє виявити змістовну (когнітивну) спорідненість. В [6] рівень спорідненості ідентифікують за кількістю цитувань зі статті певного дослідника на статті у журналах з Web of Science, де кожен із журналів віднесено до однієї із предметних областей Web of Science. В [7] ідентифікація здійснюється за допомогою показників різноманітності за розподілом посилань між спеціальностями з урахуванням відстані між ними. Деякі дослідження [8, 9] використовують одночасно аналіз цитувань та лінгвістичний аналіз для виявлення спорідненості. Дещо менш поширеним підходом є ідентифікація за аналізом графу колаборацій, який описує належність співавторів до наукових спеціальностей [10]. Як математична модель використовується індекс різноманітності Стірлінга.

Усі вищезгадані підходи вимагають опрацювання великих інформаційних ресурсів. Окрім того, методи, що базуються на аналізі цитування, є досить інерційними. Неможливо миттєво оцінити міждисциплінарність нової роботи, оскільки для її цитування потрібен певний час. На противагу згаданим підходам, у статті [11] запропоновано швидкий метод оцінювання спорідне-

ності наукових спеціальностей відповідно до системи класифікації ANZSRC-2008. Метод оснований на текстовому аналізі з використанням сервісів інформаційної системи Dimensions. Розрахунок спорідненості спеціальностей здійснюється за індексом Жаккара як відношення кількості спільних публікацій двох спеціальностей до загальної кількості публікацій за цими двома спеціальностями. Нещодавно Dimensions перейшла на оновлену систему класифікації ANZSRC-2020 [1]. Також база публікацій значно оновилася. Відповідно отримані в [11] оцінки рівня спорідненості втратили актуальність. Тому нижче здійснюється ідентифікація спорідненості наукових спеціальностей за методом [11] на новій джерельній базі і за новою системою класифікації наук. Ідентифікація здійснюється за публікаціями за період 2019–2023 рр.

Початкові дані для ідентифікації спорідненості спеціальностей

ANZSRC-2020 - це трирівнева класифікація наук за схемою: галузі (Divisions), спеціальності (Groups) та області (Fields). Уся наука розділена на такі 22 галузі:

- 30 Agricultural, Veterinary and Food Sciences;
- 31 Biological Sciences;
- 32 Biomedical and Clinical Sciences;
- 33 Built Environment and Design;
- 35 Commerce, Management, Tourism and Services;
- 34 Chemical Sciences;
- 36 Creative Arts and Writing;
- 37 Earth Sciences;
- 38 Economics;
- 39 Education;
- 40 Engineering;
- 41 Environmental Sciences;
- 47 Language, Communication and Culture;
- 48 Law and Legal Studies;
- 42 Health Sciences;
- 43 History, Heritage and Archaeology;
- 44 Human Society;
- 46 Information and Computing Sciences;
- 49 Mathematical Sciences;
- 50 Philosophy and Religious Studies;

