

## E-DITORIAL – ВІРТУАЛІЗАЦІЯ РОБОТИ РЕДАКЦІЙНОЇ КОЛЕГІЇ ФАХОВОГО ПЕРІОДИЧНОГО ВИДАННЯ

*Р.Р. Даревич<sup>1</sup>, Д.Г. Досин<sup>1</sup>, В.В. Литвин<sup>1,2</sup>, О.І. Мриглод<sup>2,3</sup>, Н.В. Шкутяк<sup>1</sup>*

<sup>1</sup> Фізико-механічний інститут ім. Г.В. Карпенка НАН України,  
79601, Львів, вул. Наукова, 5, тел. 263 3088, rpinasu@ipm.lviv.ua;

<sup>2</sup> Національний університет «Львівська політехніка»,

<sup>3</sup> Інститут фізики конденсованих систем НАН України

Розглядається архітектура віртуальної самокерованої редакційної колегії фахового періодичного видання як системи штучного інтелекту, яка містить методи і засоби автоматичної підготовки до електронного друку поданих статей. Запропоновано для знаходження найближчого за тематикою рецензента використовувати метрику між концептуальними графами – моделями поданих до друку статей. Для цього вводиться поняття центру семантичної ваги концептуального графа, а відстань від сфери компетентності рецензента до статті визначається як відстань між їх центрами семантичної ваги відповідних концептуальних графів.

This article considers architecture of virtual self-controlled journal editorial board as AI system with means and methods of automated preparation of scientific articles for electronic publication. It is proposed to use metrics between conceptual graphs – models of submitted articles to estimate nearest scientist, which field of scientific interests allow to ask him for making article review. The idea of “article center of semantic weight” is used for solving this task.

### Вступ

Наукознавство або «наука про науку» [1] нині стало сферою зосередження особливої уваги, адже розвиток інформаційних технологій надає нові можливості для того, щоб по-новому підійти до вирішення проблем у науковій сфері. Проблема аналізу наукової продуктивності та обліку праці наукового працівника є однією з найдавніших, проте й нині найактуальніших проблем наукознавства. Досі одним із найуживаніших підходів до вирішення згаданої проблеми є аналіз сукупності опублікованих наукових праць як основної форми вираження та фіксації наукових результатів. Саме цим визначається актуальність досліджень із оцінювання та рейтингування окремих наукових статей та наукових видань загалом.

Поруч із задачами обліку та оцінювання вже опублікованих наукових видань, не менш важливою задачею є сам процес їх підготовки. Адже саме на проміжку між поступленням рукописів у редакцію та опублікуванням прийнятих статей відбувається важливий процес відбору якісних матеріалів, їх вичитування та рецензування, а отже – забезпечення якісного рівня поширюваної наукової інформації. В основі процесу редакційного опрацювання рукописів лежить, в основному, робота експертів, у ролі яких виступають члени редакційної колегії. Серед основних функцій редколегії є призначення рецензентів, прийняття рішень про відповідність надісланих матеріалів профілю видання, тематична класифікація рукописів, тощо. Такого роду завдання можуть виконувати фахові спеціалісти, тобто досвідчені науковці, що є активними членами наукової спільноти та добре ознайомлені з множиною публікацій у тій чи іншій тематичній області. Такий спосіб організації роботи з підготовки наукових періодичних видань є класичним та перевіреним часом, проте він має ряд недоліків. З одного боку, участь науковців у роботі редакційної колегії з часом стає рутинним виконанням ряду функцій та забирає час, необхідний для творчої роботи. З іншого – при роботі експертів у будь-якій галузі завжди існує проблема фактора суб'єктивності, тобто впливу емоцій, настрою чи особистих упереджень на прийняття рішень.

Нині існують достатні інформаційні потужності для того, щоб ставити задачу інформатизації та автоматизації процесів підготовки наукових періодичних видань, зокрема – задачу віртуалізації роботи редакційної колегії за допомогою програмних засобів із застосуванням мережових технологій та досягнень в області штучного інтелекту. При цьому розроблення подібної системи має враховувати акумульований досвід наукознавства, існуючі стандарти в галузі наукової періодики, міжнародні вимоги. Головна мета інформаційної системи для автоматизації редакційного опрацювання наукових статей залишається співзвучною з головним призначенням будь-якого наукового періодичного видання – поширювати нову та актуальну наукову інформацію в тій чи іншій області.

