

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БУДІВНИЦТВА І АРХІТЕКТУРИ

КУЧАНСЬКИЙ ОЛЕКСАНДР ЮРІЙОВИЧ



УДК 004.5:378.4

МЕТОДОЛОГІЯ ФОРМУВАННЯ ІНФОРМАЦІЙНИХ ПРОСТОРІВ СУБ'ЄКТІВ
НАУКОВОЇ ДІЯЛЬНОСТІ У СТАЛОМУ РОЗВИТКУ
ЗАКЛАДІВ ВИЩОЇ ОСВІТИ

Спеціальність 05.13.06 – інформаційні технології

Автореферат
дисертації на здобуття наукового ступеня
доктора технічних наук

Київ – 2021

Дисертацією є рукопис.

Робота виконана на кафедрі інформаційних технологій Київського національного університету будівництва і архітектури Міністерства освіти і науки України.

Науковий консультант: доктор технічних наук, професор
Білощицький Андрій Олександрович, професор кафедри інформаційних технологій Київського національного університету будівництва і архітектури, МОН України, м. Київ.

Офіційні опоненти: доктор технічних наук, професор
Криворучко Олена Володимирівна, завідувач кафедри інженерії програмного забезпечення та кібербезпеки Київського національного торговельно-економічного університету, МОН України, м. Київ;

доктор технічних наук, професор
Корж Роман Орестович, проректор з науково-педагогічної роботи Національного університету «Львівська політехніка», МОН України, м. Львів;

доктор технічних наук, професор
Антощук Світлана Григорівна, директор Інституту комп'ютерних систем Одеського національного політехнічного університету МОН України, м. Одеса.

Захист відбудеться «16» квітня 2021 року о 09:30 на засіданні спеціалізованої вченої ради Д 26.056.01 в Київському національному університеті будівництва і архітектури Міністерства освіти і науки України (03037, м. Київ, Повітрофлотський проспект, 31, ауд. 366).

З дисертацією можна ознайомитись у бібліотеці Київського національного університету будівництва і архітектури за адресою: 03037, м. Київ, Повітрофлотський проспект, 31.

Автореферат розісланий «15» березня 2021 р.

Вчений секретар
спеціалізованої вченої ради,
доктор технічних наук, доцент



М.І. Цюцюра

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми досліджень. Для вдосконалення наукової діяльності, стимулювання фундаментальних та прикладних досліджень, розвитку інновацій в Україні, для зміцнення та просування наукового потенціалу держави в світі, важливу роль відіграє розвиток моделей та методів організації наукових досліджень в напрямку створення процедур забезпечення формування раціонального складу та структури суб'єктів наукової діяльності, що спираються на об'єктивні показники продуктивності та нормативні принципи інтеграції української науки до світового наукового співтовариства.

Кожен суб'єкт наукової діяльності має власний інформаційний простір, в якому цей суб'єкт існує та трансформується, в якому накопичується і зберігається інформація стосовно його розвитку. Інформаційний простір суб'єкта наукової діяльності уявляється як шлейф з інформацією про результати наукової діяльності, оцінки продуктивності, цитування наукових публікацій тощо. Продуктивність суб'єктів наукової діяльності залежить від розуміння принципів формування і функціонування його інформаційного простору. Крім того, створення наукових або освітніх проєктів, які включають суб'єкти наукової діяльності, забезпечує синергетичну взаємодію їх інформаційних просторів, переходу цінностей з одного інформаційного простору в інший.

Ключовою передумовою сталості розвитку закладу вищої освіти та планування його науково-дослідної діяльності є забезпечення прозорого оцінювання результатів цієї діяльності. Проте, як правило, процедура оцінювання передбачає збір даних значного обсягу, що може містити неточну або ненадійну інформацію. Формування структури та складу наукових проєктів часто відбувається суб'єктивно, без врахування або з неповним врахуванням продуктивності суб'єктів наукової діяльності, що в цілому може здійснювати негативний вплив на сталість розвитку закладу вищої освіти та ефективність науково-технічного розвитку держави в цілому. Оскільки суб'єкти наукової діяльності динамічно розвиваються в своїх інформаційних просторах, то забезпечення сталості розвитку закладів вищої освіти, які формуються з суб'єктів наукової діяльності, визначається в тому числі керованістю їх інформаційних просторів. Це ускладнюється обмеженим ресурсним забезпеченням, що виділяється державою для проведення наукових досліджень. Таким чином, можна визначити актуальну проблему: створити методологію формування інформаційних просторів суб'єктів наукової діяльності, що є передумовою сталого розвитку закладів вищої освіти, і яка включатиме методи прозорого оцінювання продуктивності суб'єктів наукової діяльності та її аналізу в динаміці. Це дозволить державі стимулювати мотивацію суб'єктів наукової діяльності шляхом забезпечення раціонального фінансування найбільш продуктивних з них, що важливо в умовах обмеженого ресурсного забезпечення. Тому завдання забезпечення сталості функціонування закладів вищої освіти шляхом вивчення інформаційних просторів суб'єктів наукової діяльності, що формують заклад, є сьогодні надзвичайно актуальним.

Вагомий внесок у розвиток концептуальних та методологічних основ аналізу наукових мереж, мереж цитування та оцінювання продуктивності наукової діяльності

зробили такі українські та іноземні вчені: J. Hirsch, E. Garfield, L. Leydesdorff, B. Gipp, W. Glänzel, Z. Zhang, A.-L. Barabási, В.Ю. Биков, Г.М. Добров. Дослідження формування та розвитку команд проєктів, а також створенню ментальних просторів команд проєктів присвячені праці С.Д. Бушуєва та О.В. Веренич.

Створення методології формування інформаційних просторів суб'єктів наукової діяльності, вивчення взаємозв'язків між інформаційними просторами для забезпечення сталості функціонування закладів вищої освіти, є важливим для розроблення стратегічних рішень щодо розвитку організацій та підвищення наукового потенціалу в Україні в цілому. Отже, можна констатувати, що обрана тема дослідження є **актуальною** та необхідною на сучасному етапі розвитку освітньо-наукової сфери України.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконувалась відповідно до національної стратегії розвитку освіти в Україні на 2011-2027 роки, відповідно до плану науково-дослідних робіт Київського національного університету будівництва і архітектури та Київського національного університету імені Тараса Шевченка у межах державних бюджетних тем МОН України: науково-дослідна робота для Міністерства освіти та науки України «Розробка методів аналізу якості науково-дослідної роботи вчених, ЗВО МОН України та окремих структурних підрозділів», Київський національний університет імені Тараса Шевченка, № ДР 0119U100187 (2019-2021), керівник – Білощицький А.О., відп. вик. – Кучанський О.Ю.; «Розробка комбінованих методів ідентифікації неповних дублікатів та виявлення повноти висвітлення наукових результатів дисертаційних досліджень, опублікованих автором», Київський національний університет будівництва і архітектури, № ДР 0119U002579 (2019-2021), керівник – Лізунов П.П., відп. вик. – Кучанський О.Ю.; науково-дослідна робота для Міністерства освіти та науки України «Методологічні основи створення інформаційного середовища управління науковими дослідженнями структурних одиниць ЗВО МОН України», Київський національний університет будівництва і архітектури (номер державної реєстрації – 0115U000330, 2016 – 2017 рр.); науково-дослідна робота для Міністерства освіти та науки України «Створення системи перевірки ступеню унікальності наукових робіт», Київський національний університет будівництва і архітектури (номер державної реєстрації – 0114U000126, 2016 – 2017 рр.). Здобувачем розроблена методологія формування інформаційних просторів суб'єктів наукової діяльності та інформаційна технологія, що реалізує дану методологію.

Об'єктом дослідження є процеси, що формують методологію формування інформаційних просторів суб'єктів наукової діяльності в сталому розвитку закладів вищої освіти.

Предметом дослідження є методологія, що є основою для побудови інформаційної технології для підтримки сталого розвитку закладів вищої освіти.

Метою дослідження є розроблення методології формування інформаційних просторів суб'єктів наукової діяльності в сталому розвитку закладів вищої освіти.

Завдання дисертаційного дослідження:

1. Провести аналіз стану завдань формування інформаційних просторів суб'єктів наукової діяльності, кількісного оцінювання продуктивності суб'єктів

наукової діяльності, формування предметних наукових просторів, пошук партнерів для співпраці в рамках освітніх та наукових проєктів.

2. Розробити моделі ідентифікації індивідуальних та колективних суб'єктів наукової діяльності, а саме: науковців та закладів вищої освіти. Розробити формалізовану множинну модель інформаційного простору суб'єктів наукової діяльності, концептуальну схему їх взаємодії.

3. Розробити методологію формування інформаційних просторів суб'єктів наукової діяльності, включаючи відповідний понятійний апарат.

4. Розробити та верифікувати інформаційну технологію, яка реалізує методологію формування інформаційних просторів суб'єктів наукової діяльності. Описати модулі візуалізації та інтерфейсу користувача в програмному комплексі формування інформаційних просторів суб'єктів наукової діяльності.

5. Вдосконалити методи формування предметних наукових просторів на основі кластерного аналізу параметрів моделі ідентифікації науковця як суб'єкта наукової діяльності та метод найменування кластерів наукових публікацій для побудови предметних наукових просторів.

6. Вдосконалити методи кількісного оцінювання продуктивності та перспекції продуктивності суб'єктів наукової діяльності, що враховують самоцитування наукових публікацій.

7. Вдосконалити модель та метод вибору потенційних партнерів для співпраці в рамках наукових та освітніх проєктів, що враховує організаційну будову проєктів та результати кількісного оцінювання продуктивності суб'єктів наукової діяльності.

8. Вдосконалити векторний підхід до формування просторів суб'єктів наукової діяльності методології проєктно-векторного управління в освітніх середовищах.

Гіпотеза дослідження. Забезпечення сталого розвитку закладів вищої освіти залежить від розуміння принципів формування та взаємодії інформаційних просторів суб'єктів наукової діяльності, що є в складі даних закладів вищої освіти.

Методи дослідження. При побудові моделей та методів методології формування інформаційних просторів суб'єктів наукової діяльності використовувались теорії множин, систем та системного аналізу. Також використовувались методи математичного моделювання, кластерного аналізу, ймовірнісного латентно-семантичного аналізу, n-грам аналізу, багатокритеріальних методів прийняття рішень тощо.

Наукова новизна дисертаційного дослідження полягає у отриманні теоретичних та практичних результатів щодо розроблення методології формування інформаційних просторів суб'єктів наукової діяльності для забезпечення сталого розвитку закладів вищої освіти. До вагомих результатів наукового дослідження, що визначається науковою новизною та виносяться на захист, належать такі:

Вперше:

- розроблено множинні моделі ідентифікації науковців та закладів вищої освіти, які визначають ключові показники діяльності і включаються до складу інформаційних просторів, що дозволяє побудувати формалізовану модель інформаційного простору;

- розроблено формалізовану множинну модель інформаційного простору

суб'єктів наукової діяльності, що включає складові моделі ідентифікації науковців та закладу вищої освіти і дозволяє систематизувати розуміння інформаційного простору для створення методології формування інформаційних просторів суб'єктів наукової діяльності;

- розроблено методологію формування інформаційних просторів суб'єктів наукової діяльності, яка дозволяє формувати команди освітніх та наукових міжнародних проєктів зі складу провідних вчених світу, що включає методи ідентифікації предметних наукових просторів та оцінювання продуктивності суб'єктів наукової діяльності в рамках даних просторів, яка дозволяє підвищити обґрунтованість прийняття рішень з вибору виконавців наукових проєктів, забезпечити сталість розвитку закладів вищої освіти і, як наслідок, збільшити продуктивність суб'єктів наукової діяльності в цілому;

- розроблено інформаційну технологію, яка реалізує методологію формування інформаційних просторів суб'єктів наукової діяльності з використанням технології збору інформації з відкритих джерел, розрахунку та формування рейтингів суб'єктів наукової діяльності у відповідних предметних наукових просторах, створення інфографіки на основі зібраної інформації, що дозволяє дослідити складові інформаційних просторів суб'єктів наукової діяльності і використати їх для забезпечення сталого розвитку закладів вищої освіти, в складі яких є дані суб'єкти.

Набули подальшого розвитку:

- методологія проєктно-векторного управління в освітніх середовищах, в частині, що вдосконалює векторний підхід до формування просторів суб'єктів наукової діяльності, включаючи методи кількісного оцінювання наукової продуктивності, ідентифікації предметних наукових просторів, моделі та методи формування робочих груп наукових проєктів та складу структурних підрозділів закладів вищої освіти тощо;

- метод формування предметних наукових просторів за кластеризацією графу цитування з врахуванням шляху між вершинами на графі цитування, з використанням тематичної моделі та на основі встановлення подібності між публікаціями за методом локально-чутливого гешування з використанням n-грамів, що дозволяє виконати вимоги для застосування методів кількісного оцінювання продуктивності суб'єктів наукової діяльності і відрізняється від відомих підвищеною стійкістю кластеризації графу цитування наукових публікацій;

- понятійний апарат у галузі управління науково-дослідною діяльністю, зокрема введено поняття «модель ідентифікації науковця», «суб'єкти наукової діяльності», «предметний науковий простір», що дозволяє систематизувати інформацію про інформаційні простори.

Вдосконалено:

- метод кількісного оцінювання продуктивності суб'єктів наукової діяльності в предметному науковому просторі, що дозволяє врахувати самоцитування наукових публікацій, а також метод оцінювання перспекції продуктивності суб'єктів наукової діяльності, які відрізняються комплексністю охоплення ключових показників, що впливають на продуктивність наукової діяльності;

- метод найменування кластерів наукових публікацій для побудови предметних наукових просторів, що дозволяє встановити вербальний зв'язок між кластерами наукових публікацій та науковими напрямками, які поєднують суб'єктів наукової діяльності на основі нечіткого експертного опитування;
- математичну модель та метод багатокритеріального вибору потенційних партнерів для співпраці в рамках наукових та освітніх проєктів, що враховує результати кількісного оцінювання продуктивності суб'єктів наукової діяльності відповідного предметного наукового простору, яка відповідає робочому пакету проєкту і дозволяє сформулювати робочі групи та склад структурних підрозділів наукових установ та закладів вищої освіти, а також відрізняється від існуючих врахуванням організаційної будови наукових проєктів.

Практичне значення одержаних результатів полягає у тому, що розроблена методологія формування інформаційних просторів суб'єктів наукової діяльності є в основі вдосконалення науково-дослідної діяльності та забезпечення сталості функціонування наукових організацій та закладів вищої освіти. Вдосконалені методи кількісного оцінювання наукової продуктивності, аналізу її в динаміці, створення предметних наукових просторів є важливими для формування інформаційних просторів суб'єктів наукової діяльності. Теоретичні та практичні напрацювання використовуються суб'єктами наукової діяльності, тобто науковцями, науковими організаціями та їх структурними одиницями, а також профільними регіональними відомствами. Отримані результати як теоретичного, так і практичного характеру, є підґрунтям для подальших досліджень, спрямованих на удосконалення та забезпечення сталості розвитку закладів вищої освіти, що визначається нормативною базою в цій галузі, керується особливостями провадження наукової діяльності та відповідає вимогам міжнародних систем оцінювання наукової продуктивності відповідно до методології формування інформаційних просторів суб'єктів наукової діяльності.

Основні положення та результати дисертаційного дослідження впроваджено і застосовано у діяльності ВГО «Українська асоціація фахівців інформаційних технологій». Теоретичні та методичні розробки впроваджено у навчальний процес Київського національного університету будівництва і архітектури.

Особистий внесок здобувача. Основні результати роботи отримані автором особисто. З наукових праць, які опубліковані у співавторстві, у дисертаційному дослідженні описано ті положення, які є результатом особистої роботи автора: метод оцінювання результатів наукової-дослідної діяльності науковців та закладів вищої освіти [3; 5], метод комплексного оцінювання продуктивності суб'єктів наукової діяльності [7; 10; 11], методи кластеризації наукових публікацій для ідентифікації предметних наукових просторів [4; 9; 23], застосування латентно-семантичного аналізу для задач ідентифікації предметних наукових просторів [12]; концепція застосування методу для формування плану реалізації наукових проєктів, враховуючи розподіл ресурсів та час виконання проєктів [8], моделі і методи оцінювання тенденції зміни дискретних показників [1; 6; 13; 15], методи ідентифікації неповних дублікатів та встановлення подібності між текстовими даними, що необхідні для порівняння анотацій публікацій [2; 17; 18], методи побудови

інформаційно-аналітичних систем [14], концепція, завдання та особливості реалізації наукових і освітніх проєктів, а також підвищення їх ефективності [16]; аналіз національних та міжнародних рейтингів закладів вищої освіти та модель інформаційної технології оцінювання результатів наукової діяльності [19; 20], вимоги для розробки інформаційної технології вибору потенційних партнерів виконавців наукових і освітніх проєктів [21], огляд проблем, які виникають в межах аналізу наукових мереж [22], метод, який базується на основі ранжування посилянь та враховує самоцитування науковців для задачі оцінювання науково-дослідної діяльності суб'єктів наукового простору [24].

Апробація результатів дисертації. Основні результати роботи доповідалися, обговорювалися й одержали позитивну оцінку на: IEEE конференціях “The Experience of Designing and Applications of CAD Systems in Microelectronics” (CADSM-2017); “Problems of Infocommunications. Science and Technology” (PICST-2017, PICST-2018, PICST-2019); “Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering” (TCSET-2018); 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT-2018; IEEE 4th Intern. Symp. on Wireless Systems within the Int. Conf. On Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS); Advances in Intelligent Systems and Computing 2019, а також на міжнародних конференціях «Управління розвитком технологій» (2015, 2016, 2017, 2020), «Інформаційні технології та взаємодії» (2015, 2016, 2018), «Обчислювальний інтелект» (2015, 2017) та інших.

Публікації. За матеріалами дисертації опубліковано 68 наукових праць, з яких: 23 наукових статей у фахових виданнях України, 12 з яких опубліковані у виданнях, що індексуються у наукометричних базах даних Scopus та/або Web of Science, 1 публікація в зарубіжному виданні, 3 монографії, 41 матеріалів конференцій, з яких 10 опубліковані у виданнях, що індексуються у наукометричних базах даних Scopus та/або Web of Science.