- 51 Physical Sciences;
- 52 Psychology.
- Кожна галузь об'єднує від 3 до 19 спеціальностей. Назва спеціальності складається з цифрового коду та змістовної частини. Код складається з номера галузі та порядкового номера спеціальності в межах галузі. Всього є 171 спеціальність, список наведено нижче:
- 3001 Agricultural Biotechnology;
- 3002 Agriculture, Land and Farm Management;
- 3003 Animal Production;
- 3004 Crop and Pasture Production;
- 3005 Fisheries Sciences;
- 3006 Food Sciences;
- 3007 Forestry Sciences;
- 3008 Horticultural Production;
- 3009 Veterinary Sciences;
- 3101 Biochemistry and Cell Biology;
- 3102 Bioinformatics and Computational Biology;
- 3103 Ecology;
- 3104 Evolutionary Biology;
- 3105 Genetics;
- 3106 Industrial Biotechnology;
- 3107 Microbiology;
- 3108 Plant Biology;
- 3109 Zoology;
- 3201 Cardiovascular Medicine and Haematology;
- 3202 Clinical Sciences;
- 3203 Dentistry;
- 3204 Immunology;
- 3205 Medical Biochemistry and Metabolomics;
- 3206 Medical Biotechnology;
- 3207 Medical Microbiology;
- 3208 Medical Physiology;
- 3209 Neurosciences;
- 3210 Nutrition and Dietetics;
- 3211 Oncology and Carcinogenesis;
- 3212 Ophthalmology and Optometry;
- 3213 Pediatrics;
- 3214 Pharmacology and Pharmaceutical Sciences;
- 3215 Reproductive Medicine;
- 3301 Architecture;
- 3302 Building;
- 3303 Design;
- 3304 Urban and Regional Planning;
- 3401 Analytical Chemistry;
- 3402 Inorganic Chemistry;
- 3403 Macromolecular and Materials Chemistry;
- 3404 Medicinal and Biomolecular Chemistry;
- 3405 Organic Chemistry;
- 3406 Physical Chemistry;
- 3407 Theoretical and Computational Chemistry;
- 3501 Accounting, Auditing and Accountability;
- 3502 Banking, Finance and Investment;
- 3503 Business Systems in Context;
- 3504 Commercial Services;
- 3505 Human Resources and Industrial Relations;
- 3506 Marketing;
- 3507 Strategy, Management and Organisational Behaviour;
- 3508 Tourism;
- 3509 Transportation, Logistics and Supply Chains;
- 3601 Art History, Theory and Criticism;
- 3602 Creative and Professional Writing;
- 3603 Music;
- 3604 Performing Arts;
- 3605 Screen and Digital Media;
- 3606 Visual Arts;
- 3701 Atmospheric Sciences;
- 3702 Climate Change Science;
- 3703 Geochemistry;
- 3704 Geoinformatics;
- 3705 Geology;
- 3706 Geophysics;
- 3707 Hydrology;
- 3708 Oceanography;
- 3709 Physical Geography and Environmental Geoscience;
- 3801 Applied Economics;
- 3802 Econometrics;
- 3803 Economic Theory;
- 3901 Curriculum and Pedagogy;
- 3902 Education Policy, Sociology and Philosophy;
- 3903 Education Systems;
- 3904 Specialist Studies in Education;
- 4001 Aerospace Engineering;
- 4002 Automotive Engineering;
- 4003 Biomedical Engineering;
- 4004 Chemical Engineering;
- 4005 Civil Engineering;
- 4006 Communications Engineering;

- | | | | |
|------|--|------|---|
| 4007 | Control Engineering, Mechatronics and Robotics; | 4605 | Data Management and Data Science; |
| 4008 | Electrical Engineering; | 4606 | Distributed Computing and Systems Software; |
| 4009 | Electronics, Sensors and Digital Hardware; | 4607 | Graphics, Augmented Reality and Games; |
| 4010 | Engineering Practice and Education; | 4608 | Human-Centred Computing; |
| 4011 | Environmental Engineering; | 4609 | Information Systems; |
| 4012 | Fluid Mechanics and Thermal Engineering; | 4610 | Library and Information Studies; |
| 4013 | Geomatic Engineering; | 4611 | Machine Learning; |
| 4014 | Manufacturing Engineering; | 4612 | Software Engineering; |
| 4015 | Maritime Engineering; | 4613 | Theory of Computation; |
| 4016 | Materials Engineering; | 4701 | Communication and Media Studies; |
| 4017 | Mechanical Engineering; | 4702 | Cultural Studies; |
| 4018 | Nanotechnology; | 4703 | Language Studies; |
| 4019 | Resources Engineering and Extractive Metallurgy; | 4704 | Linguistics; |
| 4101 | Climate Change Impacts and Adaptation; | 4705 | Literary Studies; |
| 4102 | Ecological Applications; | 4801 | Commercial Law; |
| 4103 | Environmental Biotechnology; | 4802 | Environmental and Resources Law; |
| 4104 | Environmental Management; | 4803 | International and Comparative Law; |
| 4105 | Pollution and Contamination; | 4804 | Law in Context; |
| 4106 | Soil Sciences; | 4805 | Legal Systems; |
| 4201 | Allied Health and Rehabilitation Science; | 4806 | Private Law and Civil Obligations; |
| 4202 | Epidemiology; | 4807 | Public Law; |
| 4203 | Health Services and Systems; | 4901 | Applied Mathematics; |
| 4204 | Midwifery; | 4902 | Mathematical Physics; |
| 4205 | Nursing; | 4903 | Numerical and Computational Mathematics; |
| 4206 | Public Health; | 4904 | Pure Mathematics; |
| 4207 | Sports Science and Exercise; | 4905 | Statistics; |
| 4208 | Traditional, Complementary and Integrative Medicine; | 5001 | Applied Ethics; |
| 4301 | Archaeology; | 5002 | History and Philosophy of Specific Fields; |
| 4302 | Heritage, Archive and Museum Studies; | 5003 | Philosophy; |
| 4303 | Historical Studies; | 5004 | Religious Studies; |
| 4401 | Anthropology; | 5005 | Theology; |
| 4402 | Criminology; | 5101 | Astronomical Sciences; |
| 4403 | Demography; | 5102 | Atomic, Molecular and Optical Physics; |
| 4404 | Development Studies; | 5103 | Classical Physics; |
| 4405 | Gender Studies; | 5104 | Condensed Matter Physics; |
| 4406 | Human Geography; | 5105 | Medical and Biological Physics; |
| 4407 | Policy and Administration; | 5106 | Nuclear and Plasma Physics; |
| 4408 | Political Science; | 5107 | Particle and High Energy Physics; |
| 4409 | Social Work; | 5108 | Quantum Physics; |
| 4410 | Sociology; | 5109 | Space Sciences; |
| 4601 | Applied Computing; | 5110 | Synchrotrons and Accelerators; |
| 4602 | Artificial Intelligence; | 5201 | Applied and Developmental Psychology; |
| 4603 | Computer Vision and Multimedia Computation; | 5202 | Biological Psychology; |
| 4604 | Cybersecurity and Privacy; | 5203 | Clinical and Health Psychology; |
| | | 5204 | Cognitive and Computational Psychology; |
| | | 5205 | Social and Personality Psychology. |