### Постановка проблеми у загальному вигляді

Для того, щоб краще зрозуміти множину функцій, які виконуються на рівні редакційної колегії, необхідно розглянути систему керування науковим періодичним виданням та ієрархію цілей, що покладені в її основу. З точки зору керування наукове періодичне видання є системою, яку можна розглядати на різних рівнях [2]: стратегічне планування здійснюється на рівні видавця, тактичні завдання вирішуються органом редакційної колегії, а оперативні технічні задачі виконуються працівниками редакції (рис. 1) [3]. Щоб наукове періодичне видання могло досягнути конкурентоспроможного якісного рівня та ефективно функціонувати в перспективі,

потрібно поступово реалізовувати ті чи інші стратегічні проекти. Наприклад, стратегічним завданням може бути включення до міжнародного списку авторитетних наукових видань Thomson Scientific, що визначає низку тактичних задач: забезпечення чіткої періодичності виходу, організацію та підтримку веб-сторінки журналу, наявність певних елементів статей англійською мовою тощо. Кожна тактична задача передбачає ряд оперативних дій на технічному рівні, наприклад, щоб забезпечити вчасність виходу журналу, необхідно контролювати в часі процес прийняття статей до друку, їх рецензування та технічну підготовку, забезпечувати вчасну наповненість «портфеля» тощо.

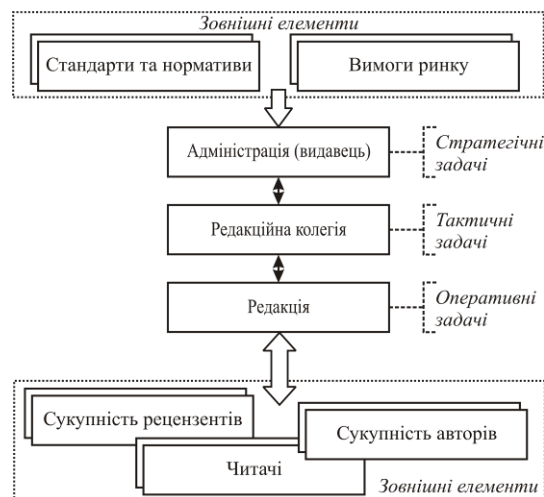


Рис. 1. Ієрархія компонентів адміністративної системи наукового видання з точки зору задач керування

Безумовно, найважливішою функцією редакційної колегії наукового періодичного видання є забезпечення видання матеріалами, що містять нову та актуальну наукову інформацію, тобто відбір статей, що найкраще відображають фахову спрямованість видання та відповідають його профілю.

Процес редакційного опрацювання надісланих рукописів є багатоетапним. На першому етапі відбувається відсіювання матеріалів, що не відповідають поставленим технічним вимогам, а тому не можуть далі опрацьовуватись редакцією. Це є невідповідність формату, мови, відсутність важливих елементів статті, технічні проблеми при роботі з електронними версіями тощо. Наступний етап передбачає розгляд рукопису редакційною колегією або її окремими членами на предмет відповідності профілю та спрямованості даного періодичного видання (наприклад, відхилення експериментальних робіт поданих до видання, що публікує теоретичні результати). Подальша робота з надісланими матеріалами пов'язана безпосередньо з організацією процесу рецензування, що є, безумовно, найбільш складним та відповідальним етапом. Процес визначення рецензентів враховує цілий ряд факторів, виходячи як із особливостей надісланого матеріалу, так і з інтересів самого видання:

- 1) тематика надісланої статті (при цьому необхідно визначити якомога вужчу область, до якої належить дана робота та визначити рецензентів, що працюють найближче до цієї області);
- 2) належність авторів та потенційних рецензентів до певної наукової школи або робочої групи (буває, що різні групи авторів, працюючи в одній і тій же науковій тематиці, стоять на конкуруючих позиціях);
- 3) соціальні аспекти (часом такі фактори, як, наприклад, національність авторів та рецензентів мають важливе значення);
- 4) попередній досвід рецензування (аналіз історії рецензування часто дозволяє передбачити перебіг майбутньої роботи: швидкість виконання, детальність відгуку та інші особливості);
- 5) сукупність процитованих у рукописі авторів (позитивним досвідом є призначення рецензентів із авторів робіт, що цитуються);
- 6) міцність взаємозв'язків редакції з потенційними рецензентами (наявність контактної інформації, можливість контролю над часовими рамками процесу рецензування, тощо).