Структура і обсяг роботи. Дисертація складається зі вступу, шести розділів, висновків по розділах, основних висновків, списку використаної літератури та додатків. Загальний обсяг дисертації становить 305 сторінок, з них основна частина складає 268 сторінок, у тому числі 51 рисунок, 8 таблиць, список літератури з 280 найменувань та 2 додатків.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертації, визначено мету, завдання і методи дослідження, визначено наукову новизну і практичне значення одержаних результатів, наведено загальну характеристику дисертаційної роботи.

У **першому розділі** розглянуто стан завдань формування інформаційних просторів суб'єктів наукової діяльності, що включає дослідження принципів створення та функціонування, а також оцінювання результатів співпраці в мережах співавторів або мережах наукової співпраці, побудова та дослідження мереж цитування, дослідження мереж спільного або перехресного цитування та бібліографічних з'єднань, ідентифікація предметних наукових просторів та аналіз тематик, методи прозорого кількісного оцінювання продуктивності суб'єктів наукової діяльності. Виявлено, що дослідження інформаційних просторів суб'єктів

наукової діяльності, зокрема їх перетинів та взаємодії між ними, є важливою передумовою сталого розвитку наукових та освітніх організацій та закладів вищої освіти як колективних суб'єктів наукової діяльності. Розвиток визначається функцією від приросту ідентифікаторів суб'єктів наукової діяльності, що складають шлейф їх інформаційних просторів. При чому розвиток індивідуальних суб'єктів наукової діяльності є каталізатором розвитку колективних суб'єктів наукової діяльності. Керованість ідентифікаторів суб'єктів наукової діяльності забезпечує систематично керований розвиток, що визначає сталість розвитку загалом.

З використанням силових алгоритмів візуалізації побудовані мережі наукової співпраці та мережі цитування реальної групи науковців, що виконували деякі науково-дослідні роботи з бюджетним фінансуванням в Київському національному університеті будівництва і архітектури з 2015 по 2020 роки (рис.1 та рис.2). Виявлено, що першочерговими факторами, які можуть здійснювати вплив на ефективність управління науковою діяльністю є забезпечення прозорого кількісного оцінювання продуктивності цієї діяльності, зокрема в рамках конкретних напрямів наукових досліджень, ідентифікація яких є окремою задачею.

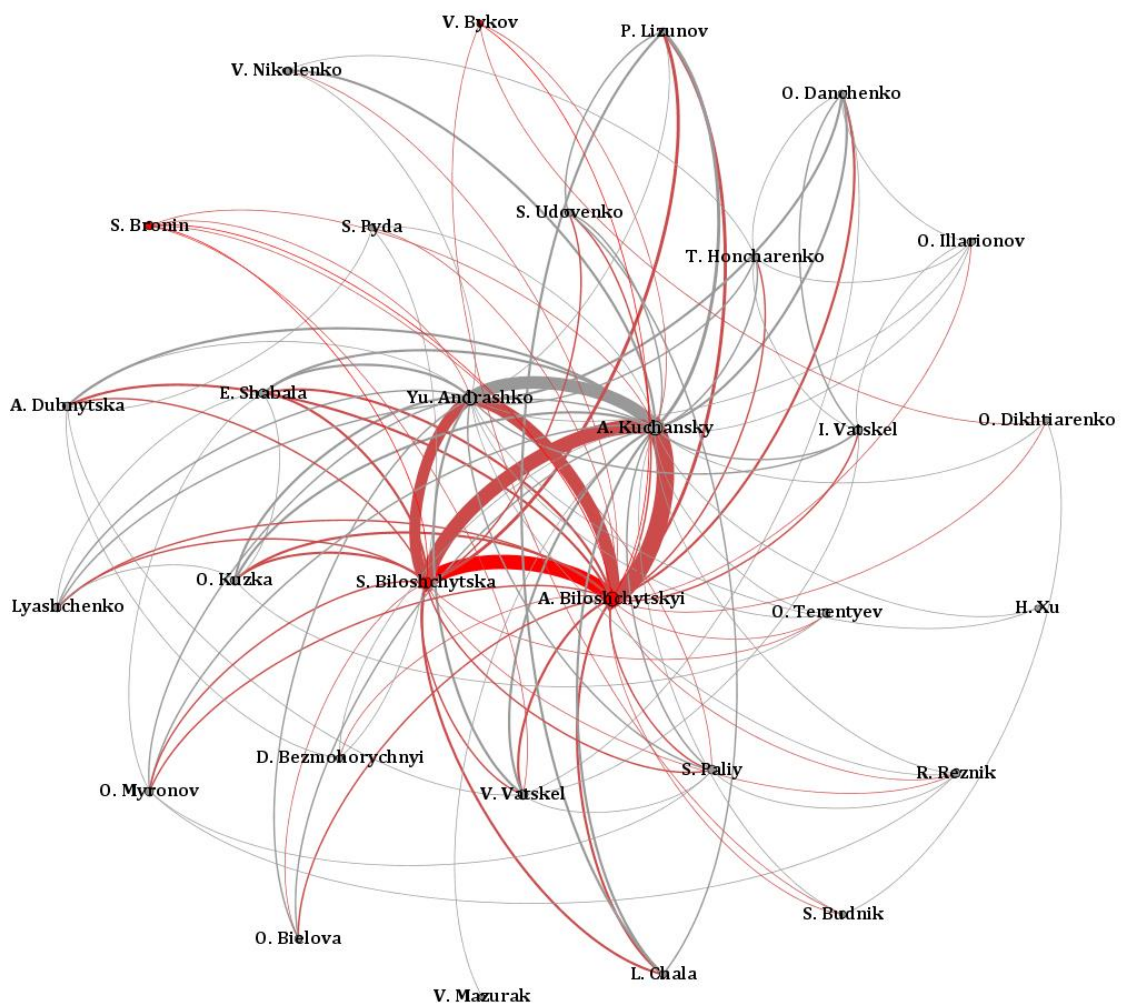


Рисунок 1. – Мережа наукової співпраці групи науковців Київського національного університету будівництва і архітектури

Виділено стратегічні орієнтири для вдосконалення науково-дослідної діяльності в Україні з врахуванням світових підходів. Виявлено, що проведення прозорого кількісного оцінювання продуктивності науковців та їх організаційних об'єднань є запорукою якісного підбору виконавців та учасників наукових та освітніх проєктів, створення консорціумів та подання заявок на отримання міжнародних грантів, що важливо для підвищення репутації України в світовому науковому просторі, стимулює науково-технічний розвиток та створення інновацій.

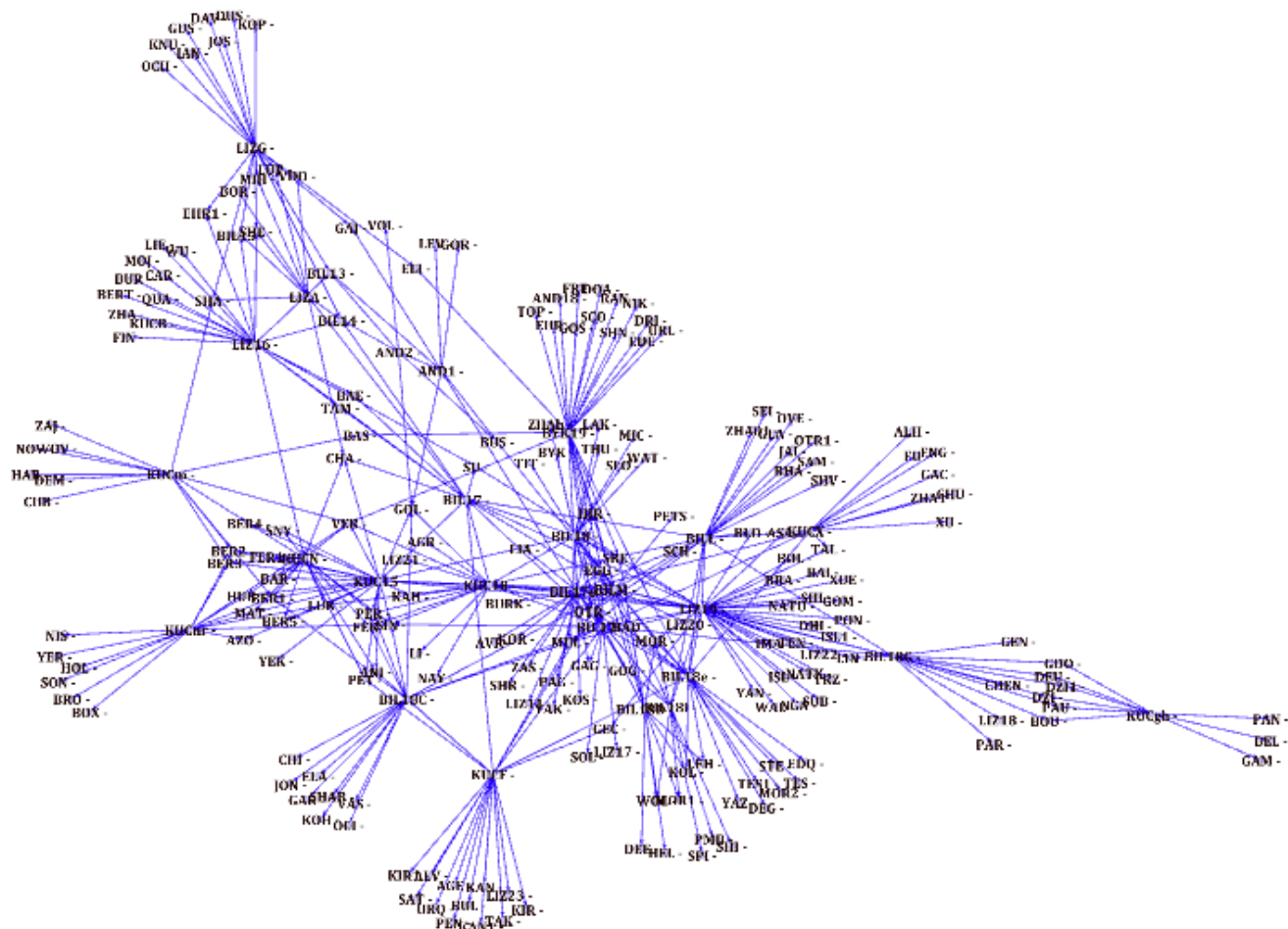


Рисунок 2. – Мережа цитування наукових публікацій групи науковців Київського національного університету будівництва і архітектури, які опубліковані в межах наукових проєктів

Глобальною метою розроблення методології формування інформаційних просторів суб'єктів наукової діяльності є домогтися прогресу у сталості розвитку наукових організацій та закладів вищої освіти, вільному обігу знань, дослідників та технологій, здійснення реформ у науково-дослідній та інноваційній діяльності, а також забезпечити процеси цифрової трансформації в Україні та інтеграції у Європейський науково-дослідний простір.

Другим розділом закладається поняттєво-термінологічний апарат та концепція методології формування інформаційних просторів суб'єктів наукової діяльності.

Означення 1. Суб'єктами наукової діяльності вважатимемо заклади вищої

освіти та їх структурні підрозділи, що безпосередньо займаються освітньо-науковою діяльністю: інститути, факультети, кафедри, науковий та професорсько-викладацький склад цих підрозділів, групи науковців, об'єднані виконанням деяких проєктів та окремі науковці. Одним з обов'язкових компонентів діяльності таких суб'єктів є наукова діяльність. В контексті цього виникає задача кількісного оцінювання продуктивності цієї діяльності. Індивідуальні суб'єкти наукової діяльності – це науковці. Колективні суб'єкти наукової діяльності – це наукові організації та заклади вищої освіти, які формуються з індивідуальних суб'єктів наукової діяльності. Кожен суб'єкт наукової діяльності представляється моделлю ідентифікації та інформаційним простором.

Означення 2. Інформаційний простір суб'єктів наукової діяльності – це інформаційний простір, складниками якого є інформаційні об'єкти, які організовані у певну структуру з ідентифікаторами, що визначаються за результатами освітньої та наукової діяльності певних суб'єктів цієї діяльності. Тобто інформаційний простір суб'єктів наукової діяльності включає суб'єкти наукової діяльності та ретроспективний шлейф з ідентифікаторів цих суб'єктів, що визначаються за результатами їх наукової діяльності. Інформаційні простори різних суб'єктів наукової діяльності можуть перетинатися та сегментуються при включенні до предметних наукових просторів.

Означення 3. Продуктивність суб'єктів наукової діяльності – це кількісний показник результативності суб'єктів наукової діяльності. В процесі розрахунку цього показника можуть використовуватись як розроблені, так і відомі індекси наукової активності, залежно від потреб розрахунку та його застосування. Продуктивність суб'єктів наукової діяльності часто розглядається як відношення результатів праці відносно часу фіксації наукових показників або в порівнянні з іншими суб'єктами наукової діяльності. Тому для розрахунку цього показника потрібно чітко дослідити інформаційний простір суб'єктів наукової діяльності, його зв'язки з іншими інформаційними просторами. Це можна зробити, виділивши частину інформаційного простору, яка відповідає за певний предметний напрямок досліджень.

Означення 4. Предметний науковий простір – це сукупність суб'єктів наукової діяльності та частин їх інформаційних просторів, що упорядковані та об'єднані за критерієм спільного спрямування тематики науково-дослідної діяльності. В основі цього простору є суб'єкт наукової діяльності та його частини інформаційних просторів з відповідними ідентифікаторами.

Нехай $A = \{a_1, a_2, \dots, a_n\}$ – множина суб'єктів наукової діяльності, які займаються активною публікаційною діяльністю, n – кількість суб'єктів наукової діяльності. Нехай $Q = \{q_1, q_2, \dots, q_m\}$ – множина наукових публікацій, які опубліковані суб'єктами наукової діяльності з множини A , m – кількість наукових публікацій.

Нехай задано множину всіх пар, між елементами множин A та Q

$$A \times Q = \{(a, q) \mid a \in A \wedge q \in Q\} \quad (1)$$

та бінарне відношення $U \subset A \times Q$, що визначає авторство публікацій $q \in Q$. Також задане бінарне відношення $S \subset Q \times Q$, що визначає цитування всіх публікацій $q \in Q$,

$$Q \times Q = \{(q_i, q_j) \mid q_i, q_j \in Q, i \neq j\}. \quad (2)$$

Множину всіх публікацій суб'єкта наукової діяльності $a \in A$ в момент часу t позначимо через

$$Q(a_i, t) = \{q_j \in Q \mid (a_i, q_j) \in U, t \in T\}, \quad i = \overline{1, n}, \quad j = \overline{1, m}. \quad (3)$$

Множина наукових публікацій, які цитує кожен суб'єкт наукової діяльності $a \in A$ в момент часу t позначимо через $\overline{C}(a_i, t)$, та множину публікацій, у яких цитуються публікації суб'єкта наукової діяльності в момент часу t – через $C(a_i)$, таким чином:

$$\overline{C}(a_i, t) = \{q_j \in Q \mid (q_y, q_j) \in C, q_y \in Q(a_i), t \in T, y = \overline{1, m}\}, \quad i = \overline{1, n}, \quad j = \overline{1, m}, \quad (4)$$

$$C(a_i, t) = \{q_j \in Q \mid (q_j, q_y) \in C, q_y \in Q(a_i), t \in T, y = \overline{1, m}\}, \quad i = \overline{1, n}, \quad j = \overline{1, m}. \quad (5)$$

Для кожної публікації $q \in Q$ визначимо множину її авторів

$$A(q_j) = \{a_i \in A \mid (a_i, q_j) \in U\}, \quad i = \overline{1, n}, \quad j = \overline{1, m}, \quad (6)$$

а також множину наукових публікацій, які цитує задана публікація $q \in Q$ в момент часу t – $\overline{C}(q_j)$ та множину публікацій, у яких цитується публікація $q \in Q$ в момент часу t – $C(q)$:

$$\overline{C}(q_j, t) = \{q_j \in Q \mid (q_j, q_y) \in C, t \in T, y = \overline{1, m}\}, \quad j = \overline{1, m}, \quad (7)$$

$$C(q_j, t) = \{q_j \in Q \mid (q_y, q_j) \in C, t \in T, y = \overline{1, m}\}, \quad j = \overline{1, m}. \quad (8)$$

Кожній публікації $q \in Q$ поставимо у відповідність її анотацію $q^A \in Q^A$.

Скалярна оцінка продуктивності суб'єкта наукової діяльності – це деяке функціональне відображення Φ :

$$\Phi : A \rightarrow R, \quad (9)$$

де R – множина дійсних чисел.

Індивідуальний суб'єкт наукової діяльності в момент часу t представляється сукупністю ідентифікаторів, які його визначають:

$$\aleph^I(a_i, t) = \langle Q(a_i, t), \overline{C}(a_i, t), C(a_i, t), Q^A(a_i, t), \Phi(a_i, t), M(a_i, t) \rangle, \quad (10)$$

де $Q(a_i, t)$ – множина публікацій індивідуального суб'єкта наукової діяльності a_i в момент часу t ;

$\overline{C}(a_i, t)$ – множина наукових публікацій, які цитує суб'єкт наукової діяльності a_i в момент часу t ;

$C(a_i, t)$ – множина публікацій, у яких цитуються публікації суб'єкта наукової діяльності a_i в момент часу t ;

$Q^A(a_i, t)$ – множина анотацій публікацій суб'єкта наукової діяльності a_i в момент часу t ;

$\Phi(a_i, t)$ – оцінка продуктивності суб'єкта наукової діяльності a_i в момент часу t , наприклад індекс Гірша;

$M(a_i, t)$ – кількість міжнародних проєктів, в яких бере участь суб'єкт наукової діяльності a_i в момент часу t .

Також до моделі ідентифікації індивідуального суб'єкта наукової діяльності можна включити імпакт-фактор його публікацій та інші параметри.