На сьогодні в системі Dimensions проіндексовано понад 140 млн наукових публікацій. Біля 80% від усіх публікацій категоризовано за спеціальностями, тобто віднесено до однієї чи кількох спеціальностей [1]. Категоризація публікацій у Dimensions здійснена на основі машинного навчання, переважно за аналізом змісту назв, анотацій та ключових слів. За останні 5 років у системі Dimensions проіндексовано 33.8 млн публікацій, які є джерельною базою дослідження. Публікації за спеціальностями розподілені нерівномірно (рис. 1). Децильний коефіцієнт дорівнює 14.9. Найбільше публікацій – 3188112 віднесено до спеціальності 3202, найменше – 3339 – до спеціальності 3606. Навіть найменш популярна спеціальність має достат-

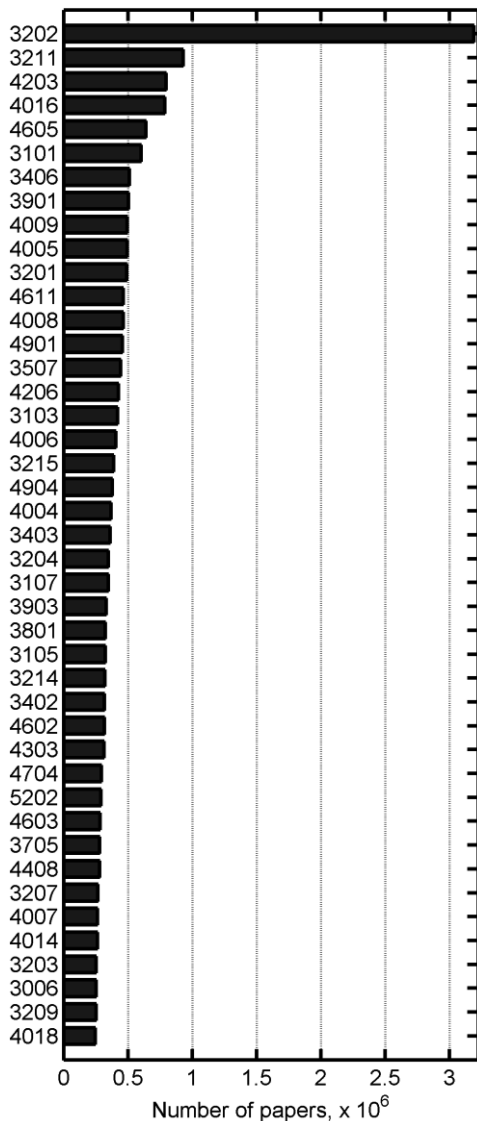


Рис. 1. Перший квартиль розподілу публікацій за спеціальностями за 2019–2023 рр.