Наступні кроки процесу «проходження» статей у редакції наукового періодичного видання є вже технічнішими – це комунікація з авторами та рецензентами, надсилання відгуків і перероблених версій та безпосередньо технічна підготовка статей до друку.

Таким чином, при розробці архітектури інтелектуальної інформаційної системи, що реалізує функції віртуальної редакційної колегії, основними її компонентами можна визначити такі:

- підсистема попереднього аналізу («технічне відсіювання»);
- підсистема співставлення профілю видання (визначення типу статті та відповідності тематиці видання);

- підсистема організації процесу рецензування (визначення вузької тематики статті, виділення кола потенційних рецензентів та їх призначення, технічний супровід процесу рецензування);
- підсистема технічної підготовки (додрукарські процеси для друкованого видання або ж підготовка електронних публікацій для розміщення в інформаційній мережі).

З огляду на те, що перша та остання з наведених компонент вимагають технічної реалізації відомими методами, на них не будемо зупинятись. Друга та третя підсистеми вже потребують інтеграції певних функцій інтелектуального аналізу, тому будуть розглянуті більш детально у наступних підрозділах.

### Огляд підходів до проблеми формалізації аналізу змісту текстових документів

Формалізація аналізу та представлення змісту текстового документа залишається актуальною науково-технічною проблемою, пов'язаною, з одного боку, зі створенням (вибором) адекватних проблемній області та природній мові документів формально-логічних виразних засобів [4], а з другого – розробленням методів та процедур розпізнавання (зокрема, класифікації та кластеризації) змісту текстових документів, що аналізуються та опрацьовуються з певною метою. У роботі [5] представлено огляд стану досліджень у цій галузі, розглядаються проблеми організації просторів знань з використанням Інтернет-технологій, методи та засоби видобування знань з текстів на природних мовах, а також питання використання просторів знань при створенні прикладних інтелектуальних систем, що функціонують в мережі Інтернет. Стверджується, що стандартом de facto, запровадженим W3C на специфікацію онтологій, стала мова OWL. Натомість мова SPARQL поступово стає стандартом de facto, як мова запитів до баз знань.

В області моделей і методів аналізу природомовних текстів на даний час отримані теоретично й практично значущі морфологічні моделі аналізу/синтезу лексем; розроблені моделі синтаксичного аналізу основних ПМ-конструкцій; запропоновано ряд методів реалізації основних моделей аналізу ПМ-конструкцій, які можуть використовуватися на практиці; виділені основні прийоми евристичної реалізації окремих моделей інтерпретації ПМ-висловлювань; опрацьовані окремі моделі концептуального синтезу ПМ-текстів; запропоновані і практично перевірені моделі та методи лінгвістичного синтезу. В області моделей розуміння розроблені багаторівневі моделі, що враховують не тільки лінгвістичні, а й когнітивні складові цього процесу. В області реалізації розроблені прототипи інтелектуальних ПМ-систем; є промислові реалізації ПМ-систем різного класу (див. таблицю), які в більшості випадків лише «імітують» повномасштабне розуміння природної мови.