Коллективний суб'єкт наукової діяльності представляється сукупністю ідентифікаторів, які його визначають:

$$\aleph^C(a_i, t) = \langle \Phi_1(a_i, t), \Phi_2(a_i, t), \Phi_3(a_i, t), \Phi_4(a_i, t), \Phi_5(a_i, t) \rangle, \quad (11)$$

де $\Phi_1(a_i, t)$ – нормована оцінка міжнародної активності суб'єкта наукової діяльності a_i в момент часу t , наприклад, кількість стажувань співробітників колективного суб'єкта наукової діяльності за кордоном, кількість проєктів з іноземним фінансуванням тощо;

$\Phi_2(a_i, t)$ – нормована оцінка контингенту здобувачів вищої освіти колективного суб'єкта наукової діяльності, a_i в момент часу t , в разі, якщо він надає освітні послуги, наприклад є закладом вищої освіти (ЗВО). Оскільки оцінювання ЗВО має враховувати освітню складову, контингент студентів: середній бал ЗНО зарахованих на навчання, прохідний мінімум тощо, що є важливими оцінками, що визначають потенціал ЗВО;

$\Phi_3(a_i, t)$ – нормована оцінка наукового або науково-педагогічного складу колективного суб'єкта наукової діяльності a_i в момент часу t , може бути усередненою кількісною оцінкою продуктивності співробітників за певний період часу;

$\Phi_4(a_i, t)$ – нормована кількісна оцінка науково-дослідної діяльності колективного суб'єкта наукової діяльності в момент часу t ;

$\Phi_5(a_i, t)$ – нормована оцінка наявного ресурсного забезпечення в момент часу t , зокрема матеріально-технічного забезпечення колективного суб'єкта наукової діяльності.

Ідентифікатори суб'єкта наукової діяльності в поточний момент часу t_x , включаючи їх ретроспективні значення в моменти часу $t_{x-\Delta}$, $\Delta = \overline{1, x}$, t_0 – початковий момент часу, формують інформаційний простір суб'єкта наукової діяльності, причому значення ідентифікаторів є накопичувальними. Шлейф значень ідентифікаторів індивідуальних суб'єктів наукової діяльності a_i представляється багаторівневим часовим рядом:

$$(\aleph^I(a_i, t_0), \aleph^I(a_i, t_1), \dots, \aleph^I(a_i, t_{x-1}), \aleph^I(a_i, t_x)). \quad (11)$$

Шлейф значень ідентифікаторів колективних суб'єктів наукової діяльності a_i представляється багаторівневим часовим рядом:

$$(\aleph^C(a_i, t_0), \aleph^C(a_i, t_1), \dots, \aleph^C(a_i, t_{x-1}), \aleph^C(a_i, t_x)). \quad (12)$$

Потенціал розвитку суб'єктів наукової діяльності визначається швидкістю зміни показників ідентифікаторів цих суб'єктів, що розраховується як відсоткове порівняння поточного значення кожного ідентифікатора з значенням одного з попередніх періодів. Для колективних суб'єктів наукової діяльності:

$$S^C(a_i, t_x, t_b) = \frac{1}{5} \sum_{j=1}^5 \frac{\Phi_j(a_i, t_x) - \Phi_j(a_i, t_{x-b})}{\Phi_j(a_i, t_{x-b})} \cdot 100\%, \quad (13)$$

де $S^C(a_i, t_x, t_b)$ – швидкість зміни показників ідентифікаторів колективних суб'єктів наукової діяльності a_i в момент часу t_x відносно моменту часу t_b , тобто

потенціал колективних суб'єктів наукової діяльності a_i .

Аналогічно для індивідуальних суб'єктів наукової діяльності швидкості зміни показників визначаються за формулами:

$$Q^N(a_i, t_x) = \frac{|Q(a_i, t_x)| - \min_{j=1, n} (|Q(a_j, t_x)|)}{\max_{j=1, n} (|Q(a_j, t_x)|) - \min_{j=1, n} (|Q(a_j, t_x)|)}, \quad (14)$$

$$C^N(a_i, t_x) = \frac{|C(a_i, t_x)| - \min_{j=1, n} (|C(a_j, t_x)|)}{\max_{j=1, n} (|C(a_j, t_x)|) - \min_{j=1, n} (|C(a_j, t_x)|)}, \quad (15)$$

$$M^N(a_i, t_x) = \frac{|M(a_i, t_x)| - \min_{j=1, n} (|M(a_j, t_x)|)}{\max_{j=1, n} (|M(a_j, t_x)|) - \min_{j=1, n} (|M(a_j, t_x)|)}, \quad (16)$$

$$\Phi^N(a_i, t_x) = \frac{\Phi(a_i, t_x) - \min_{j=1, n} (\Phi(a_j, t_x))}{\max_{j=1, n} (\Phi(a_j, t_x)) - \min_{j=1, n} (\Phi(a_j, t_x))}, \quad (17)$$

$$S^I(a_i, t_x, t_b) = \frac{1}{4} \cdot \left(\frac{Q^N(a_i, t_x) - Q^N(a_i, t_{x-b})}{Q^N(a_i, t_{x-b})} + \frac{C^N(a_i, t_x) - C^N(a_i, t_{x-b})}{C^N(a_i, t_{x-b})} + \frac{M^N(a_i, t_x) - M^N(a_i, t_{x-b})}{M^N(a_i, t_{x-b})} + \frac{\Phi^N(a_i, t_x) - \Phi^N(a_i, t_{x-b})}{\Phi^N(a_i, t_{x-b})} \right) \cdot 100\%, \quad (18)$$

де $S^I(a_i, t_x, t_b)$ – швидкість зміни нормованих показників ідентифікаторів індивідуальних суб'єктів наукової діяльності a_i в момент часу t_x відносно моменту часу t_b , тобто потенціал індивідуальних суб'єктів наукової діяльності a_i .

Нехай $A^I \subset A$ – множина індивідуальних суб'єктів наукової діяльності, $A^C \subset A$ – множина колективних суб'єктів наукової діяльності. Нехай $\lambda \in \Omega \subset A^C$, де Ω – множина, якій належать індивідуальні суб'єкти наукової діяльності і яка визначає колективний суб'єкт наукової діяльності a_Ω , то

$$\exists \lambda \in \Omega, \forall t_y \in T, S^I(\lambda, t_y, t_b) > 0 \Rightarrow S^C(a_\Omega, t_y, t_b) \geq 0, \quad (19)$$

тобто якщо існують такі індивідуальні суб'єкти наукової діяльності, потенціал яких додатній, то з цього випливає, що потенціал колективного суб'єкта наукової діяльності або організації, що включає дані індивідуальні суб'єкти, буде мати невід'ємний потенціал, тобто швидкість зміни показників ідентифікаторів колективного суб'єкта наукової діяльності буде додатня.

Отже, розвиток індивідуальних суб'єктів наукової діяльності спричиняє розвиток колективних суб'єктів наукової діяльності або організацій. Причому частини інформаційних просторів суб'єктів наукової діяльності перетинаються з інформаційними просторами інших суб'єктів наукової діяльності, наприклад спільні публікації, в яких обидва суб'єкти є авторами.

Суб'єкти наукової діяльності та частини їх інформаційних просторів включаються до предметних наукових просторів, що об'єднують суб'єктів, які працюють в спільному науковому напрямі. У випадку створення спільних проєктів,

інформаційні простори індивідуальних суб'єктів наукової діяльності взаємодіють, утворюючи нову якість, що є каталізатором розвитку колективних суб'єктів наукової діяльності, підвищуючи власний потенціал розвитку (рис. 3).

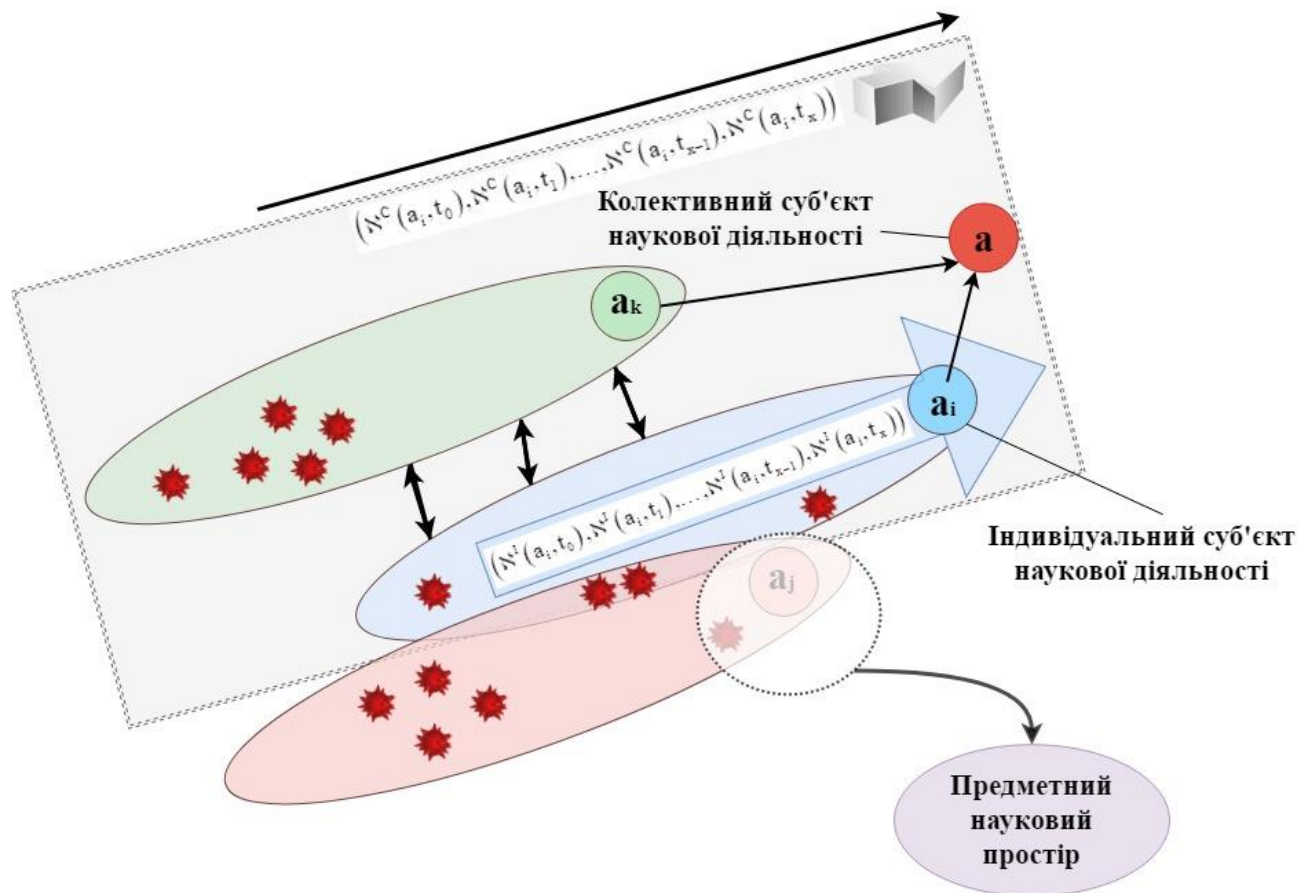


Рисунок 3. – Концептуальна схема взаємодії суб'єктів наукової діяльності

Систематично керований розвиток є запорукою сталості розвитку колективних суб'єктів наукової діяльності: наукових організацій та закладів вищої освіти. Для того, щоб це показати, необхідно провести дослідження, схема якого вказана на рисунку 4:

1. Побудова та дослідження множинної моделі ідентифікації індивідуальних та колективних суб'єктів наукової діяльності.
2. Побудова та застосування методів ідентифікації предметних наукових просторів.
3. Побудова за застосування методів кількісного оцінювання рівня та характеру розвитку продуктивності суб'єктів наукової діяльності.
4. Побудова та застосування математичної моделі та методу багатокритеріального вибору суб'єктів наукової діяльності для організації спільних наукових та освітніх проєктів.
5. Забезпечення сталості розвитку колективних суб'єктів наукової діяльності: наукових організацій, закладів вищої освіти та їх структурних підрозділів.

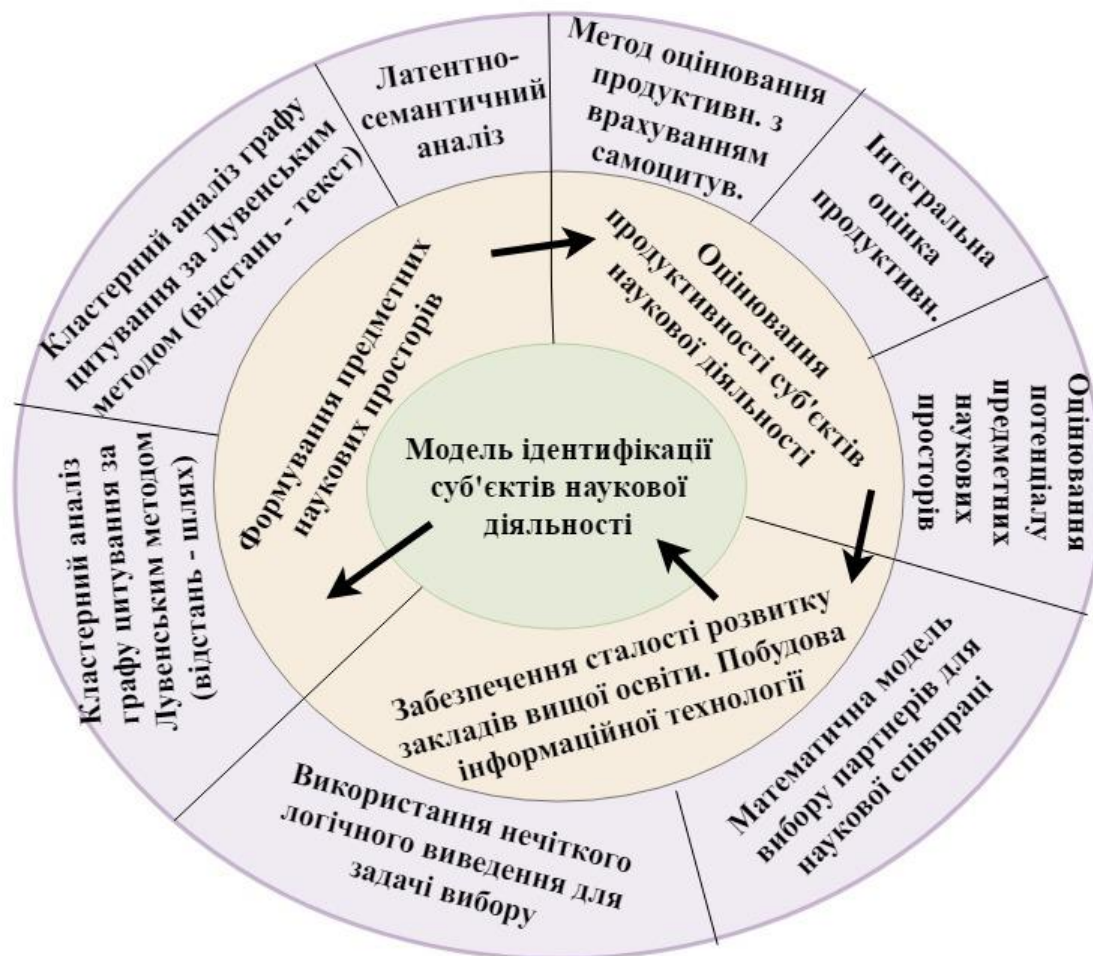


Рисунок 4. – Структура схеми дослідження з формування інформаційних просторів суб'єктів наукової діяльності

Третій розділ присвячено розробленню методів формування предметних наукових просторів на основі кластерного аналізу параметрів моделей ідентифікації суб'єктів наукової діяльності, що входять до інформаційних просторів цих суб'єктів.

Ідентифікація предметних наукових просторів необхідна для:

1. Визначення того, в якому напрямі досліджень працює кожен індивідуальний суб'єкт наукової діяльності та який його вклад в ідентифікований напрям.
2. Дослідження ідентифікаторів суб'єктів такого наукового напрямку.
3. Оцінювання потенціалу наукового напрямку, враховуючи динаміку розвитку суб'єктів, які складають науковий простір, що відображається цим напрямом.

Схема методу ідентифікації предметних наукових просторів вказана на рис.5.

Нехай w_s – деякий предметний науковий простір, $w_s \subset W$, де W – загальний освітньо-науковий простір.

Задача полягає у побудові таких предметних просторів w_s , які б склалися з суб'єктів наукової діяльності, публікаційна активність яких відбувається у спільному науковому напрямі. Для цього пропонується виконати ряд етапів:

1. Задати метричний простір та розглянути методи встановлення відстані між публікаціями індивідуальних суб'єктів наукової діяльності.

2. Провести кластерний аналіз публікацій індивідуальних суб'єктів наукової діяльності за заданими критеріями та побудувати кластери подібних наукових публікацій.

3. Проаналізувати результати кластеризації. За потреби об'єднати близькі кластери.

4. Провести найменування кластерів, тобто встановити відповідність між кластерами публікацій та відповідною його назвою. Побудувати на основі ідентифікованих кластерів, предметних наукових просторів, що складатимуться з суб'єктів наукової діяльності.



Рисунок 5. – Схема методу ідентифікації предметних наукових просторів

Розглянемо метод визначення відстані між вершинами графу цитування на основі знаходження довжини мінімального шляху між ними. Нехай задано метричний простір (Q, g) , який складається з множини наукових публікацій Q та деякої відстані g , визначеної для будь-якої пари елементів цієї множини. Відстань g визначається як відображення з множини декартового квадрату множини Q на множину дійсних чисел, тобто:

$$g: Q \times Q \rightarrow \mathbf{R}, \quad (20)$$

де \mathbf{R} – множина дійсних чисел.

Відстань між публікаціями g є невід'ємною дійсною функцією $g(q_i, q_j) \geq 0$, $i = \overline{1, m}$, $j = \overline{1, m}$, яка визначається для $\forall q_i, q_j \in Q$. Для функції $g(q_i, q_j)$, $i \neq j$ виконуються аксіоми тотожності, симетрії та нерівність трикутника.

Відстань $g(q_i, q_j)$ між довільними вершинами $q_i, q_j \in Q$ визначатиметься як довжина мінімального шляху з вершини q_i у вершину q_j , якщо він існує.