ню кількість публікацій, щоб на їх основі отримати достовірні статистичні висновки.

Ідентифікація спорідненості спеціальностей

Індекс спорідненості пари спеціальностей (A, B) розраховується таким чином [11]:

$$J(A, B) = \frac{N_{A \cap B}}{N_A + N_B - N_{A \cap B}}, \quad (1)$$

де N_A – кількість публікацій за спеціальністю A ;

N_B – кількість публікацій за спеціальністю B ;

$N_{A \cap B}$ – кількість публікацій, які одночасно віднесено як до спеціальності A , так і до спеціальності B .

Проілюструємо застосування формули (1) на прикладі розрахунку спорідненості спеціальностей 4602 та 4608. За спеціальністю 4602 є 316910 публікацій, за спеціальністю 4608 – 239664 публікації. За обома спеціальностями одночасно категоризовано 17757 публікацій. Відповідно рівень спорідненості спеціальностей 4602 та 4608 становить: $J(4602, 4608) = \frac{17757}{316910 + 239664 - 17757} = 0.03$.

Dimensions можна за API [12]. Дані можна отримати використовуючи звичайний HTTP запит, де тілом є текст у форматі Dimensions Search Language [13]. Наприклад, для того, щоб отримати розподіл кількості публікацій за кожною спеціальністю за останні 5 років необхідно відправити запит із таким тілом:

```
search publications
where year in
[2019, 2020, 2021, 2022, 2023]
return category_for_2020
```

Результат повертається у JSON-форматі, в якому для кожної спеціальності вказано кількість публікацій, категоризованих до відповідної спеціальності. Для того, щоб знайти кількість публікацій віднесених одночасно до двох спеціальностей у запиті необхідно додатково вказати ідентифікатор спеціальності. Як відповідь повертається кількість публікацій, які категоризовані одночасно до вказаної спеціальності і

до кожної іншої спеціальності. Приклад такого запиту для спеціальності 3202, яка має внутрішній ідентифікатор 80045, наведено нижче:

```
search publications
where year in
[2019, 2020, 2021, 2022, 2023]
and category_for.id=80045
return category_for_2020
```

Надіславши API-запит для кожної спеціальності отримуємо необхідні початкові дані для розрахунку рівня спорідненості всіх пар спеціальностей. Ранговий розподіл пар спеціальностей за індексом Жаккара (1) наведено на рис. 2. Високу спорідненість має 20 пар спеціальностей, індекс Жаккара для яких перевищує 0.2 (табл. 1). Поміж 20 сильно споріднених пар 3 пари утворено спеціальностями з різних галузей знань. Середня спорідненість має місце для 41 пари спеціальностей (табл. 2). Поміж 41 пари із середньою спорідненістю 9 пар утворено спеціальностями з різних галузей знань. Спорідненість пар із високими та середніми індексами Жаккара варто врахувати під час автоматичного підбору рецензентів чи виявлення схожих наукових та освітніх установ. Можливо, деякий ефект буде і від врахування пар спеціальностей із низьким рівнем спорідненості. Таких пар виявилось 70. Решта пар спеціальностей, а саме $14535 - 131 = 14404$ має шумову спорідненість – їхній індекс Жаккара менше 0.05. Поріг щодо шуму встановлено на підставі обчислювальних експериментів [14].