Таблиця. Розробники практично реалізованих ПМ-засобів

1	BASIS Tech	<a href="http://www.basistech.com">http://www.basistech.com</a>
2	ClearForest Corporation	<a href="http://www.clearforest.com">http://www.clearforest.com</a>
3	CognIT	<a href="http://www.cognit.no">http://www.cognit.no</a>
4	Compris Intelligence GmbH	<a href="http://www.kompass.com">http://www.kompass.com</a>
5	Convera (formerly Excalibur)	<a href="http://www.convera.com">http://www.convera.com</a>
6	Delphes	<a href="http://www.delphes.com">http://www.delphes.com</a>
7	Megaputer Intelligence Inc.	<a href="http://www.megaputer.com">http://www.megaputer.com</a>
8	Insightful Corporation & InFact	<a href="http://www.insightful.com">http://www.insightful.com</a>
9	Inxight Software Inc.	<a href="http://www.inxight.com/">http://www.inxight.com/</a>
10	MITRE	<a href="http://www.mitre.org">http://www.mitre.org</a>
11	Ontotext	<a href="http://www.ontotext.com">http://www.ontotext.com</a>
12	SRA International Inc.	<a href="http://www.sra.com">http://www.sra.com</a>
13	TEMIS	<a href="http://www.temis.com">http://www.temis.com</a>
14	Teragram Corporation	<a href="http://www.teragram.com">http://www.teragram.com</a>

На думку авторів, комп'ютерна лінгвістика, штучний інтелект і, тим більше, інформаційні технології досі не мають ефективних моделей обробки текстів на природних мовах, а завдання автоматичного опрацювання довільних ПМ текстів далі не має рішення, яке могло б мати якесь практичне значення. Ряд діючих проєктів застосовують евристичні алгоритми статистичного опрацювання текстів, обмежені онтології, зорієнтовані на вузьку конкретну наперед задану проблемну область. Наприклад, у конфігурації NameTag продукту **NetOwl** фірми **SRA International Inc.** обробляються 7 основних і близько 70 підтипів важливих сутностей, зокрема, *people, organizations, places, artifacts, phone, social security numbers, dollar amounts, dates, addresses*. У конфігурації Link and Event виділяються більше 150 типів зв'язків і подій на основі вбудованої онтології.

### Модель текстового документа та введення на її основі метрики

Пропонується підхід до побудови моделі текстового документа (наукової статті), що базується на спеціальних процедурах його формалізації. Для цього використано результати досліджень природної мови дистрибутивним методом та концептуальні граfi (КГ). Якщо на отриманий таким чином КГ накласти ваги понять, взяті з адаптивної онтології ПО, то отримаємо зважений концептуальний граф (ЗКГ), який використовуватимемо для подальшого опрацювання.

Для знаходження різного роду необхідних семантичних відстаней, як віддаленості змісту окремих понять чи тематики повідомлень (текстових документів), знайдемо центр ваг ЗКГ, який є моделлю деякого текстового документа. Для цього виконаємо такі кроки:

- 1) порівнювані тексти подаємо у вигляді їх концептуальних графів;
- 2) графи доповнюємо відповідним контекстом та коефіцієнтами важливості з адаптивної онтології. Детально про процедури адаптації онтології описано в [6];
- 3) відстань між двома вершинами графа  $C_i$  та  $C_j$ , якщо ці вершини з'єднані дугою, визначаємо як:

$$d_{ij} = \frac{Q}{L_{ij}(W_i + W_j)}, \quad (1)$$

де  $W_i$  та  $W_j$  – коефіцієнти важливості вершин  $C_i$  та  $C_j$  відповідно;  $L_{ij}$  – коефіцієнт важливості зв'язку між вершинами;  $Q$  – константа, яка залежить від конкретної онтології. Приймемо, що  $L_{ii} = \infty$ , тоді  $d_{ii} = 0$ .

- 4) Знаходимо центр ваг концептуального графа. Це вершина  $C_{i^*}$ , для якої середня відстань  $\bar{d}_i$  найменша:

$$\bar{d}_{i^*} = \min_i \bar{d}_i. \quad (2)$$

Середня відстань  $\bar{d}$  для вершини  $C_i$  обчислюється згідно формули:

$$\bar{d}_i = \frac{\sum_{j=1, j \neq i}^n d_{ij}^*}{n-1}, \quad (3)$$

де  $n$  – кількість вершин графа,  $d_{ij}^*$  – найкоротший шлях між вершинами  $C_i$  та  $C_j$ , який обчислюється за допомогою відомих алгоритмів, наприклад, Форда, Дейкстри, Флойда-Уоршалла [6]. Визначена таким чином відстань задовольняє аксіомам метрики [7].