Найкоротший шлях може бути знайдений за алгоритмом Беллмана-Форда зі складністю $O(QC)$ або Дейкстри зі складністю $O(Q^2)$.

Розглянемо метод визначення відстані між вершинами графу цитування на основі встановлення ступеня близькості за змістом анотацій публікацій. Кожній публікації $q \in Q$ поставимо у відповідність її анотацію $q^A \in Q^A$. Побудуємо повний зважений граф (Q^A, C) , Q^A – множина вузлів або вершин графа, C – множина ребер графа, тобто пар (q_i^A, q_j^A) , $q_i^A, q_j^A \in Q^A$. Відстань g між публікаціями (q_i, q_j) або вага ребра (q_i^A, q_j^A) буде розглядатися як відстань між їх анотаціями. Завдання полягає у визначенні цих відстаней між всіма вершинами графа.

Розглянемо метод визначення відстані між вершинами графу цитування на основі m -грам аналізу. Текст анотацій $q^A, q_y^A \in Q^A$ в цьому випадку представляють собою послідовності m -грамів. Задача полягає в знаходженні відстані між анотаціями шляхом порівняння цих m -грамів. Побудуємо для рядкової послідовності $S = \{s_1, s_2, \dots, s_w\}$ вхідної анотації $q^A \in Q^A$ m -грам зі слів в канонізованій формі:

$$W = \{S_1^{\beta_1}, S_2^{\beta_2}, \dots, S_m^{\beta_m}\}, \quad (21)$$

де β_j , $j = \overline{1, m}$ – довжини слів, а m – їх кількість.

Нехай S_a та S_b деякі слова, тоді частотою m -грама $T(S_a, S_{a+1}, \dots, S_{a+m-2}, S_b)$ назвемо відношення кількості входжень m -грама в даний текст до загальної кількості m -грамів в тексті. Тоді для кожної пари слів S_a та S_b розрахуємо середню частоту m -грам, що починаються та закінчуються відповідними словами за формулою:

$$\mu(S_a, S_b) = \frac{1}{2} \sum_{i_1=1}^u \dots \sum_{i_{m-2}=1}^u \left(T(S_a, S_{i_1}, \dots, S_{i_{m-2}}, S_b) + T(S_b, S_{i_1}, \dots, S_{i_{m-2}}, S_a) \right), \quad (22)$$

$\underbrace{\hspace{10em}}_{m-2 \text{ рази}}$

де S_{i_j} , $j = \overline{1, m-2}$, $i_j = \overline{1, u}$, – слова в канонізованій формі, які зустрічаються в анотації після канонізації, u – кількість слів.

Для подібних анотацій, різниця середніх частот m -грам не перевищує наперед заданого порогового значення.

Пропонуються застосовувати Лувенський метод для кластеризації графу цитування. Причому, оскільки кількість класерів після такої кластеризації буде достатньо великою, після кластеризації кластери потрібно об'єднати за критерієм близькості центрів ваги цих кластерів.

Позначимо множину кластерів через $K = \{k_1, k_2, \dots, k_z\}$, де z – кількість кластерів, на які розбивається Q . Для кластерів виконуються умови:

1. Кожна публікація обов'язково належить одному із кластерів, тобто $\bigcup_{i=1}^z k_i = Q$
2. Кожна публікація належить єдиному кластеру, тобто $k_i \cap k_j = \emptyset$, $\forall i \neq j$.
3. В кожному кластері об'єднані достатньо близькі (в розумінні відстані g) публікації.

Центром ваги кластеру $k_r = \{q_1^r, q_2^r, \dots, q_{\mu_r}^r\}$, $r = \overline{1, Z}$, називається такий з об'єктів кластеру Ω^r , що сумарна відстань до інших об'єктів даного кластеру є мінімальною:

$$\Omega^r = \arg \min \left(\sum_{i=1}^{\mu_r} g(q_i^r, q_j^r), j = \overline{1, \mu_r} \right), \quad (23)$$

де $\mu_r = \text{card}(k_r)$ – кількість об'єктів, які належать до кластеру k_r , $r = \overline{1, Z}$.

Розглянемо тематичну модель для кластеризації графу цитування і побудови предметних наукових просторів. Предметні наукові простори є динамічними структурами. Це пов'язано з тим, що перелік термінів, які характеризують простір постійно коригується внаслідок проведення нових досліджень та втрати актуальності інших досліджень. Відповідно склад предметних наукових просторів динамічно змінюється. Тому для виявлення меж цих просторів варто застосовувати інструменти, які можуть виділяти характеристики просторів без попереднього їх задання. Для цієї задачі можна скористатися методами ймовірнісного латентно-семантичного аналізу.

Нехай існує скінченна множина тем досліджень T . Якщо частота появи певних термів, які визначають предметний науковий простір V , в тексті вищій, в порівнянні з частотою появи термів інших напрямків, то текст належить до предметного наукового простору V . Темою документу будемо розуміти ймовірнісний дискретний розподіл на множині термів Ω .

Тобто існує прихована залежність між термами, темами та текстовим документом. Для відображення цієї залежності представимо текстові документи як множину точок (q_i, ω_i, t_i) , $i = \overline{1, Y}$, $Y = |Q| \cdot |\Omega| \cdot |T|$ у дискретному ймовірнісному просторі $Q \times \Omega \times T$ з невідомою функцією ймовірності $p(q, \omega, t)$. Значення функції $p(q, \omega, t)$ можуть бути оцінені на основі статистичного означення ймовірності:

$$p(q, \omega, t) = \frac{n_{q\omega t}}{\sum_{q=1}^{|Q|} \sum_{\omega=1}^{|\Omega|} \sum_{t=1}^{|T|} n_{q\omega t}}, \quad (24)$$

де $p(q, \omega, t)$ – ймовірність використання терму ω в текстовому документі q за темою t , $n_{q\omega t}$ – кількість точок (q, ω, t) в просторі $Q \times \Omega \times T$ для заданої множини текстових документів. Іншими словами $p(q, \omega, t)$ – це кількість використань терму ω в текстовому документі q за темою t . Будемо вважати, що застосування термів залежить тільки від теми. Позначимо через

$$p(\omega|q) = \sum_{t=1}^{|T|} p(\omega|t)p(t|q) = \sum_{t=1}^{|T|} \phi_{\omega t} \theta_{tq} \quad (25)$$

– ймовірнісну модель, що описує процес формування колекції документів на основі відомих розподілів $p(\omega|t)$ та $p(t|q)$. На основі колекції документів Q необхідно знайти частотні оцінки розподілів або параметри $\phi_{\omega t}$ та θ_{tq} . Параметри визначені таким чином, щоб модель (1) наближала оцінки умовних ймовірностей

$$p(\omega|q) = \frac{n_{q\omega}}{n_q}, \quad (26)$$

де $n_{\omega q}$ – це число входжень терма ω в документ q , n_q – довжина документа q . Нехай Φ – матриця, що представляє належність термів до тем $\Phi = (\phi_{\omega t})_{\Omega \times T}$, а Θ –

матриця належності тем до документів $\Theta = (\theta_{tq})_{T \times Q}$.

Для оцінювання невідомих параметрів ймовірнісних моделей можна використати принцип максимуму правдоподібності з регуляризатором $R(\Phi, \Theta)$.

$$\sum_{q=1}^{|Q|} \sum_{\omega \in q} n_{q\omega} \ln \sum_{t=1}^{|T|} \phi_{\omega t} \theta_{tq} + R(\Phi, \Theta) \rightarrow \max_{\Phi, \Theta} \quad (27)$$

$$\sum_{\omega=1}^{|\Omega|} \phi_{\omega t} = 1, \sum_{t=1}^{|T|} \theta_{tq} = 1, \phi_{\omega t} \geq 0, \theta_{tq} \geq 0. \quad (28)$$

Нехай множина $C \subset Q \times Q$ задає відношення цитування між публікаціями суб'єктів наукової діяльності. Взаємозв'язок між публікаціями та їх цитуваннями можна відобразити у вигляді орієнтованого графа (Q, C) , де публікації з множини Q є вершинами, а цитування C – дугами графа. Для знаходження ваг $n_{q\omega}$ можна використати величину, обернену до довжини мінімального маршруту між відповідними вершинами графу (Q, C) . Якщо ж між вершинами не існує маршруту, то $n_{q\omega} = 0$:

$$\sum_{q=1}^{|Q|} \sum_{\omega \in q} n_{q\omega} \ln \sum_{t=1}^{|T|} \phi_{\omega t} \theta_{tq} + \tau \sum_{q=1}^{|Q|} \sum_{c=1}^{|Q|} n_{qc} \sum_{t=1}^{|T|} \theta_{tq} \theta_{tc} \rightarrow \max_{\Phi, \Theta}, \quad (29)$$

$$\sum_{\omega=1}^{|\Omega|} \phi_{\omega t} = 1, \sum_{t=1}^{|T|} \theta_{tq} = 1, \phi_{\omega t} \geq 0, \theta_{tq} \geq 0. \quad (30)$$

В результаті отримаємо нечіткий розподіл належностей публікацій до кластерів, що описується матрицею Θ :

$$\Theta = \begin{pmatrix} \theta_{11} & \theta_{12} & \dots & \theta_{1m} \\ \theta_{21} & \theta_{22} & \dots & \theta_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{\psi 1} & \theta_{\psi 1} & \dots & \theta_{\psi m} \end{pmatrix}. \quad (31)$$

Позначимо через $\overline{Q}_r(a_i)$ множину публікацій автора a_i , $i = \overline{1, n}$, що включена до кластеру k_r , $r = \overline{1, z}$:

$$\overline{Q}_r(a_i) = \{q_j \in Q \mid (a_i, q_j) \in U, q_j \in k_r\}, i = \overline{1, n}, r = \overline{1, z}, \quad (32)$$

$U \subset A \times Q$ відображає авторство публікацій $q \in Q$.

Тоді належність публікацій індивідуального суб'єкта наукової діяльності a_i , $i = \overline{1, n}$ в кожному з кластерів k_r , $r = \overline{1, z}$ визначатимуться як:

$$R_r(a_i) = \frac{\sum_{q \in \overline{Q}_r(a_i)} \theta_{zq}}{\text{card}(\overline{Q}_r(a_i))}, r = \overline{1, z}, \quad (33)$$

де $R_r(a_i)$ – належність публікацій індивідуального суб'єкта наукової діяльності a_i до кластеру k_r .

Розглянемо метод найменування кластерів наукових публікацій. Після кластеризації графу (Q, C) потрібно встановити відповідність між кластером та

назвою предметного наукового простору, який цей кластер представляє. Особливістю розв'язку задачі є те, що предметні наукові простори перетинаються внаслідок того, що деякі індивідуальні суб'єкти наукової діяльності потрапляють одразу до кількох просторів, тобто їх публікації можуть належати до кількох наукових напрямків. В цьому випадку визначити чітко належність індивідуального суб'єкта наукової діяльності до відповідного предметного наукового простору неможливо і ця належність буде нечіткою.

Якщо всі публікації належать своїм кластерам, то

$$m = \sum_{i=1}^n \sum_{r=1}^z R_r(a_i) \quad (34)$$

і можна задати множину впорядкованих пар, що складаються з елементів $a \in A$ і відповідної функції належності $\mu_r(a)$, $r = \overline{1, z}$, тобто множини $E^r = \{(a, \mu_r(a)) \mid a \in A\}$, де функція належності $\mu_r(a)$ вказує на те, в якій мірі елемент $a \in A$, тобто індивідуальний суб'єкт наукової діяльності, належить нечіткій множині E^r , що відповідає визначеному напрямку наукових досліджень або предметному науковому простору. Функція $\mu_r(a)$ приймає значення на деякій множині належностей $[0, 1]$.

Носієм нечіткої множини E^z , $\text{supp} E^z$ називається така множина:

$$\text{supp} E^r = \{a \mid a \in A, \mu_r(a) > 0\}. \quad (35)$$

Визначимо функцію $\mu_r(a)$ так:

$$\mu_r(a_i) = \begin{cases} 0, & \text{supp} E^r \cap a_i = \emptyset \\ \frac{R_r(a_i)}{R(a_i)}, & \text{supp} E^r \cap a_i \neq \emptyset \end{cases}, \quad (36)$$

де $R(a_i) = \sum_{r=1}^z R_r(a_i)$, $i = \overline{1, n}$.

Значення $\mu_r(a_i) = 1$, коли $\bigcup_{\substack{r=1, z, \\ r \neq i}} \text{supp} E^r \cap a_i = \emptyset$, причому $\sum_{r=1}^z \mu_r(a_i) = 1$ для

фіксованого $i = \overline{1, n}$.

Множина предметних наукових просторів індивідуального суб'єкта наукової діяльності a_i визначається так:

$$\Theta(a_i) = \{\Phi(k_r) = \eta, \eta \in W_s \mid k_r \cap Q(a_i) \neq \emptyset\}, \quad (37)$$

де $\Theta(a_i)$ – множина предметних наукових просторів індивідуального суб'єкта наукової діяльності a_i , $i = \overline{1, n}$, $\Phi: K \rightarrow W_s$

Сукупність множин $\Theta(a_i)$ для всіх індивідуальних суб'єктів наукової діяльності є розв'язком задачі ідентифікації предметних наукових просторів. Отримані результати дають змогу розв'язати споріднену задачу знаходження індивідуальних суб'єктів наукової діяльності, що здійснюють дослідження у заданому напрямі. Схематична належність індивідуальних суб'єктів наукової діяльності до кількох

предметних наукових просторів вказана на рисунку 6. Результати найменування кластерів – рис. 7.

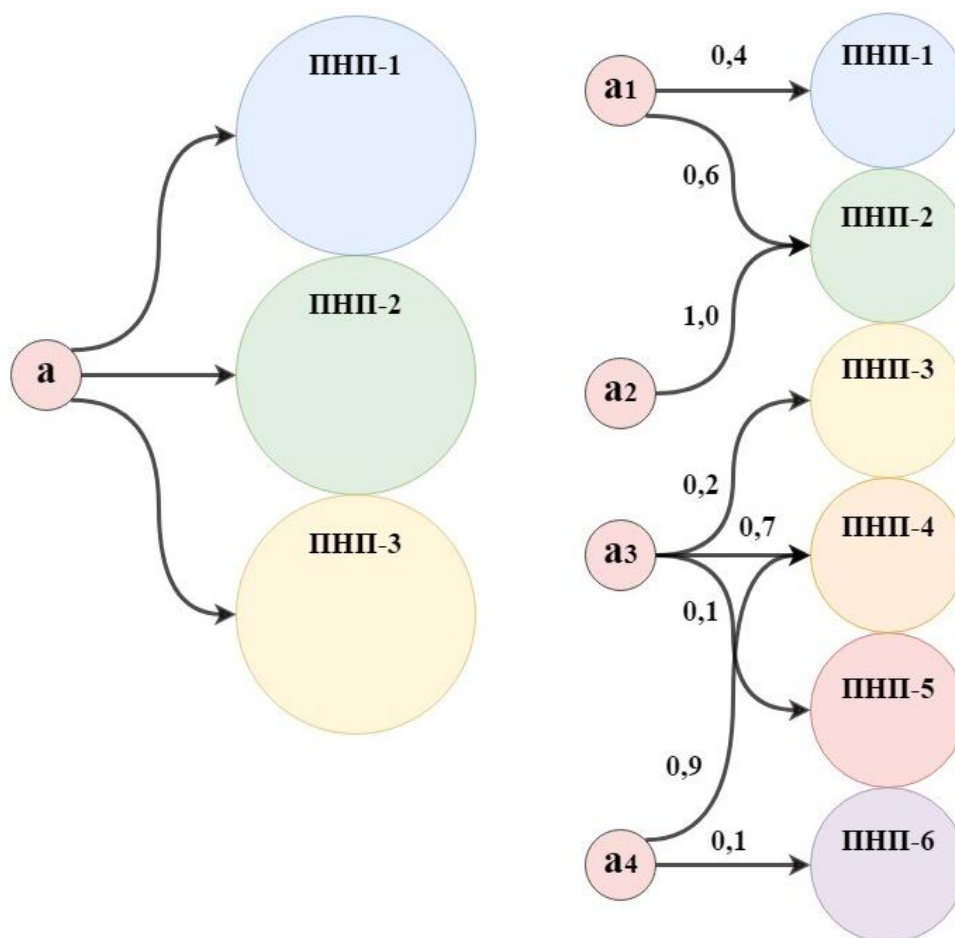


Рисунок 6. – Схематична належність індивідуальних суб'єктів наукової діяльності разом з інформаційними просторами до кількох предметних наукових просторів

Для експериментального аналізу розробленого методу сформовано базу даних з інформацією про 215082 публікації 58834 українських авторів. В результаті застосування методу кластеризації графу цитування було виділено близько 20 тисяч кластерів. При проведенні процедури об'єднання кластерів було виявлено, що зі зменшенням кількості кластерів зростає нестабільність кластеризації графу цитування. Причиною нестабільності є недостатня зв'язність графу. Для вирішення даної проблеми необхідне доповнення вхідних даних публікаціями закордонних авторів. Застосування методу об'єднання кластерів, що базується на знаходженні відстані між анотаціями, зменшує кількість кластерів до близько 500. Після застосування латентно-семантичного аналізу було отримано близько 250 кластерів. Це значно спрощує подальший аналіз та оцінювання даних напрямів.

Використання кожного способу окремо приводить до виникнення ізольованих кластерів публікацій. Зокрема існують групи вчених, які паралельно досліджують одну тематику, але між їх публікаціями є мало цитувань, або вони взагалі відсутні. Аналогічно, враховуючи різні авторські стилі написання, існують анотації публікацій, що акцентуються на різних аспектах однієї і тієї ж проблеми, а тому вони

є достатньо далекими за змістом. Використання обидвох способів оцінювання відстані між анотаціями надає змогу більш точно визначати публікації що належать до спільного напрямку досліджень.