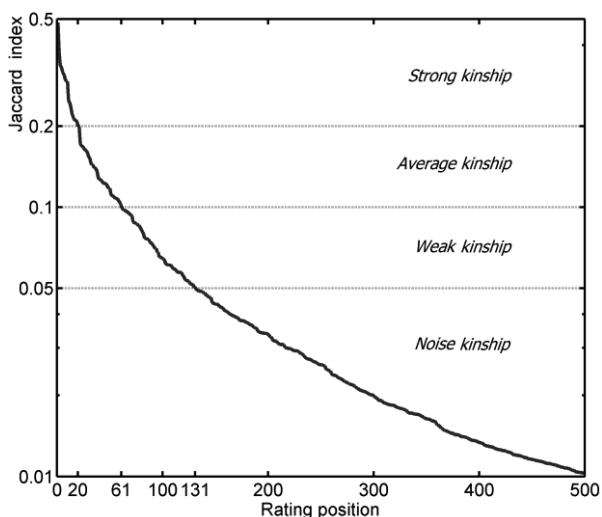


Рис. 2. Фрагмент рангового розподілу спорідненості пар спеціальностей (напівлогарифмічний формат)

Таблиця 1. Спеціальності з сильною спорідненістю

Спеціальності	Індекс Жаккара
3504 3508	0.485
4703 4704	0.378
3506 3508	0.339
3504 3506	0.332
4008 4009	0.322
4002 4017	0.313
3002 3004	0.309
3901 3903	0.297
5106 5107	0.294
4402 4805	0.293
3503 3506	0.247
4901 4904	0.245
4007 4010	0.236
5106 5110	0.224
3705 3706	0.218
3703 3706	0.212
3703 3705	0.211
4902 5107	0.21
3605 4701	0.208
4004 4011	0.2

Таблиця 2. Спеціальності з середньою спорідненістю

Спеціальності	Індекс Жаккара
3802 3803	0.199
3302 4005	0.172
3502 3801	0.169
4901 4903	0.168
3501 3502	0.167
3801 3802	0.164
5102 5108	0.163
5004 5005	0.162
4902 4904	0.159
3402 3405	0.154
4006 4613	0.152
4203 4205	0.146
4604 4606	0.144
3103 3104	0.144
5201 5205	0.14
4803 4807	0.14
3304 3509	0.139
3102 3105	0.138
3209 5202	0.129
3705 3709	0.127
4407 4408	0.127
5201 5203	0.126
4605 4606	0.124
3801 3803	0.123
4006 4009	0.123

Спеціальності		Індекс Жаккара
4602	4605	0.122
5101	5109	0.122
3902	3903	0.119
3204	3211	0.118
4804	4807	0.118
3404	3405	0.112
3403	4016	0.111
3406	4016	0.11

Спеціальності		Індекс Жаккара
5202	5204	0.109
4903	4904	0.108
4803	4804	0.108
3702	3709	0.107
3503	4609	0.106
4301	4303	0.105
5002	5003	0.104
3107	3207	0.101