### Визначення відповідності статті профілю наукового періодичного видання

Фахова спрямованість видання задається адаптивною онтологією предметної області (ПО) для яких спрямоване це видання. Так деякі видання є фаховими в галузі комп'ютерних наук, деякі в хімічних і т.д. Адаптивну онтологію визначимо як п'ятірку [7]:

$$O = \langle X, R, F, W, L \rangle,$$

де  $X$  – скінченна множина концептів (понять, термінів) предметної області, яку задає онтологія  $O$ ;  $R$  – скінченна множина відношення між концептами (поняттями, термінами) заданої предметної області;  $F$  – скінченна множина функцій інтерпретації (аксіоматизація), заданих на концептах або відношеннях онтології  $O$ , де  $W$  – важливість понять  $X$ ,  $L$  – важливість відношень  $R$ .

За отриманою відстанню визначається найближча тематична група, якій відповідає дана стаття. Блок-схема алгоритма показана на рис. 2.

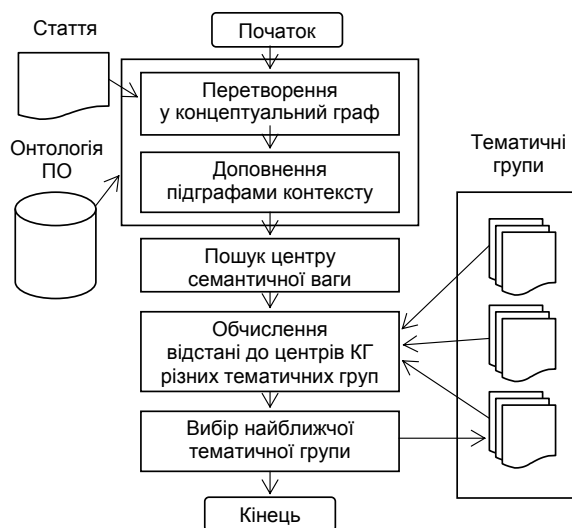


Рис. 2. Блок-схема алгоритму пошуку найближчої за змістом тематичної групи текстових документів з використанням онтології

### Виділення кола потенційних рецензентів

База потенційних рецензентів наукового періодичного видання, в основному, складається із сукупності її авторів. На відміну від реальної редакційної колегії, що має можливість залучати нових рецензентів на основі особистого досвіду експертів, віртуальна редакційна колегія володіє лише тими даними, що є зафіксованими у її базі. З іншого боку, згадана база даних може поповнюватись новими прізвищами шляхом пошуку в глобальній мережі Інтернет.

Перед тим як використати методики більш точного оцінювання спорідненості тематики надісланої статті та потенційних рецензентів, необхідно визначити скінченний перелік прізвищ учених, що працюють в околі даної предметної області. Для більш загального групування авторів за їх тематикою, можна застосувати різні підходи, зокрема на основі інформації про наукове співавторство. Використання саме даних про співавторство виправдане тим, що їх легко можна одержати з мережі Інтернет. З іншого боку, відомості про співавторство у конкретному науковому періодичному виданні переважно вже у готовому вигляді містяться в базі даних цього видання.

Нехай  $St = \{St_1, St_2, \dots, St_n\}$  – множина статей направлених для опублікування в фаховому виданні. Кожна стаття представляється у вигляді моделі зваженого концептуального графа (ЗКГ). У свою чергу кожний ЗКГ містить центр семантичної ваги (об'єкт досліджень статті). Множину таких об'єктів позначимо  $Ob = \{Ob_1, Ob_2, \dots, Ob_m\}$ . Нехай  $Z = \{z_1, z_2, \dots, z_l\}$ , яких використовує фахове видання для рецензування статей. Очевидно, що результат рецензування тим кращий, чим компетентніший рецензент у відповідній проблематиці досліджень. Проблематику досліджень фахового видання задамо у вигляді онтології. Очевидно, що така онтологія має включати об'єкти досліджень. А моделлю онтології є свій ЗКГ. Компетентність рецензента щодо об'єкта досліджень задамо у вигляді нечіткого відношення  $F(Z, Ob) = \{(z_i, Ob_j | \lambda_{ij})\}$ . Чим більше значення  $\lambda_{ij}$ , тим компетентніший рецензент. Таке нечітке відношення може бути побудоване на основі аналізу наукових праць рецензентів і / або шляхом їх опитування.