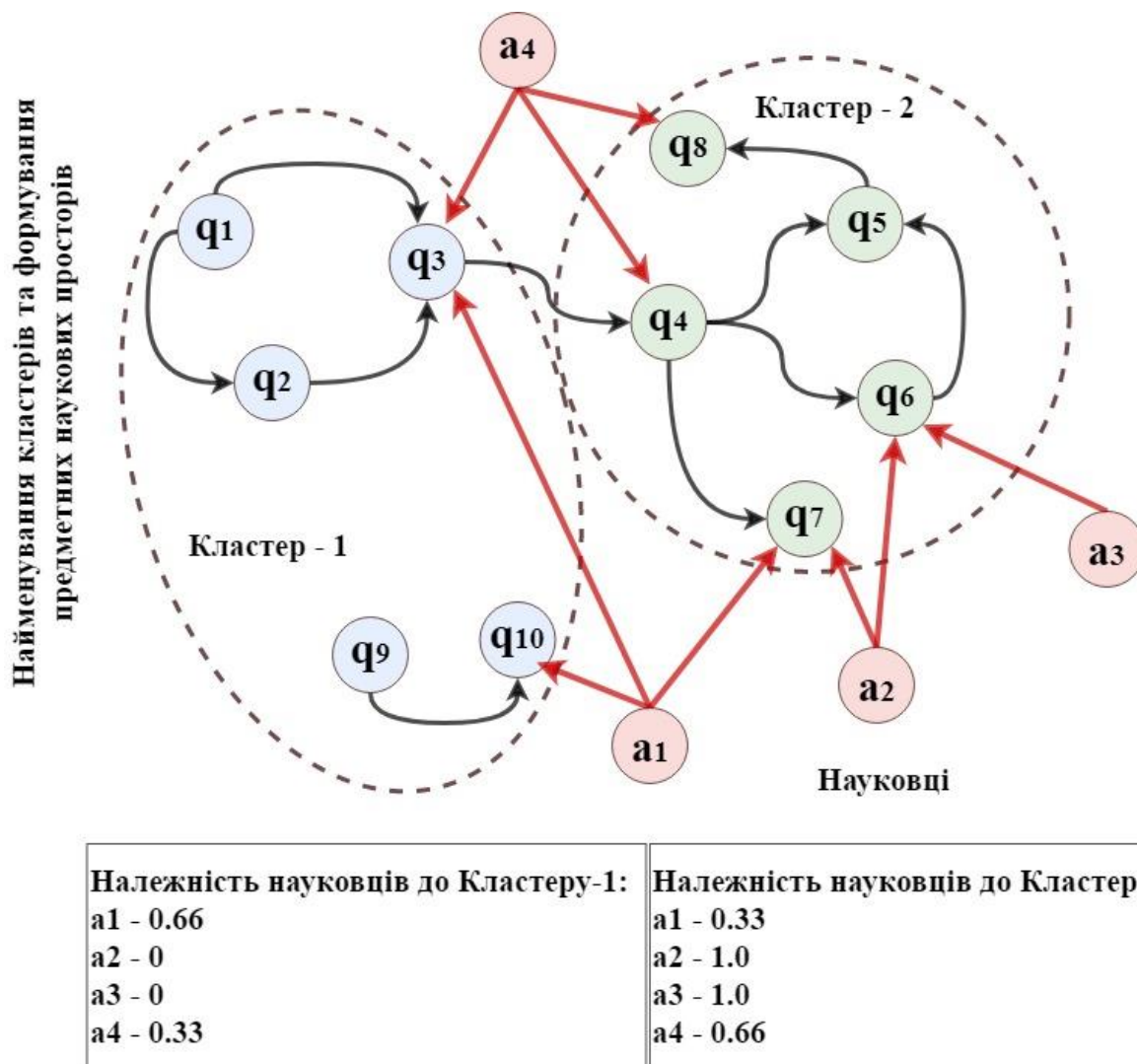


Рисунок 7. – Схема найменування кластерів для ідентифікації предметних наукових просторів

Основними складнощами, які виникають при ідентифікації предметних наукових просторів є:

1. Повнота відомостей про цитування публікацій є необхідною умовою реалізації методу. Саме тому існує необхідність доступу до наукометричних баз даних, які містять ці відомості. Враховуючи тенденції впровадження національних наукометричних баз даних в країнах ЄС, виникає проблема включення до розгляду публікацій іноземних науковців. В разі відсутності таких публікацій граф цитування (Q,C), є неповним і тому виникають незв'язані кластери публікацій, що негативно впливає на результати кластеризації наукових публікацій.

2. Необхідність знаходження локально-чутливого гешу для анотації кожної з розглядуваних публікацій та знаходження матриці відстаней між ними. Кожна із цих

дій має значну обчислювальну складність. Для зменшення навантаження на систему пропонується знаходити геш анотації при додаванні публікації до системи, а надалі зберігати його як одну із характеристик цієї публікації.

3. Визначення порогового значення відстані δ , при якому кластери слід об'єднувати. Проблема полягає в тому, що різні галузі характеризуються різною теоретичною середньою відстанню між науковими публікаціями.

4. Побудова відповідності між кластерами публікацій та напрямками наукових досліджень потребує залучення експертів. В ході функціонування системи виникає база даних основних понять галузі, яка будується на основі ключових слів публікацій та термінів, які найчастіше зустрічаються в анотаціях цих публікацій.

Враховуючи вказані складнощі, для кластеризації наукових публікацій більш доцільно використовувати відстані між публікаціями, визначені на основі цитування. Знаходження таких відстаней потребує меншого обсягу обчислень, ніж знаходження відстаней між публікаціями на основі близькості публікацій за змістом. Також слід врахувати, що граф зв'язку цитування (Q,C) є досить розрідженим: з кожної вершини, як правило, виходить близько 10–20 дуг. А граф, побудований на основі відстані між анотаціями за змістом, є майже повним, тобто кількість ребер що виходить із кожної вершини, трохи менша ніж кількість вершин в графі. Кластеризація розрідженого графу є більш простою задачею, ніж кластеризація повного.

При об'єднанні кластерів, відстань визначається тільки між центрами ваги кластерів, яких значно менше ніж публікацій в цілому. Тому на цьому етапі більш доцільно застосовувати метод розрахунку відстаней між анотаціями наукових публікацій. В такій послідовності дій кластеризація наукових публікацій потребує меншого обсягу обчислень.

Ідентифікація предметних наукових просторів є необхідною задачею для проведення якісного оцінювання продуктивності суб'єктів наукової діяльності, що визначається інформаційним простором даних суб'єктів.

Четвертий розділ присвячено розробленню методів кількісного оцінювання продуктивності суб'єктів наукової діяльності на основі аналізу складників інформаційного простору суб'єктів наукової діяльності.

Розглянемо метод кількісного оцінювання продуктивності суб'єктів наукової діяльності з врахуванням самоцитування. В дослідженні модифіковано ідею, яка використовується для впорядкування веб-сторінок пошуковими системами для оцінювання продуктивності суб'єктів наукової діяльності. Цей метод базується на цитуваннях наукових публікацій суб'єктів наукової діяльності, при чому враховує всі цитування без виключення, на відміну від відомих наукометричних індексів, які враховують тільки ті цитування публікацій, що розміщені в середині ядра. Кількісна оцінка продуктивності суб'єктів наукової діяльності розраховується за такою формулою:

$$\Phi(a_i, t) = \sum_{\substack{j=1 \\ i \neq j}}^n h_{ij} w_j \Phi(a_j, t), \quad i = \overline{1, n}, \quad (38)$$

де

- $\Phi(a_i, t)$ – оцінка продуктивності суб'єкта наукової діяльності a_i ;
- h_{ij} – кількість цитувань публікацій суб'єкта наукової діяльності a_i в наукових публікаціях суб'єкта наукової діяльності a_j ;
- w_j – коефіцієнт, який дозволяє отримати ненульовий розв'язок системи;
- $\Phi(a_j, t)$ – оцінка науково-дослідної діяльності суб'єкта наукової діяльності a_j .

Невідомі оцінки продуктивності знаходимо як розв'язок системи алгебраїчних рівнянь

$$(E - H)\phi = 0, \quad (39)$$

де E – одинична матриця, $H = \{h_{ij}w_j\}_{i,j=1}^n$, $\phi = \{\Phi(a_i, t)\}_{i=1}^n$ – вектор-стовпець оцінок продуктивності.

Ненульовий розв'язок існує, якщо $|E - H| = 0$.

Коефіцієнти системи можна визначити за формулами:

$$h_{ij} = \|C(a_i) \cap \bar{C}(a_j)\|, \quad (40)$$

$$w_j = \|\bar{C}(a_j)\|^{-1}, \quad i = \overline{1, n}, \quad j = \overline{1, n}, \quad (41)$$

де h_{ij} – кількість публікацій суб'єкта наукової діяльності a_j , в яких даний суб'єкт цитує публікації суб'єкта a_i , а w_j – величина, обернена до загальної кількості публікацій суб'єкта наукової діяльності a_j .

Також коефіцієнти системи можна розрахувати так:

$$h_{ij} = \sum_{s=1}^m \frac{\|C(a_i) \cap \{q_s\}\|}{\|A(q_s)\|}, \quad (42)$$

$$w_j = \left(\sum_{s=1}^n \frac{1}{\|A(q_s)\|} \right)^{-1}, \quad i = \overline{1, n}, \quad j = \overline{1, n}, \quad (43)$$

де коефіцієнти h_{ij} та w_j враховують в оцінці кількість співавторів кожної наукової публікації q_s .

Якщо існує нетривіальний розв'язок системи, тоді існує безліч розв'язків, пропорційних даному. Отже після знаходження оцінок продуктивності доцільно здійснити нормування цих оцінок за однією з формул:

$$\Phi'(a_i, t) = \frac{\Phi(a_i, t)}{\sum_{k=1}^n \Phi(a_k, t)}, \quad i = \overline{1, n}, \quad (44)$$

$$\Phi''(a_i, t) = \frac{\Phi(a_i, t)}{\max_{k=1, n} \Phi(a_k, t)}, \quad i = \overline{1, n}, \quad (45)$$

де $\Phi(a_i, t)$ – оцінка продуктивності суб'єкта наукової діяльності a_i , $\Phi'(a_i, t)$ – нормована за сумою оцінка $\Phi(a_i, t)$, $\Phi''(a_i, t)$ – нормована за максимумом оцінка $\Phi(a_i, t)$.

Розглянемо метод зменшення впливу самоцитування на підсумкову кількісну оцінку продуктивності суб'єктів наукової діяльності. Нехай $f(\beta)$ деяка неспадна функція, для якої виконується умова $0 \leq f(\beta) \leq \beta$, для будь-яких значень β . При знаходженні оцінки продуктивності суб'єктів наукової діяльності самоцитування впливає виключно на коефіцієнти h_{ii} . Застосуємо до кожного з них розглянуту функцію $f(\beta)$. Вона дає змогу зменшити вплив самоцитування на визначення кінцевої оцінки за цим методом. Нехай $\tilde{h}_{ii} = f(h_{ii})$, $\tilde{h}_{ij} = h_{ij}$, $i \neq j$, тоді для збереження стійкості методу необхідно покласти \tilde{w}_{ij} таким чином, щоб забезпечувалось виконання умови $|E - \tilde{H}| = 0$. Тоді оцінки продуктивності суб'єктів наукової діяльності з урахуванням самоцитування знайдемо як розв'язок системи лінійних алгебраїчних рівнянь:

$$(E - \tilde{H})\phi = 0, \quad (46)$$

де E – одинична матриця, $\tilde{H} = \{\tilde{h}_{ij} \tilde{w}_{ij}\}_{i, j=1}^n$, $\phi = \{\Phi(a_i, t)\}_{i=1}^n$ – вектор-стовпець оцінок продуктивності.

Нехай $f(\beta) = \frac{\beta}{2}$ тобто важливість самоцитування зменшено в два рази в порівнянні з цитуваннями іншими авторами. Тоді

$$\tilde{h}_{ij} = \|C(a_i) \cap C(a_j)\|, \quad i \neq j, \quad (47)$$

$$\tilde{h}_{ii} = \frac{1}{2} \|C(a_i) \cap \tilde{C}(a_i)\|. \quad (48)$$

Покладемо

$$\tilde{w}_{ii} = \frac{1}{\|\tilde{C}(a_i)\|}, \quad \tilde{w}_{ij} = \frac{1}{\|\tilde{C}(a_j)\|} + \frac{\tilde{h}_{ii}}{(n-1)\tilde{h}_{ij}\|\tilde{C}(a_i)\|}. \quad (49)$$

Рядки матриці $E - \tilde{H}$ будуть лінійно залежні, а отже система лінійних алгебраїчних рівнянь матиме нетривіальні розв'язки.

Нехай Q' – замкнута система публікацій суб'єктів наукової діяльності A' , тобто для них відомі всі публікації та всі цитування C' між ними: $C' = Q' \times Q'$. Назвемо кортеж $\langle A', P', C' \rangle$ ядром. Ядро є замкнутою системою.

Враховуючи, що існують цитування, які не входять до ядра системи, то оцінити продуктивність суб'єктів наукової діяльності за цим методом неможливо.

Припустимо, що продуктивність всіх суб'єктів наукової діяльності, що не входять до ядра системи є рівноцінними. Тоді, не обмежуючи загальності, можемо визначити оцінку їх продуктивності рівну 1. Тоді оцінку продуктивності суб'єктів наукової діяльності $a_i \in A'$, можна розрахувати за формулою:

$$\Phi(a_i, t) = \sum_{\substack{j=1 \\ i \neq j}}^{n'} h_{ij} w_j \Phi(a_j, t) + \sum_{\substack{j=n' \\ i \neq j}}^n h_{ij} w_j, \quad i = \overline{1, n'}, \quad (50)$$

де $\Phi(a_i, t)$ – оцінка продуктивності суб'єкта наукової діяльності a_i , n' – кількість суб'єктів наукової діяльності, що входять до ядра.

В роботі також описано метод знаходження комплексної оцінки продуктивності суб'єктів наукової діяльності, зокрема ЗВО. Згідно з цим методом оцінки за різними категоріями розглядаються як точки в багатовимірному просторі. Комплексна оцінка продуктивності розраховується як об'єм, який утворюється в результаті описування цих точок фігурою.

Для верифікації результатів дослідження здійснено підбір показників діяльності ЗВО, які отримуються з відкритих джерел та відображають основні аспекти діяльності ЗВО. Запропоновано п'ять наборів показників, на основі яких побудовано критерії оцінки міжнародної активності, контингенту студентів, науково-педагогічного персоналу, науково-дослідної діяльності та ресурсного забезпечення. Відповідні оцінки результатів діяльності можуть бути використані ректорами, деканами, завідувачами кафедр та для аналізу ефективності функціонування різними аспектами діяльності підлеглих суб'єктів. Своєчасний моніторинг динаміки зміни результатів діяльності дає змогу, в разі потреби, внесення відповідних коректив до стратегії функціонування суб'єкту з метою її вдосконалення. Обчислювальна задача розрахунку оцінки для п'яти категорій є простою, бо потребує порядку 10^3 арифметичних операцій і може бути виконана процесором з тактовою частотою 1 ГГц за 10^{-3} с.

Розглянемо метод кількісного оцінювання перспекції продуктивності суб'єктів наукової діяльності. Нехай для кожного науковця a_i вже задано кількісні оцінки продуктивності результатів їх наукової діяльності і проведено процедуру ідентифікації предметних наукових напрямів. Нехай K_1, K_2, \dots, K_h – предметні наукові простори, які складаються з науковців, які споріднені певним науковим напрямом; h – кількість просторів. Оскільки значення показників оцінок для кожного науковця динамічно змінюються з часом, то будемо вважати, що якщо фіксувати ці значення з певною періодичністю (квартал, півріччя, рік тощо), то можемо отримати значення, які будуть відображати динаміку зміни продуктивності науково-дослідної діяльності кожного науковця в моменти часу t_1, t_2, \dots, t_s :

$$(q_{i1}^d, q_{i2}^d, \dots, q_{is}^d) = (q_i^d(t_1), q_i^d(t_2), \dots, q_i^d(t_s)), \quad (51)$$

де q_{ir}^d – оцінка результатів науково-дослідної діяльності науковців a_i , які зафіксовані в дискретні моменти часу t_r , $r = \overline{1, s}$, причому науковці a_i належать до простору K_d , $d = \overline{1, h}$, $t_r = t_1 + (r-1) \cdot \Delta t$, $r = \overline{2, s}$, $t_1 < t_2 < \dots < t_s$, t_1 – початковий момент

часу; Δt – часовий інтервал через який фіксуються кількісні оцінки продуктивності науково-дослідної діяльності. Тобто

$$q_{ip}^d = \Phi(Q(a_i, t_p), C(a_i, t_p)), \quad a_i \in K_d, \quad d = \overline{1, h}, \quad (52)$$

де $Q(a_i, t_p)$ – множина публікацій науковців a_i , які були опубліковані до моменту часу t_p , тобто розглядаються всі публікації науковця, які були опубліковані в моменти часу t_1, t_2, \dots, t_p , $p = \{1, 2, \dots, s\}$, тобто в моменти $t_c \leq t_p$ при $c = \overline{0, p-1}$; $C(a_i, t_p)$ – множина публікацій, у яких цитувалися публікації науковців a_i , в моменти часу $t_c \leq t_p$ при $c = \overline{0, p-1}$; q_{ip}^d – оцінка результатів науково-дослідної діяльності a_i , що належать до предметного наукового простору K_d за період часу, що передує моменту часу t_p .

Розрахуємо сумарні кількісні оцінки продуктивності науково-дослідної діяльності за всіма суб'єктами наукової діяльності, які належать до відповідного предметного наукового простору або кластеру K_d за період часу, що передує моменту часу t_r , $r = \overline{1, s}$:

$$\bar{q}_r^d = \sum_{i=1}^{\text{card}(K_d)} q_{ir}^d, \quad (53)$$

де \bar{q}_r^d – сумарна оцінка результатів науково-дослідної діяльності суб'єктів наукової діяльності, які належать предметного наукового простору K_d за період часу, що передує моменту часу t_r ; $\text{card}(K_d)$ – кількість суб'єктів наукової діяльності, які належать до предметного наукового простору K_d .