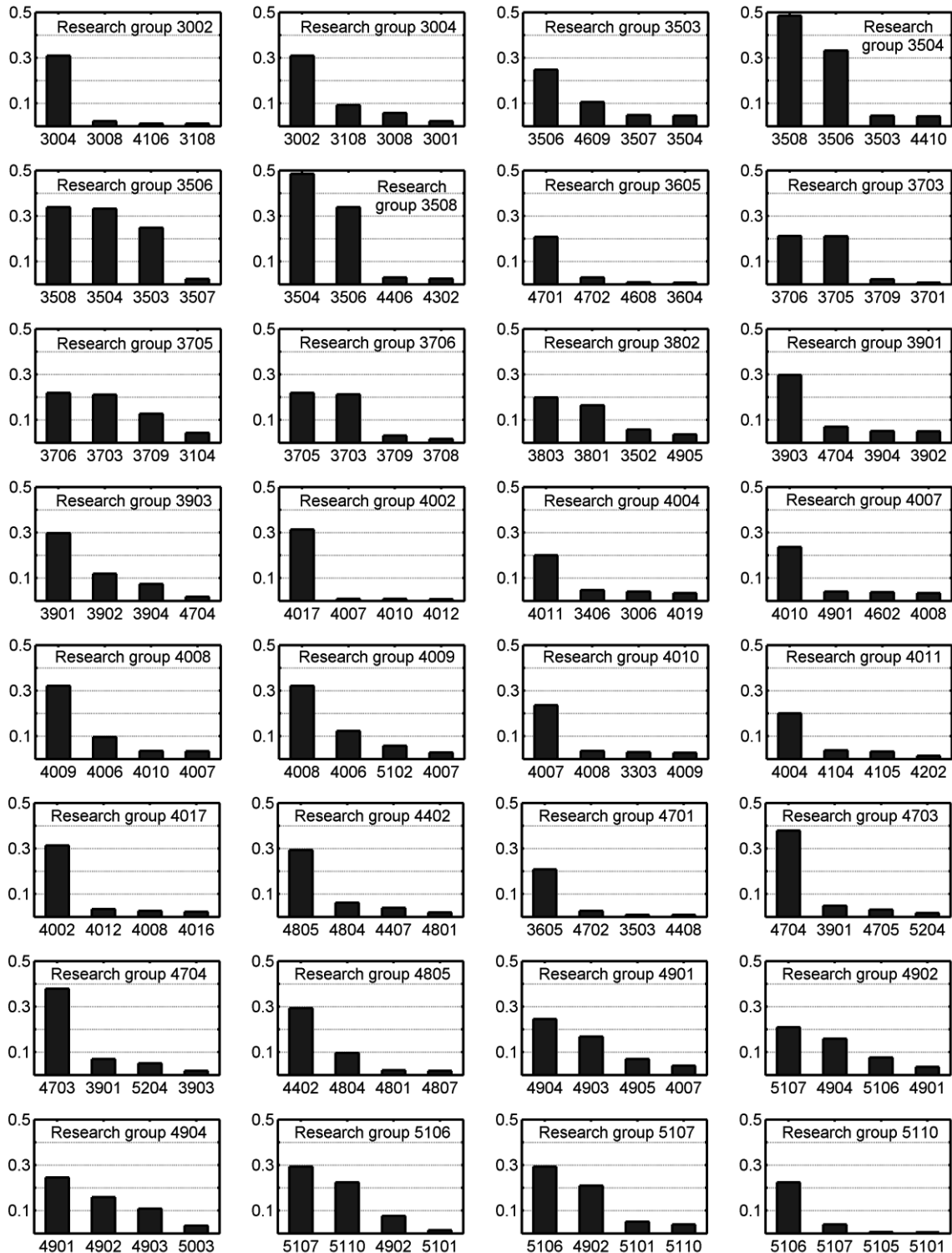


Рис. 3. Діаграми спорідненості спеціальностей з топ-20 за індексом Жаккара

На рис. 3 наведені діаграми спорідненості для 32 спеціальностей, які утворили пари з максимальними індексами Жаккара. Більшість із них мають сильну спорідненість лише з однією спеціальністю. Водночас, 7 спеціальностей сильно взаємодіють з двома спеціальностями. Спеціальність 3506 має високу спорідненість аж із трьома спеціальностями – 3504, 3508 та 3503. Усі вони належать до спільної галузі 35 – Commerce, Management, Tourism and Services.

Висновки

Ідентифіковано рівні спорідненості спеціальностей за системою класифікації наукових спеціальностей ANZSRC-2020. Ідентифікація здійснена з використанням інформаційної системи Dimensions шляхом аналізу 33.8 млн публікацій за 2019–2023 рр. Рівень спорідненості оцінено за індексом Жаккара. Встановлено, що із 14535 можливих пар спеціальностей, лише 131 пара має значнішу спорідненість з індексом Жаккара, що перевищує 0.05. З них для 20 пар спеціальностей рівень спорідненості є високим, а для 61 пари – середнім. Поміж пар із високою спорідненістю переважають спеціальності з однакових галузей, але 3 пари із 20 утворено із спеціальностей з різних галузей.

Отримані оцінки рівня спорідненості спеціальностей можуть бути використані для покращення розв'язання задач автоматичного призначення рецензентів дисертацій, запитів на гранти, рукописів статей, для створення міждисциплінарних PhD-програм, удосконалення системи класифікації наук, відслідковування трендів міждисциплінарних досліджень тощо.

Подяки

Автори висловлюють подяку Digital Science & Research Solutions Inc. за надання доступу до ресурсів Dimensions за проектом DIM-371.

Література

1. Porter, S. J., Hawizy, L., & Hook, D. W. (2023). Recategorising research: Mapping from FoR 2008 to FoR 2020 in Dimensions.