Якщо об'єктом досліджень статті  $St_i \in Ob_k$ , то рецензувати цю статтю має рецензент  $z_j$ , де  $j = \arg \max_k \lambda_{jk}$ .

Зручні засоби та методи для аналізу даних про авторів видання надає теорія складних мереж [8, 9], що дає можливість представити великий масив взаємопов'язаних даних у вигляді складної математичної мережі – набору вузлів, пов'язаних зв'язками. Різноманітні алгоритми на мережах дають можливість згрупувати дані за ступенем їх близькості або ж навпаки – виявити, наскільки вони віддалені. Математичні мережі є також чудовим засобом візуалізації даних у вигляді, що є зручним для сприйняття людиною. Зобразивши ті чи інші дані у вигляді складної мережі, можна розраховувати ряд типових величин, що її характеризують, а саме: розподіл ступенів вузлів, найдовший та середній шлях із множини найкоротших шляхів між двома довільними вузлами, посередництво, коефіцієнт кластерності мережі та інші [8, 9].

На рис. 3 показано фрагмент мережі співавторства вибраного українського наукового періодичного видання. Вузли позначають його авторів, а зв'язок між двома вузлами виникає тоді, коли пара авторів має спільну опубліковану роботу в даному виданні.

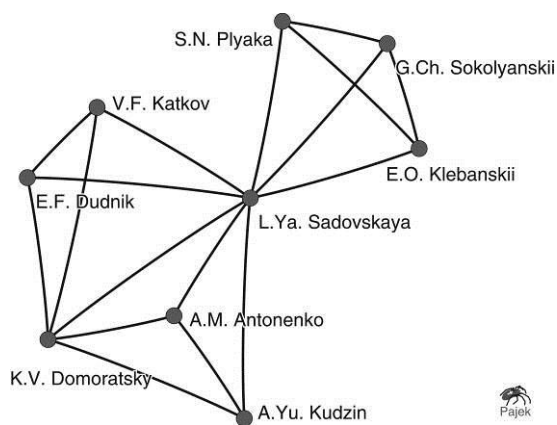


Рис. 3. Фрагмент мережі співавторства українського наукового журналу з теоретичної фізики «Condensed Matter Physics»<sup>1</sup>

Представлення даних у вигляді складної мережі дає можливість використовувати ряд алгоритмів на мережах, зокрема, алгоритм Гірвана–Ньюмена [10] для знаходження «природної структури мережі» (без попереднього задання кількості кластерів). При цьому мається на увазі виявлення зв'язаних груп вузлів, що мають більшу концентрацію зв'язків між собою, проте не представлені ізольованими кластерами (для знаходження ізольованих компонент мережі співавторства можна використовувати будь-які з відомих методів обходу графів). У випадку аналізу даних про публікації в науковому виданні знаходження структури мережі статей чи авторів, поєднаних певними зв'язками, забезпечує групування їх за спільною науковою тематикою.

<sup>1</sup> Візуалізація цієї та частини інших мереж здійснена за допомогою програмного пакету Pajek (Vlado A. Pajek: Program for large network analysis; Borgatti, S.P. 2002.)

Узагальнений алгоритм методу знаходження тематичних груп авторів видання на основі мережі співавторства передбачає виконання таких кроків:

- 1) побудова мережі співавторства;
- 2) знаходження зв'язаних компонент побудованої мережі;
- 3) аналіз кожної із знайдених компонент та виявлення їх структури методом Гірвана–Ньюмена (в рамках алгоритма автоматично відбувається перевірка того, чи мережа має виражену структуру);
- 4) виявлення домінуючих тематик для кожної із знайдених груп авторів на основі аналізу тематичних індексів.