Позначимо через Q_r^d множину всіх публікацій суб'єктів наукової діяльності, які належать до предметного наукового простору K_d . Причому наукові публікації були опубліковані в моменти часу що передують t_r . Надалі будемо вважати, що

$$\text{card}(Q_r^1) \geq \text{card}(Q_r^2) \geq \dots \geq \text{card}(Q_r^h), \quad r = \overline{1, s}, \quad (54)$$

тобто, простори суб'єктів наукової діяльності впорядковані в порядку неспадання кількості публікацій в кожен з моментів часу t_1, t_2, \dots, t_s . Тоді основні індекси цитування можуть бути розраховані за формулами:

$$h_r^d = \max_{y=1, m} \min \left\{ y, \text{card}(Q_r^d) \right\}, \quad (55)$$

де h_r^d – індекс Гірша предметного наукового простору K_d , $d = \overline{1, h}$ в момент часу t_r , $r = \overline{1, s}$,

$$g_r^d = \max_{y=1, m} \min \left\{ y, \left[\sqrt{\sum_{x=1}^y \text{card}(Q_r^d)} \right] \right\}, \quad (56)$$

де g_r^d – g-індекс предметного наукового простору K_d , $d = \overline{1, h}$ в момент часу t_r , $r = \overline{1, s}$

$$e_r^d = \sqrt{\sum_{y=1}^{h_r^d} \text{card}(Q_r^d) - (h_r^d)^2}, \quad (57)$$

де e_r^d – е-індекс предметного наукового простору K_d , $d = \overline{1, h}$ в момент часу t_r , $r = \overline{1, s}$, а h_r^d – індекс Гірша,

$$i10_r^d = \text{card}(C10_r^d), \quad (58)$$

де $i10_r^d$ – індекс і-10 предметного наукового простору K_d , $d = \overline{1, h}$ в момент часу t_r , $r = \overline{1, s}$, а $C10_r^d$ – це підмножина публікацій Q_r^d , які цитувались в інших публікаціях 10 або більше разів, тобто:

$$C10_r^d = \{q_y \in Q_r^d \mid \text{card}(C(q_y)) \geq 10\}, \quad (59)$$

де $\text{card}(C(p_y))$ – кількість публікацій, в яких цитується публікація p_y .

Для знаходження кількісних оцінок продуктивності суб'єктів наукової діяльності також можна використати формулу (38).

Якщо розрахувати такі оцінки для різних періодів часу, то отримаємо дискретний часовий ряд $(\bar{q}_1^d, \bar{q}_2^d, \dots, \bar{q}_s^d)$. Для аналізу динаміки результатів науково-дослідної діяльності за заданим предметним науковим простором знайдемо відношення $\frac{\bar{q}_2^d}{\bar{q}_1^d}, \frac{\bar{q}_3^d}{\bar{q}_2^d}, \dots, \frac{\bar{q}_s^d}{\bar{q}_{s-1}^d}$, тобто побудуємо величини:

$$V_a^d = \frac{\bar{q}_b^d}{\bar{q}_a^d}, \quad b = 2, 3, \dots, s, \quad a = 1, 2, \dots, s-1, \quad (60)$$

де V_a^d – потенціал предметного наукового простору K_d .

Тоді часовий ряд потенціалів предметних наукових просторів K_d має вигляд

$$(V_1^d, V_2^d, \dots, V_{s-1}^d), \quad d = \overline{1, h}. \quad (61)$$

Прогноз потенціалу предметного наукового простору з горизонтом 1 визначається як деякий функціонал F від ретроспективних даних:

$$\hat{V}_s^d = F(V_{s-u}^d, V_{s-u+1}^d, \dots, V_{s-1}^d), \quad (62)$$

де u – кількість ретроспективних даних, які обираються для розрахунку прогнозу.

Прогноз потенціалу предметного наукового простору з горизонтом 2 визначається так:

$$\hat{V}_{s+1}^d = F(V_{s-u+1}^d, V_{s-u+2}^d, \dots, V_{s-1}^d, \hat{V}_s^d), \quad (63)$$

\hat{V}_{s+1}^d – прогноз потенціалу предметного наукового простору на 2 точки вперед; F – функціонал, який визначає прогноз потенціалу. Наприклад, за F можна обрати лінійно-зважене плинне середнє:

$$\hat{V}_s^d = \left(\sum_{k=1}^u k \right)^{-1} \sum_{k=1}^u k \cdot V_{s-u+k-1}^d, \quad d = \overline{1, h}, \quad (64)$$

де \hat{V}_s^d – прогноз потенціалу предметного наукового простору на 1 точку вперед на основі лінійно-зваженого плинного середнього.

Описаний метод дозволяє оцінити потенціал предметного наукового простору або розрахувати перспекцію продуктивності суб'єктів наукової діяльності та предметних наукових просторів, в яких вони об'єднуються.

Якщо $\hat{V}_s^d < 1$, то предметний науковий простір K_d , має негативну динаміку та потенціал, якщо $\hat{V}_s^d > 1$ – потенціал позитивний, якщо ж $\hat{V}_s^d = 1$ – приросту потенціалу немає, отже жодних змін як позитивних, так і негативних в результатах розвитку предметного наукового простору немає.

У п'ятому розділі описано задачу вибору суб'єктів наукової діяльності для організації співпраці в межах наукових та освітніх проєктів на основі аналізу моделі ідентифікації науковця.

Розглянемо математичну модель та метод багатокритеріального вибору потенційних партнерів для співпраці в рамках наукових та освітніх проєктів. Нехай задана скінчена множина наукових та освітніх проєктів або грантів $G = \{G_1, G_2, \dots, G_n\}$, n – кількість проєктів, для яких потрібно підібрати виконавців. Нехай задано скінчена множина потенційних виконавців цих проєктів $V = \{v_1, v_2, \dots, v_t\}$, t – кількість потенційних виконавців як суб'єктів освітньо-наукового середовища. Виконавцями можуть бути науковці, менеджери проєктів, науково-дослідні установи, заклади вищої освіти тощо.

Будь який проєкт, складається з ряду робочих пакетів $G_i = \{g_1^i, g_2^i, \dots, g_{r_i}^i\}$, r_i – кількість робочих пакетів проєкту G_i , $i = \overline{1, n}$, які виконуються у визначеній послідовності і зв'язані результатами. Для виконання кожного з цих робочих пакетів необхідно вибрати виконавців, які мають досвід та компетентності для виконання пакету вчасно і ефективно. Тобто необхідно знайти такі множини потенційних виконавців:

$$W(g_j^i) = \{v_d \in V \mid (v_d, g_j^i) \in Q^i\}, \quad Q^i \subset V \times G_i, \quad j = \overline{1, r_i}, \quad i = \overline{1, n}, \quad d = \overline{1, t}. \quad (65)$$

Для кожного робочого пакета кожного проєкту має бути сформовано перелік ключових критеріїв відбору партнерів. Тобто вектори критеріїв оцінювання матимуть вигляд:

$$f^{ij}(v) = (f_1^{ij}(v), f_2^{ij}(v), \dots, f_{N_{ij}}^{ij}(v)), \quad v \in V, \quad (66)$$

N_{ij} – кількість критеріїв оцінювання потенційних партнерів робочих пакетів g_j^i проєктів G_i , $j = \overline{1, r_i}$, $i = \overline{1, n}$.

Деякі критерії максимізуються, тому множину індексів таких критеріїв позначимо через $J_1^{ij} = \{1, 2, \dots, h_{ij}\}$. Інші критерії з індексами $J_2^{ij} = \{h_{ij} + 1, h_{ij} + 2, \dots, N_{ij}\}$ – мінімізуються, $J^{ij} = \{1, 2, \dots, N_{ij}\}$, $J_1^{ij} \cup J_2^{ij} = J^{ij}$. Тоді

$$\sum_{k \in J_1^{ij}} \lambda_k f_k^{ij}(v) \rightarrow \max, \quad \sum_{k \in J_1^{ij}} \lambda_k = 1, \quad (67)$$

$$\sum_{k \in J_2^{ij}} \delta_k f_k^{ij}(v) \rightarrow \min, \quad \sum_{k \in J_2^{ij}} \delta_k = 1, \quad (68)$$

для кожного пакету і проєкту будують обмеження множини потенційних виконавців $v \in V^{ij}$, $V^{ij} = \{v \in V \mid y_u^{ij}(v) \geq \rho_u^{ij}, u = \overline{1, z_{ij}}, j = \overline{1, r_i}, i = \overline{1, n}\}$, де z_{ij} – кількість

порогових значень для вектор-функції обмежень $y_u^{ij}(v)$. Коефіцієнти λ_k та δ_k визначають важливість кожного з критеріїв у розрахунку комплексної оцінки.

Рішення про те, чи обирати виконавців для реалізації відповідного пакету приймає менеджерський склад проекту або особа, яка приймає рішення. Для визначення оптимального складу виконавців робочих пакетів кожного з проєктів можна використати метод агрегації експертних оцінок.

Нехай задана множина експертів $E = \{E_1, E_2, \dots, E_s\}$, s – кількість експертів. Кожен з експертів складає переваги потенційних виконавців з урахуванням вектору критеріїв. Допускається неповний профіль переваг. Нехай $\xi_{c,b}^{ij}$ – середня частота появи кожної з переваг між потенційними виконавцями v_c та v_b , $c \neq b$, $v_c \in V$, $v_b \in V$. Тоді отримаємо матриці переваг вигляду:

$$\Psi^{ij} = \begin{pmatrix} \xi_{1,1}^{ij} & \xi_{1,2}^{ij} & \dots & \xi_{1,t}^{ij} \\ \xi_{2,1}^{ij} & \xi_{2,2}^{ij} & \dots & \xi_{2,t}^{ij} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{t,1}^{ij} & \xi_{t,2}^{ij} & \dots & \xi_{t,t}^{ij} \end{pmatrix}, \quad (69)$$

$$j = \overline{1, r_i}, \quad i = \overline{1, n}.$$

Використовуючи методи формування колективного рішення, за матрицею парних порівнянь для кожного пакету g_j^i проєкту G_i можна отримати впорядкований перелік потенційних виконавців:

$$v_{k_1}^{ij} \succ v_{k_2}^{ij} \succ \dots \succ v_{k_t}^{ij}, \quad k_1 < k_2 < \dots < k_t, \quad k_q \in \{1, 2, \dots, t\}, \quad v_{k_q}^{ij} \in V^{ij}, \quad q = \overline{1, t}. \quad (70)$$

Враховуючи вказаний перелік, менеджер проєкту підбирає конкретних виконавців і формує робочу групу.

Загальна постановка задачі може змінюватись залежно від того які партнери розглядаються. Потенційними партнерами можуть бути університети, науково-дослідні установи, приватні компанії, державні організації, окремі науковці, тобто колективні та індивідуальні суб'єкти наукової діяльності. Залежно від цілей таких партнерів в проєкті можна підібрати відповідні критерії для їх кількісного оцінювання. Показано ефективність використання методології формування інформаційних просторів суб'єктів наукової діяльності для виявлення потенційних суб'єктів наукової діяльності для реалізації дослідницьких та освітніх проєктів за результатами кількісного оцінювання їх продуктивності, що забезпечує сталість розвитку наукових організацій.

Подібний метод може бути застосовано при формуванні організаційної структури колективних суб'єктів наукової діяльності, зокрема кафедр, факультетів та університетів. Отримані результати дали можливість вдосконалити векторний підхід в методології проєктно-векторного управління в освітніх середовищах.

У шостому розділі описано інформаційну технологію, в якій реалізовано методологію формування інформаційних просторів суб'єктів наукової діяльності та практично застосовано розроблені і вдосконалені моделі та методи. В основі технології є інформаційно-аналітична система оцінювання продуктивності

індивідуальних та колективних суб'єктів наукової діяльності.

Ключовою концепцією в реалізації системи оцінювання продуктивності наукової діяльності є забезпечення максимальної відкритості та доступності для академічної спільноти. Також важливим положенням є використання тільки надійних, перевірених джерел інформації, що мінімізують вплив суб'єктивних оцінок і фіксуються за результатами реальної роботи суб'єктів наукової діяльності. Для забезпечення цих положень, було створено архітектуру системи, в якій взаємодія системи з користувачами відбувається через мережу Інтернет.

Структурна схема інформаційної системи оцінювання продуктивності наукової діяльності (рис. 8) включає модуль збору інформації, модуль обробки інформації, базу даних та веб-сервер. Цільовою аудиторією користувачів системи є відповідальні особи Міністерства освіти і науки України, керівний склад закладів вищої освіти та окремі індивідуальні суб'єкти наукової діяльності.

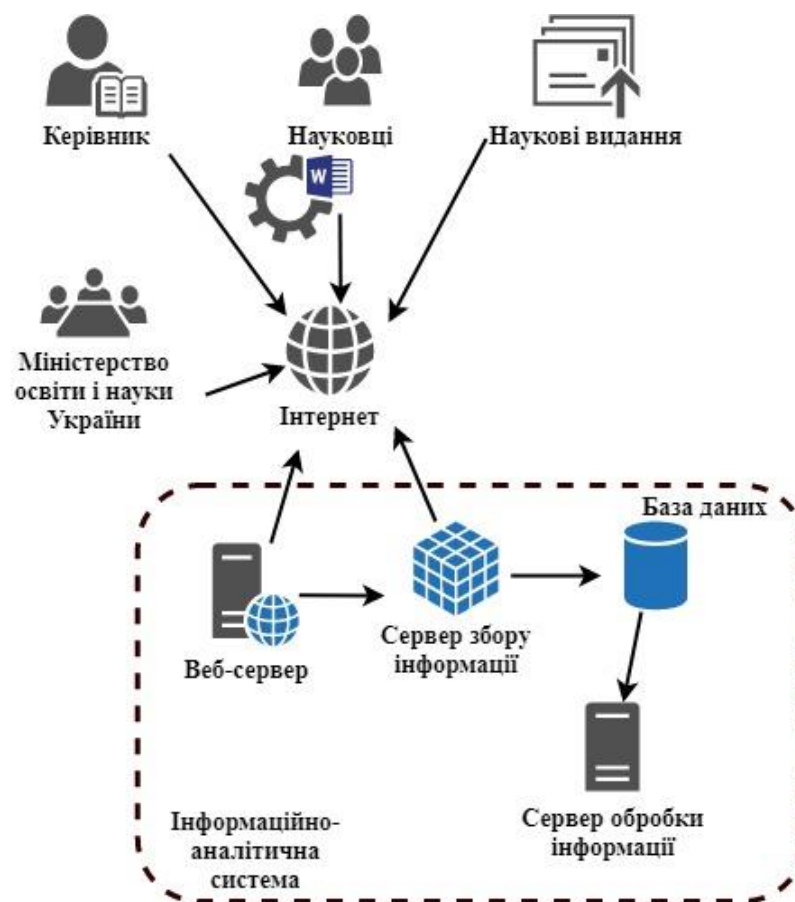


Рисунок 8. – Структурна схема інформаційної системи оцінювання продуктивності наукової діяльності

Архітектура інформаційної системи представлена блоками, що включає окремі мікросервіси для реалізації таких задач:

- комплекс для збору, опрацювання, зберігання інформації про продуктивність суб'єктів наукової діяльності;
- методи формування предметних наукових просторів;
- методи оцінювання продуктивності суб'єктів наукової діяльності;

- методи формування організаціо-функціональної будови суб'єктів наукової діяльності.

Система може працювати як на одному сервері, так і бути розділена на кілька різних серверів для оптимізації навантаження. Якщо компоненти системи знаходяться на різних серверах, між ними необхідно забезпечити зв'язок за допомогою локальної мережі або глобальної мережі Інтернет. Кожен компонент має власні API для обміну задачами між сервісами та результатами їх виконання незалежно від їх фізичного розміщення.

Вся інформація зберігається в реляційній базі даних. База має фізичну реалізацію в системі управління PostgreSQL, але керування базою здійснюється шляхом взаємодії із нею ORM Django.

Основними моделями в базі даних є:

1. Користувач (містить інформацію про логін, пароль, електронну пошту користувача та його права)
2. Науковець (містить інформацію про науковця: прізвище, ім'я, по батькові на Українській, Російській та Англійській мовах, вчене звання та науковий ступінь, відомості про місце роботи шляхом зв'язку із відповідними моделями, реалізує зв'язки із публікаціями користувача та оцінками його наукової діяльності)
3. Публікація (містить інформацію про публікацію: заголовок, джерело, посилання на файл, зв'язок із авторами та іншими публікаціями через цитування)
4. Цитата (реалізація зв'язку між публікаціями, містить також додаткову інформацію про джерело, звідки отримана інформація про цитату, бібліографічне посилання тощо)
5. Заклад вищої освіти (інформація про ЗВО: назва, тип, контактна інформація, інформація про структурні підрозділи та Також реалізує зв'язок із користувачами для визначення відповідальних осіб (адміністраторів))
6. Структурний підрозділ (кафедра чи інститут, реалізує каскадний зв'язок між ВНЗ та науковцем)
7. Кількісна оцінка продуктивності наукової діяльності (оцінка містить наукомеричні показники знайдені для відповідного об'єкта та дату їх знаходження).

Створено інформаційну технологію, в якій реалізовано методологію формування інформаційних просторів суб'єктів наукової діяльності, що включає модулі: оцінювання продуктивності наукової діяльності, побудова предметних наукових просторів. Використовуючи задачу багатокритеріального вибору, в системі реалізовано метод вибору науковців для створення наукового або освітнього проекту, що дозволяє забезпечити сталість розвитку колективних суб'єктів наукової діяльності: наукових організацій та закладів вищої освіти.

ВИСНОВКИ

В дисертаційній роботі представлені результати узагальнення й розв'язку важливої науково-технічної проблеми: створити методологію, яка дозволить здійснити формування інформаційних просторів суб'єктів наукової діяльності, що є передумовою сталого розвитку організацій, включаючи процедури прозорого оцінювання продуктивності суб'єктів наукової діяльності та їх розвитку в динаміці. Це дасть змогу вдосконалити управління освітньо-науковою сферою України, що забезпечується сталим розвитком суб'єктів наукової діяльності в освітньо-науковому

просторі України на основі результатів кількісного оцінювання їх продуктивності.