- Quantitative Science Studies*, 4(1), 127–143. <https://doi.org/10.1162/qss.a.00244>.
2. *Frascati Manual 2015*. Frascati Manual 2015. OECD. <https://doi.org/10.1787/9789264268111-ko>.
3. Legendre, A. (2019). The development of the Canadian research and development classification. *Knowledge Organization*. International Society for Knowledge Organization. <https://doi.org/10.5771/0943-7444-2019-5-371>.
4. Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, K. W., Keyton, J., Rafols I., Börner, K. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics*, 5(1), 14–26. <https://doi.org/10.1016/j.joi.2010.06.004>.
5. Dias, L., Gerlach, M., Scharloth, J., & Altmann, E. G. (2018). Using text analysis to quantify the similarity and evolution of scientific disciplines. *Royal Society Open Science*, 5(1). <https://doi.org/10.1098/rsos.171545>.
6. Porter, A. L., Cohen, A. S., David Roessner, J., & Perreault, M. (2007). Measuring researcher interdisciplinarity. *Scientometrics*, 72(1), 117–147. <https://doi.org/10.1007/s11192-007-1700-5>.
7. Van Noorden, R. (2015, September 16). Interdisciplinary research by the numbers. *Nature*. Nature Publishing Group. <https://doi.org/10.1038/525306a>.
8. Braam, R. R., Moed, H. F., & van Raan, A. F. J. (1991). Mapping of science by combined co-citation and word analysis. II: Dynamical aspects. *Journal of the American Society for Information Science*, 42(4), 252–266. [https://doi.org/10.1002/\(SICI\)1097-4571\(199105\)42:4<252::AID-ASI2>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-4571(199105)42:4<252::AID-ASI2>3.0.CO;2-G).
9. Silva, F. N., Amancio, D. R., Bardosova, M., Costa, L. da F., & Oliveira, O. N. (2016). Using network science and text analytics to produce surveys in a scientific topic. *Journal of Informetrics*, 10(2), 487–502. <https://doi.org/10.1016/j.joi.2016.03.008>.
10. Karlovčec, M., & Mladenčić, D. (2015). Interdisciplinarity of scientific fields and its evolution based on graph of project collaboration and co-authoring. *Scientometrics*, 102(1), 433–454. <https://doi.org/10.1007/s11192-014-1355-y>.
11. Shtovba, S., & Petrychko, M. (2019). Jaccard index-based assessing the similarity of

- research fields in dimensions. In CEUR Workshop Proceedings (Vol. 2533, pp. 117–128). CEUR-WS.
12. Dimensions API request. Електронний ресурс. Режим доступу: <https://www.dimensions.ai/products/all-products/dimensions-api/> (20.02.2024).
 13. Dimensions DSL. Електронний ресурс. Режим доступу: <https://docs.dimensions.ai/dsl/> (20.02.2024).
 14. Shtovba S., Petrychko M., Shtovba O. Similarity metric of categorical distributions for topic modeling problems with akin categories // CEUR Workshop Proceedings, Vol. 3392 “Proc. of the Sixth International Workshop on Computer Modeling and Intelligent Systems”. – 2023. – P. 76-85. DOI: <https://doi.org/10.32782/cmis/3392-7>.

Одержано: 22.02.2024

Про авторів:

Штовба Сергій Дмитрович,
професор, д. т. н., професор кафедри
інформаційних технологій
Донецького національного університету
імені Василя Стуса.

Кількість наукових публікацій
в українських виданнях – понад 100.
Кількість наукових публікацій
в іноземних виданнях – понад 50.
Індекс Гірша – 8.
<https://orcid.org/0000-0003-1302-4899>

Петричко Микола Володимирович,
аспірант кафедри комп’ютерних
систем управління Вінницького національного
технічного університету.
Кількість наукових публікацій
в українських виданнях – 7.
Кількість наукових публікацій
в іноземних виданнях – 6.
Індекс Гірша – 3.
<https://orcid.org/0000-0001-6836-7843>

Місце роботи авторів:

Донецький національний університет
імені Василя Стуса,
600-річчя, 21, 21021, м. Вінниця
email: s.shtovba@donnu.edu.ua
Вінницький національний
технічний університет,
Хмельницьке шосе, 95,
21021, м.Вінниця
email: mpetrychko@vntu.edu.ua