Виконані дослідження дають підстави зробити низку висновків:

1. Проведено аналіз стану завдань формування інформаційних просторів суб'єктів наукової діяльності, оцінювання продуктивності суб'єктів наукової діяльності, виявлення предметних наукових просторів, пошук партнерів для співпраці в рамках освітніх та наукових проєктів. Побудовані мережі цитування та мережі наукової співпраці реальних наукових колективів, що є основою аналізу їх продуктивності.

2. Розроблено моделі ідентифікації індивідуальних та колективних суб'єктів наукової діяльності, зокрема науковців та закладів вищої освіти. Побудовано формалізовану множинну модель інформаційного простору суб'єктів наукової діяльності, на основі якої побудовано концептуальну схему їх взаємодії та визначено принцип методології формування інформаційних просторів.

3. Розроблено методологію формування інформаційних просторів суб'єктів наукової діяльності, яка дозволяє формувати команди освітніх та наукових міжнародних проєктів зі складу провідних вчених світу, що включає моделі та методи ідентифікації предметних наукових просторів та оцінювання продуктивності суб'єктів наукової діяльності в рамках даних просторів, що дозволяє підвищити обґрунтованість прийняття рішень з вибору виконавців наукових проєктів, забезпечити сталість розвитку закладів вищої освіти і, як наслідок, збільшити продуктивність суб'єктів наукової діяльності в цілому. Сформовано понятійний апарат методології, що дозволяє систематизувати інформацію про інформаційні простори.

4. Розроблено інформаційну технологію, яка реалізує методологію формування інформаційних просторів суб'єктів наукової діяльності з використанням технології збору інформації з відкритих джерел, розрахунків та формування рейтингів суб'єктів наукової діяльності у відповідних предметних наукових просторах, створення інфографіки на основі зібраної інформації, що дозволяє дослідити складові інформаційних просторів суб'єктів наукової діяльності і використати їх для забезпечення сталого розвитку закладів вищої освіти, в складі яких є дані суб'єкти.

5. Набули подальшого розвитку методи формування предметних наукових просторів на основі кластерного аналізу параметрів моделі ідентифікації суб'єкта наукової діяльності та на основі латентно-семантичного аналізу. Кластеризація реалізується на основі встановлення метричної відстані між публікаціями суб'єктів наукової діяльності за графом цитування та з врахуванням близькості анотацій наукових публікацій суб'єктів наукової діяльності за методом локально-чутливого гешування з використанням n -грамів. Після кластеризації публікацій здійснено найменування кластерів та побудовано предметні наукові простори. Вдосконалено метод найменування кластерів наукових публікацій для побудови предметних наукових просторів, що дозволяє встановити вербальний зв'язок між кластерами наукових публікацій та науковими напрямками. Тематична модель дозволяє наперед задати теми кластерів, таким чином спрощуючи їх найменування, що виконується експертним методом.

6. Вдосконалено метод кількісного оцінювання продуктивності індивідуального та колективного суб'єктів наукової діяльності з врахуванням

самоцітування, а також метод оцінювання перспекції продуктивності суб'єктів наукової діяльності на основі ретроспективних даних.

7. Вдосконалено математичну модель та метод багатокритеріального вибору потенційних партнерів для співпраці в рамках наукових та освітніх проєктів, що враховує організаційну будову проєктів та результати кількісного оцінювання продуктивності суб'єктів наукової діяльності відповідного предметного наукового простору, яка відповідає робочому пакету проєкту, що дозволяє сформувати робочі групи та склад структурних підрозділів наукових установ та закладів вищої освіти.

8. Набув подальшого розвитку векторний підхід до формування просторів суб'єктів наукової діяльності методології проєктно-векторного управління в освітніх середовищах, що включає нові та вдосконалені методи і моделі, які дозволяють оцінювати продуктивність наукової діяльності, ідентифікувати предметні наукові простори, формувати робочі групи наукових проєктів та склад структурних підрозділів закладів вищої освіти.

9. Проведена верифікація результатів досліджень. Описано модуль візуалізації та інтерфейсу користувача в програмному комплексі формування інформаційних просторів суб'єктів наукової діяльності. Розроблені методи, моделі та інформаційна технологія можуть бути використані для практичної організації наукової діяльності та формування організаційно-функціональної будови суб'єктів наукової діяльності: закладів вищої освіти та їх структурних підрозділів (кафедр, факультетів, інститутів), наукових груп. Розроблена технологія надає інструментарій для закладів вищої освіти, Міністерства освіти і науки України та інших відомств здійснювати стале управління науковою діяльністю, раціонально розподіляючи фінансування тих або інших напрямків досліджень з потенційно максимальною продуктивністю. В умовах обмежених ресурсів це має значне практичне значення для України.

Наукові положення, висновки, пропозиції і рекомендації дисертаційної роботи можуть бути використані для практичної організації наукової діяльності з управління ЗВО. Робота впроваджена в діяльності ІТ асоціації в частині створення консорціуму партнерів для організації спільної проєктної наукової діяльності, а також у навчальний процес закладу вищої освіти.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

*Наукові праці, в яких опубліковані основні результати дисертації
Статті у міжнародних наукових виданнях і тих, що входять до
міжнародних наукометричних баз (МНБД):*

1. **Кучанський О.Ю.**, Білощицький А.О. Прогнозування часових рядів методом селективного зіставлення зі зразком. Eastern-European Journal of Enterprise Technologies. 2015. № 6/4(78). С. 13 – 18.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CАplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить метод індексації часових рядів для знаходження в їх динаміці подібних ділянок на основі методу К-найближчих сусідів, а також селективного групування цих ділянок за знаками приростів.

2. Lizunov P., Biloshchytskyi A., **Kuchansky A.**, Biloshchytska S., Chala L. Detection of near duplicates in tables based on the locality-sensitive hashing method and the

nearest neighbor method. Eastern-European Journal of Enterprise Technologies. 2016. № 6/4 (84). P. 4 – 10.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить метод ідентифікації неповних дублікатів, які наявні в даних таблиці відносно множини таблиць, які відібрані з наукових публікацій.

3. Biloshchytskyi A., Andrashko Yu., **Kuchansky A.**, Biloshchytska S., Kuzka O., Terentyev O. Evaluation methods of the results of scientific research activity of scientists based on the analysis of publication citations. Eastern-European Journal of Enterprise Technologies. 2017. № 3/2 (87). P. 4 – 10.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить метод оцінювання продуктивності індивідуальних суб'єктів наукової діяльності, який дозволяє розрахувати скалярну оцінку результатів наукової діяльності і базується на визначенні ряду дійсних коефіцієнтів, що визначають цитування одного науковця в публікаціях інших науковців.

4. Biloshchytskyi A., **Kuchansky A.**, Andrashko Yu., Biloshchytska S., Kuzka O., Shabala Ye., Lyashchenko T. A method for the identification of scientists' research areas based on cluster analysis of scientific publications. Eastern-European Journal of Enterprise Technologies. 2017. № 5/2 (87). P. 4 – 11.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2. Категорія Q2.

Автору належить метод ідентифікації напрямів наукових досліджень або предметних наукових просторів на основі кластеризації наукових публікацій.

5. Biloshchytskyi A., Myronov V., Reznik R., **Kuchansky A.**, Andrashko Yu., Paliy S., Biloshchytska S. A method to evaluate the scientific activity quality of HEIs based on a scientometric subjects presentation model. Eastern-European Journal of Enterprise Technologies. 2017. № 6/2 (90). P. 16 – 22.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить метод інтегральної оцінки якості наукової діяльності закладів вищої освіти та інших структурних науково-дослідних підрозділів або організацій, які займаються науковою діяльністю.

6. **Kuchansky A.**, Biloshchytskyi A., Andrashko Yu., Biloshchytska S., Shabala Ye., Myronov O. Development of adaptive combined models for predicting time series based on similarity identification. Eastern-European Journal of Enterprise Technologies. 2018. № 1/4 (91). P. 25 – 28.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить адаптивні комбіновані моделі гібридного та селективного типів з ідентифікацією подібностей для дослідження характеру зміни дискретних показників.

7. **Kuchansky A.**, Andrashko Yu., Biloshchytskyi A., Danchenko O., Parionov O., Vatskel I., Honcharenko T. The method for evaluation of educational environment subjects'

performance based on the calculation of volumes of m-simplexes. *Eastern-European Journal of Enterprise Technologies*. 2018. № 2/4 (92). P. 15 – 25.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить метод комплексного оцінювання результатів діяльності суб'єктів освітніх середовищ, зокрема закладів вищої освіти.

8. Biloshchytskyi A., **Kuchansky A.**, Paliy S., Biloshchytska S., Bronin S., Andrashko Yu., Shabala Ye., Vatskel V. Development of technical component of the methodology for project-vector management of educational environments. *Eastern-European Journal of Enterprise Technologies*. 2018. № 2/2 (92). P. 4 –13.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить концепція застосування методу для формування плану реалізації наукових проєктів, який є допустимим за розподілом ресурсів та близьким до оптимального за часом виконання проєктів.

9. Lizunov P., Biloshchytskyi A., **Kuchansky A.**, Andrashko Yu., Biloshchytska S. Improvement of the method for scientific publications clustering based on n-gram analysis and fuzzy method for selecting research partners. *Eastern-European Journal of Enterprise Technologies*. 2019. № 4/4 (100). P. 6 –14.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить комплексний метод, який складається з двохетапного методу кластеризації графу цитування публікацій науковців та методу нечіткого логічного виводу для узгодження думок експертів щодо вибору потенційних партнерів і включення їх до проєктної групи.

10. Вukov V., Biloshchytskyi A., **Kuchansky A.**, Andrashko Yu., Dikhtiarenko O., Budnik S. Development of information technology for complex evaluation of higher education institutions. *Information Technologies and Learning Tools*. 2019. 73(5). P. 293 – 306.

Видання індексовано в МНБД: Web of Science

Автору належить метод комплексного оцінювання продуктивності закладів вищої освіти.

11. Вukov V., Spirin O., Biloshchytskyi A., **Kuchansky A.**, Dikhtiarenko O. Open digital systems for assessment of pedagogical research results. *Information Technologies and Learning Tools*. 2020. 75(1). P. 294 – 315.

Видання індексовано в МНБД: Web of Science

Автору належить описання використання кількісних показників при оцінці результатів роботи вчених, науково-дослідних робіт та науково-дослідних установ.

12. Lizunov P., Biloshchytskyi A., **Kuchansky A.**, Andrashko Yu., Biloshchytska S. The use of probabilistic latent semantic analysis to identify scientific subject spaces and to evaluate the completeness of covering the results of dissertation studies *Eastern-European Journal of Enterprise Technologies*. 2020. № 4/4 (106). P. 14–20.

Видання індексовано в МНБД: SCOPUS, Index Copernicus, EBSCO, CAplus, DOAJ, ROAD, MIAR, Open AIRE, WorldCat, CNKI SCHOLAR, Scilit, ResearchBib, EuroPub, CORE. Категорія Q2.

Автору належить застосування латентно-семантичного аналізу для задач ідентифікації предметних наукових просторів.

Статті в наукових фахових виданнях України, що входять до наукометричних баз даних:

13. **Кучанський О.Ю.**, Ніколенко В.В. Прогнозування часових рядів методом зіставлення зі зразком. Управління розвитком складних систем, КНУБА. 2015. Вип. 22. С. 101 – 106.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належить опис теоретичних викладок для прогнозування булевих часових рядів та дискретних часових рядів.

14. **Кучанський О.Ю.**, Мазурак В.В. Перспективи розвитку в Україні систем дистанційного банківського обслуговування клієнтів. Управління розвитком складних систем, КНУБА. 2015. Вип. 23. С. 115 – 119.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належить формування пропозицій щодо побудови інформаційно-аналітичних систем для дистанційного обслуговування.

15. **Кучанський О.Ю.**, Ніколенко В.В., Раченко А.В. Методи ідентифікації тенденцій фінансових часових рядів на основі трендових моделей прогнозування. Управління розвитком складних систем, КНУБА. 2015. Вип. 24. С. 115 – 119.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належить метод, який дозволяє ідентифікувати моменти зміни тенденцій дискретних показників, тобто точки локального максимуму та мінімуму з різними потужностями правих і лівих плечей на основі комплексу з наївних та трендових моделей: плинні, ірраціональні, дробові, різницеві.

16. Білощицький А.О., **Кучанський О.Ю.**, Безмогоричний Д.М., Пида С.В., Кузьомко А.С. Формування концепцій побудови інфраструктури GameHub в українських університетах. Управління розвитком складних систем, КНУБА. 2016. Вип. 26. С. 163 – 170.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належить концепція, завдання та особливості реалізації наукових і освітніх проєктів.

17. Лізунов П.П., Білощицький А.О. **Кучанський О.Ю.**, Чала Л.Е., Білощицька С.В., Удовенко С.Г. Гібридний підхід до аналізу та розпізнавання математичних формул з метою виявлення в них подібностей. Управління розвитком складних систем. 2016. № 27. С. 145 – 155.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належить гібридний підхід, що передбачає використання шаблонів, створених відповідно до особливостей графічних редакторів, та спеціальних конверторів формул з різних форматів до канонічного формату.

18. Лізунов П.П., Білощицький А.О., **Кучанський О.Ю.**, Чала Л.Е., Білощицька С.В., Удовенко С.Г. Автоматичний аналіз подібностей схем та діаграм в електронних текстових документах. Управління розвитком складних систем. 2016. № 28. С. 147 – 156.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належить підхід, що передбачає використання результатів аналізу їх топологічної структури графічних об'єктів для встановлення подібності тексту.

19. Андрашко Ю.В., Білощицький А.О., **Кучанський О.Ю.**, Білощицька С.В., Лященко Т.О. Огляд методів оцінювання результатів діяльності науково-педагогічних працівників вищих навчальних закладів. Управління розвитком складних систем. 2017. № 29. С. 151 – 159.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належить аналіз параметрів, що використовуються при побудові національних та міжнародних рейтингів вищих навчальних закладів.

20. Білощицький А.О., **Кучанський О.Ю.**, Андрашко Ю.В., Білощицька С.В., Кузка О.І. Концептуальна модель інформаційної технології оцінювання результатів науково-дослідної діяльності. Управління розвитком складних систем. 2017. № 30. С. 163 – 168.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належить концептуальна модель інформаційної технології оцінювання продуктивності наукової діяльності науковців, закладів вищої освіти та їх структурних підрозділів.

21. Xu H., **Kuchansky A.** The problem of choice of partners for organization of cooperation in the framework of scientist of scientific and educational projects. Scientific Bulletin of Uzhhorod National University. Series of Mathematics and Informatics. 2019. 2(35), P. 134 – 142.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належать загальні вимоги для розробки інформаційної технології вибору потенційних партнерів виконавців наукових і освітніх проєктів, яка при впровадженні дозволить підвищити ефективність виконання проєктів для отримання стабільних інфраструктурних, освітніх та наукових результатів.

22. **Кучанський О.Ю.** Складові аналізу наукових мереж. Управління розвитком складних систем: збірник наукових праць. 2020. № 41. С. 115 – 126.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належать огляд задач, які виникають в межах аналізу наукових мереж: дослідження мереж наукової співпраці, дослідження мереж цитування, спільного цитування та бібліографічних з'єднань тощо.

23. **Кучанський О.Ю.** Задача ідентифікації предметних наукових просторів. Управління розвитком складних систем: збірник наукових праць. 2020. № 42. С. 70 – 78.

Видання індексовано в МНБД: Index Copernicus, BASE.

Автору належать опис етапів та визначено методи, що можуть бути застосовані для якісної ідентифікації предметних наукових просторів.

Публікації в закордонних виданнях:

24. Biloshchytskyi A., Biloshchytska S., **Kuchansky A.**, Andrashko Yu. Use of the link ranking method to evaluate scientific activities of scientific space subjects. Scientific Journal of Astana IT University. 2020. 1. P. 12 – 21.

Видання індексовано в МНБД: Google Scholar, Ulrichs Web, CNKI.

Автору належить метод, який базується на основі ранжування посилань та враховує самоцитування науковців для задачі оцінювання продуктивності суб'єктів наукової діяльності.

Монографії:

25. Білощицький А.О., Лізунов П.П., **Кучанський О.Ю.**, Андрашко Ю.В., Миронов О.В., Білощицька С.В. Методологічні основи створення інформаційного середовища управління науковими дослідженнями. Монографія. К.: КНУБА, 2017. 138 с.

26. Лізунов П.П., Білощицький А.О., **Кучанський О.Ю.**, Чала Л.Е., Діхтяренко О.В., Білощицька С.В. Інформаційна технологія пошуку неповних

дублікатів у графічних даних, математичних формулах та таблицях. Монографія. К.: КНУБА, 2017. 136 с.

27. Дзябенко О., Зінюк Л., Якубів В., **Кучанський О.Ю.** Концепція освіти для Serious Games індустрії: Комплексне дослідження в рамках проекту GameHub. Монографія. І.-Ф.: ПНУ ім. В. Стефаника, 2017. 98 с.

Матеріали міжнародних наукових конференцій:

28. Biloshchytskyi A., **Kuchansky A.**, Biloshchytska S., Dubnytska A. Conceptual Model of Automatic System of Near Duplicates Detection on Electronic Documents. Видання IEEE “The Experience of Designing and Applications of CAD Systems in Microelectron.” (CADSM), IEEE. 2017. P. 381 – 384.

Видання індексовано в МНБД: SCOPUS, Web of Science.

29. Biloshchytskyi A., **Kuchansky A.**, Andrashko Yu., Biloshchytska S., Dubnytska A., Vatskel V. The Method of the Scientific Directions Potential Forecasting in Infocommunication Systems of an Assessment of the Research Activity Result. Видання IEEE “Problems of Infocommunications. Science and Technology” (PICST), IEEE. 2017. P. 65 – 68.

Видання індексовано в МНБД: SCOPUS, Web of Science.

30. Biloshchytskyi A., **Kuchansky A.**, Biloshchytska S., Andrashko Yu., Bielova O. Infocommunication system of scientific activity management on the basis of project-vector methodology. Видання IEEE “Advanced Trends in Radioelectronics Telecommunications and Computer Eng.” (TCSET), IEEE. 2018. P. 27 – 30.

Видання індексовано в МНБД: SCOPUS, Web of Science.

31. Biloshchytskyi A., **Kuchansky A.**, Andrashko Yu., Bielova O. Learning space conceptual model for computing games developers. 2018 IEEE 13th International Scientific and Technical Conf. on Comp. Sciences and Information Technologies (CSIT). Lviv, 2018. Vol. 1. P. 97–102.

Видання індексовано в МНБД: SCOPUS, Web of Science.

32. **Kuchansky A.**, Biloshchytskyi A., Andrashko Yu., Vatskel V., Biloshchytska S., Danchenko O., Vatskel I. Combined models for forecasting the air pollution level in infocommunication systems for the environment state monitoring. 2018 IEEE 4th Intern. Symp. on Wireless Systems within the Int. Conf. On Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS). Lviv, 2018. P. 125–130.

Видання індексовано в МНБД: SCOPUS, Web of Science.

33. Biloshchytskyi A., **Kuchansky A.**, Andrashko Yu., Biloshchytska S., Danchenko O. Development of infocommunication system for scientific activity administration of educational environment’s subjects. 2018 IEEE Inter. Scient.-Pract. Conf. Problems of Infocommunications. Science and Technology (PIC S&T). Kharkiv, 2018. P. 369–373.

Видання індексовано в МНБД: SCOPUS, Web of Science.

34. Biloshchytskyi A., **Kuchansky A.**, Andrashko Yu., Biloshchytska S., Honcharenko T., Nikolenko V. Fractal Time Series Analysis in Non-Stationary Environment. 2019 IEEE Inter. Scient.-Pract. Conf. Problems of Infocommunications. Science and Technology (PIC S&T). Kyiv, 2019. P. 236–240.

Видання індексовано в МНБД: SCOPUS, Web of Science.

35. **Kuchansky A.**, Biloshchytskyi A., Bronin S., Biloshchytska S., Andrashko Y. Use of the fractal analysis of non-stationary time Series in mobile foreign exchange trading for m-learning. Internet of Things, Infrastructures and Mobile Applications (IMCL 2019). Advances in Intelligent Systems and Computing, vol 1192. 2021. Springer, P. 950 – 961.

Видання індексовано в МНБД: SCOPUS, Web of Science.

36. Bronin S., **Kuchansky A.**, Biloshchytskyi A., Zinyuk O., Kyselov V. Concept of Digital Competences in Service Training Systems. Internet of Things, Infrastructures and Mobile Applications (IMCL 2019). Advances in Intelligent Systems and Computing, vol 1192. 2021. Springer, P. 379 – 388.

Видання індексовано в МНБД: SCOPUS, Web of Science.

37. Biloshchytskyi A., **Kuchansky A.**, Andrashko Y., Mukhatayev A., Toxanov S., Faizullin A. Methods of Assessing the Scientific Activity of Scientists and Higher Education Institutions. 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT). Kyiv, 2021. P. 162 – 167.

Видання індексовано в МНБД: SCOPUS, Web of Science.

38. **Кучанський О.Ю.** Передпрогнозна індексація фінансових часових рядів. Матеріали міжнародної науково-практичної конференції «Обчислювальний інтелект (результати, проблеми, перспективи)», Черкаси. 2015. С. 93.

39. Білощицький А.О., **Кучанський О.Ю.**, Діхтяренко О.В., Лісневська І.Г. Про подібність фрагментів електронних документів. Матеріали II-ї Міжнародної науково-практичної конференції «Управління розвитком технологій», Київ, КНУБА. 2015. С. 26 – 27.

40. Білощицький А.О., Білощицька С.В., **Кучанський О.Ю.** Прогнозування методом співставлення зі зразком в управлінні інвестиційними проектами. Матеріали XII-ї міжнародної конференції «Управління проектами у розвитку суспільства», Київ, КНУБА. 2015. С. 148 – 149.

41. **Кучанський О.Ю.** Застосування методу зіставлення зі зразком при прогнозуванні часових рядів. Матеріали XI Міжнародної науково-практичної конференції «Управління проектами: стан та перспективи», м. Миколаїв, НУК ім. адм. Макарова. 2015. С. 181.

42. **Кучанський О.Ю.**, Мазурак В.В. Перспективи розвитку Інтернет-банкінгу в Україні. Матеріали XXI-ї Всеукраїнської наукової конференції “Сучасні проблеми прикладної математики та інформатики” СППМІ-2015, Львів, ЛНУ ім. І. Франка. 2015. С. 203 – 204.

43. **Кучанський О.Ю.** Прогнозування фінансових часових рядів методом зіставлення зі зразком. Матеріали II-ї Міжнародної науково-практичної конференції «Інформаційні технології та взаємодії». 2015. С. 98 – 100.

44. **Кучанський О.Ю.** Застосування індексації фінансових часових рядів для ідентифікації подібностей. Матеріали I-ї Всеукраїнської науково-практичної конференції «БудМайстерКлас». 2015. С. 191.

45. Білощицький А.О., **Кучанський О.Ю.**, Діхтяренко О.В., Мазурак В.В. Модулі попередньої обробки та індексації імпортованих документів. Матеріали III-ї Міжнародної науково-практичної конференції «Управління розвитком технологій», Київ. 2016. С. 10 – 12.

46. Лізунов П.П., Білощицький А.О., **Кучанський О.Ю.**, Діхтяренко О.В. Огляд методів аналізу та розпізнавання графічних даних для ідентифікації в них подібностей і дублікатів. Матеріали III-ї Міжнародної науково-практичної конференції «Управління розвитком технологій», Київ, 2016. С. 57 – 59.
47. **Кучанський О.Ю.**, Ніколенко В.В., Лященко Т.О. Використання наївних та трендових моделей для ідентифікації моментів зміни тенденцій часових рядів. Матеріали III-ї Міжнародної науково-практичної конференції «Управління розвитком технологій», Київ, 2016. С. 48 – 50.
48. Biloshchytskyi A., Biloshchytska S., **Kuchansky A.**, Vatskel' V. Sings of clustering documents in the system determining the fuzzy matches in the content of electronic documents. Матеріали XIII-ї Міжнародної конференції «Управління проектами у розвитку суспільства», Київ, 2016. Р. 18 – 20.
49. Білощицький А.О., **Кучанський О.Ю.**, Безмогоричний Д.М., Кузьомко А.С., Пида С.В. Концепція розбудови інфраструктури GameHub в українських університетах. Матеріали VIII-ї міжнародної школи-семінару “Теорія прийняття рішень”, Ужгород, ДВНЗ «УжНУ», 2016. С. 48 – 49.
50. Лізунов П.П., Білощицький А.О., **Кучанський О.Ю.**, Чала Л.Е. Знаходження неповних дублікатів математичних формул. Матеріали VIII-ї міжн. школи-семінару “ТПР”, Ужгород, ДВНЗ «УжНУ», 2016. С. 165.
51. Білощицький А.О., **Кучанський О.Ю.**, Чала Л.Е., Лізунов П.П. Гібридний метод знаходження неповних дублікатів математичних формул. Матеріали III-ї Міжнародної науково-практичної конференції «Інформаційні технології та взаємодії». 2016. С. 158 – 159.
52. Biloshchytskyi A., **Kuchansky A.**, Kuzomko A., Bezmogorychnyi D., Pyda S. Implementation plan of the GameHub in the part of integration process in training programs for computer game design. Матеріали III-ї Міжнародної науково-практичної конференції «Інформаційні технології та взаємодії». 2016. Р. 161 – 162.
53. Lizunov P., Biloshchytskyi A., **Kuchansky A.**, Dubnytska A. Methods of analysis of the image, mathematical formulas and tables for finding incomplete duplicates. Матеріали науково-практичної конференції «БудМайстерКлас». 2016. Р. 111 – 112.
54. Biloshchytskyi A., **Kuchansky A.**, Dubnytska A., Pyda S. Conceptual bases in gamehub infrastructure development in Kyiv National University of Construction and Architecture. Матеріали науково-практичної конференції «БудМайстерКлас». 2016. Р. 59 – 60.
55. **Kuchansky A.**, Chyzhmotrya O. Multiple transport task with two criteria. Матеріали науково-практичної конференції «БудМайстерКлас». 2016. Р. 66 – 67.
56. Білощицький А.О., **Кучанський О.Ю.**, Чала Л.Е. Гібридний метод визначення неповних дублікатів у таблицях. Матеріали I-ї Міжнародної науково-практичної конференції «Project, Program, Portfolio, р3management». Одеса, 2016. С. 77 – 80.
57. Biloshchytskyi A., **Kuchansky A.**, Chala L. Detection of Near-Duplicates in Tables. Матеріали Міжнародної науково-практичної конференції «Актуальні питання забезпечення кібербезпеки та захисту інформації». Воловець, 2017. Р. 193 – 194.
58. Білощицький А.О., **Кучанський О.Ю.** Методи виявлення плагіату та незаконного поширення інформації в кіберпросторі. Матеріали II-ї науково-

практичної конференції «Проблеми кібербезпеки інформаційно-комунікаційних систем». Київ, 2017. С. 129 – 132.

59. Удовенко С.Г., **Кучанський О.Ю.** Метод пошука дублікатів зображень в форматі PDF. Матеріали Міжнародної науково-практичної конф. «ІТ в металургії та машинобудуванні». Дніпро, 2017. С. 162.

60. Білощицький А.О., **Кучанський О.Ю.** Виявлення неповних дублікатів у рукописах дисертаційних досліджень. Матеріали IV-ої науково-практичної конференції «Управління розвитком технологій». Київ, 2017. С. 36.

61. Білощицький А.О., **Кучанський О.Ю.**, Андрашко Ю.В. Методи оцінювання результатів наукової діяльності науково-педагогічних працівників. Матеріали науково-практичної конференції «Обчислювальний інтелект». 2017. С. 65 – 66.

62. Білощицький А.О., **Кучанський О.Ю.**, Андрашко Ю.В. Мережеві технології оцінки результатів науково-дослідної діяльності структурних підрозділів ВНЗ МОН України. Матеріали науково-технічної конференції «Актуальні проблеми інформаційних технологій». Київ, 2017. С. 39 – 40.

63. Білощицький А.О., **Кучанський О.Ю.**, Андрашко Ю.В. Методи оцінки результатів науково-дослідної діяльності структурних підрозділів ВНЗ. Матеріали IV-ої Міжнародної науково-практичної конференції «Інформаційні технології та взаємодії». Київ, 2017. С. 25 – 26.

64. Білощицький А.О., **Кучанський О.Ю.**, Андрашко Ю.В. Методи визначення відстаней в задачі кластеризації публікацій науковців за науковими напрямками. Матеріали IV-ої Міжнародної науково-практичної конференції «Інформаційні технології та взаємодії». Київ, 2017. С. 150 – 151.

65. Білощицький А.О., Білощицька С.В., **Кучанський О.Ю.**, Вацкель В.Ю., Вацкель В.І. Проблеми реалізації концепції Інтернет речей «Розумний будинок» на прикладі IT Lynx Smarthouse. Перша науково-практична конференція «Інтернет речей: проблеми правового регулювання та впровадження». Київ, 2017. С. 55 – 58.

66. Білощицький А.О., **Кучанський О.Ю.**, Андрашко Ю.В. Застосування методології проектно-векторного управління науково-дослідною діяльністю. Інформаційні технології та взаємодії (IT&I–2018). Матеріали доповідей. Київ, 2018. С. 74–75.

67. Biloshchytskyi A., **Kuchansky A.**, Bronin S. Digital competencies in Ukrainian education. V Міжнародна науково-практична конференція «Інформаційні технології в культурі, мистецтві, освіті, науці, економіці та бізнесі», Київ, 2020. P. 341 – 343.

68. **Кучанський О.Ю.** Елементи аналізу мереж наукової співпраці. VII Міжнародна науково-практична конференція "Управління розвитком технологій", Київ, 2020, С. 67 – 68.

Наукові публікації з достатньою повнотою відображають винесені на захист наукові положення, моделі і розробки, одержані практичні результати та висновки.

АНОТАЦІЯ

Кучанський Олександр Юрійович. Методологія формування інформаційних просторів суб'єктів наукової діяльності у сталому розвитку закладів вищої освіти.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 «Інформаційні технології». – Київський національний університет будівництва і архітектури, Київ, 2021.

В дисертаційній роботі наведені результати узагальнення й розв'язку важливої проблеми: створення методології формування інформаційних просторів суб'єктів наукової діяльності, яка є важливою складовою забезпечення сталого розвитку закладів вищої освіти і включає моделі та методи прозорого кількісного оцінювання продуктивності суб'єктів наукової діяльності у відповідних предметних наукових просторах. Застосування методології дозволяє стимулювати мотивацію суб'єктів наукової діяльності шляхом забезпечення раціонального фінансування найбільш продуктивних з них, що важливо в умовах обмеженого ресурсного забезпечення.

Проведено аналіз стану завдань формування інформаційних просторів та кількісного оцінювання продуктивності суб'єктів наукової діяльності, формування предметних наукових просторів, пошук партнерів для співпраці в рамках освітніх та наукових проєктів. Розроблено множинні моделі ідентифікації індивідуальних та колективних суб'єктів наукової діяльності, а також формалізовану множинну модель інформаційного простору суб'єктів наукової діяльності та концептуальну схему їх взаємодії. Дані моделі, а також моделі та методи ідентифікації предметних наукових просторів, оцінювання продуктивності суб'єктів наукової діяльності та математична модель і метод багатокритеріального вибору потенційних партнерів для співпраці лягли в основу створення методології формування інформаційних просторів суб'єктів наукової діяльності.

Набули подальшого розвитку методи формування предметних наукових просторів на основі кластерного аналізу параметрів моделей ідентифікації суб'єктів наукової діяльності та на основі латентно-семантичного аналізу. Вдосконалено метод кількісного оцінювання продуктивності індивідуального та колективного суб'єктів наукової діяльності з врахуванням самоцитування, а також метод оцінювання перспекції їх продуктивності на основі ретроспективних даних. Вдосконалено математичну модель та метод багатокритеріального вибору потенційних партнерів для співпраці в рамках наукових та освітніх проєктів, що враховує організаційну будову проєктів та результати кількісного оцінювання продуктивності суб'єктів наукової діяльності відповідного предметного наукового простору.

Набув подальшого розвитку векторний підхід до формування просторів суб'єктів наукової діяльності методології проєктно-векторного управління в освітніх середовищах, що включає нові та вдосконалені методи і моделі, які дозволяють оцінювати продуктивність наукової діяльності, ідентифікувати предметні наукові простори, формувати робочі групи наукових проєктів та склад структурних підрозділів закладів вищої освіти.

Розроблено інформаційну технологію, яка реалізує методологію формування інформаційних просторів суб'єктів наукової діяльності з використанням технології збору інформації з відкритих джерел, розрахунку та формування рейтингів суб'єктів наукової діяльності у відповідних предметних наукових просторах тощо.

Отримані результати як теоретичного, так і практичного характеру, можуть бути підґрунтям для подальших досліджень, спрямованих на забезпечення сталості, удосконалення організації та управління науковою діяльністю, що визначається нормативною базою в цій галузі, керується особливостями ресурсного забезпечення та відповідає вимогам міжнародних систем оцінювання продуктивності наукової діяльності суб'єктів інформаційного освітньо-наукового простору.

Ключові слова: інформаційні технології, інформаційний простір, суб'єкт наукової діяльності, оцінювання продуктивності, заклад вищої освіти.

ABSTRACT

Kuchanskyi Oleksandr. The methodology for forming information spaces of scientific subjects for sustainable development in higher education institutions. – Qualified scientific work on the rights of the manuscript.

Thesis for a Doctor of Science Degree in specialty 05.13.06 "Information Technologies". – Kyiv National University of Construction and Architecture, Kyiv, 2021.

The dissertation presents the results of generalization and solution of a significant problem: the creation of a methodology for forming information spaces of scientific subjects, which is an essential component of sustainable development of higher education institutions and includes models and methods of a transparent quantitative assessment of the productivity of scientific subjects. The methodology's application allows stimulating the motivation of the subjects of scientific activity by providing rational financing of the most productive of them, which is essential in the conditions of limited resource provision.

The analysis of the state of tasks of formation of information spaces and quantitative estimation of productivity of scientific activity subjects, the formation of scientific subject spaces, search of partners for cooperation within the limits of educational and scientific projects are carried out. Multiple models of identification of individual and collective subjects of scientific activity are developed. They formalized multiple models of information space of scientific activity subjects and a conceptual scheme of their interaction. These models and models and methods of identifying scientific subject spaces, evaluating the productivity of scientific subjects and mathematical model, and method of multi-criteria selection of potential partners for cooperation formed the basis for creating a methodology for forming information spaces of scientific subjects.

Methods of forming scientific subject spaces based on cluster analysis of parameters of models of identification of scientific activity subjects and based on latent-semantic analysis have been further developed. The method of a quantitative assessment of the productivity of individual and collective subjects of scientific activity, taking into account self-citation and a method of estimation of prospecting of their productivity based on retrospective data is improved. The mathematical model and method of multi-criteria selection of potential partners for cooperation in research and educational projects have been improved, which considers the organizational structure of projects and the quantitative evaluation results of the productivity of scientific subjects of the relevant subject space.

The vector approach to the formation of spaces of scientific activity subjects of the methodology of project-vector management in educational environments has been further developed. It includes new and improved methods and models that allow to assess scientific activity productivity, identify subject scientific spaces, form working groups of scientific projects, and the composition of structural units of higher education institutions.

Information technology has been developed, which implements the methodology of forming information spaces of scientific subjects using the technology of collecting

information from open sources, calculating and forming ratings of scientific subjects in the relevant subject scientific spaces, etc.

The obtained results, both theoretical and practical, can be the basis for further research to ensure sustainability, improve the organization, and manage scientific activities. It is determined by the regulatory framework in this area, guided by the specifics of resource provision and meets the requirements of international systems for assessing the productivity of scientific activities.

Keywords: information technologies, information space, subject of scientific activity, productivity assessment, higher education institution.