Описаний метод був випробуваний на основі реальних даних про співавторство у виданні «Condensed Matter Physics». На основі даних про публікації в журналі, за 10 років, було побудовано мережу співавторства (838 вузлів) та виділено її зв'язані компоненти. 116 вузлів мережі увійшли в найбільшу зв'язану компоненту, тоді як наступні за розміром містять, відповідно, 74 та 15 вузлів. На рис. 4 показано результат застосування методу виділення тематичних груп авторів у найбільшій зв'язаній компоненті, що вже має достатньо виражену структуру. Для визначення домінуючих тематик для знайдених груп авторів використовувались дані про вказані в публікаціях тематичні номери PACS (запропонована Американським інститутом фізики класифікаційна схема для окреслення тематики наукових публікацій з фізики та астрономії, акронім від «Physics and Astronomy Classification Scheme»), які, подібно до УДК, мають ієрархічну структуру. Наприклад, домінуючою тематикою однієї із визначених груп авторів (група із 6 авторів, розміщена справа вгорі на рис. 4: V.G. Morozov, G. Ropke, N.V. Davydov, V.A. Davydov, M.N. Stolyarov та T.Yamaguchi) є статистична фізика, термодинаміка та нелінійні динамічні системи.

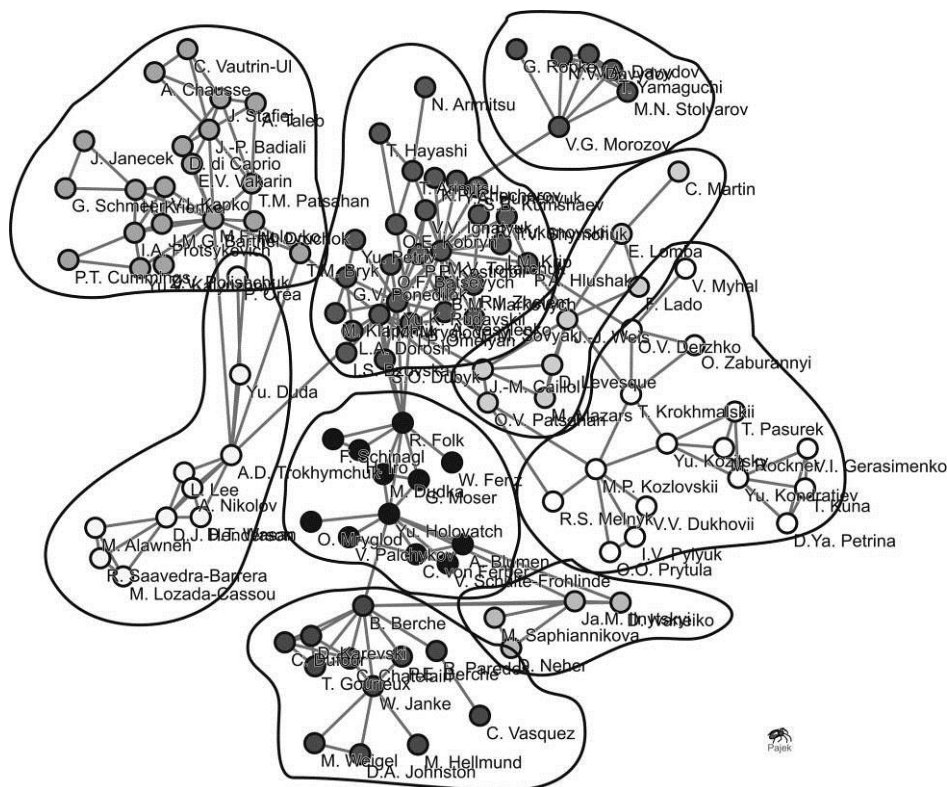


Рис. 4. Найбільша зв'язана компонента мережі співавторства видання «Condensed Matter Physics» із виділеною її природною структурою.

### Архітектура інтелектуальної системи E-ditorial

Розроблені методи покладені в основу пакета E-ditorial – інтелектуальної системи аналізу наукових праць з метою включення їх до електронного збірника за певним профілем наукових досліджень, як однієї із служб віртуального автоматизованого робочого місця наукового працівника, що розробляється у Фізико-механічному інституті НАН України.

Взаємозв'язок основних елементів системи E-ditorial здійснюється через реляційну базу даних, в якій зберігається вся інформація про опубліковані, відхилені та подані до друку статті, дані про авторів, їх публікації, спів- та взаємоцитуння, про рецензентів (рис. 5).

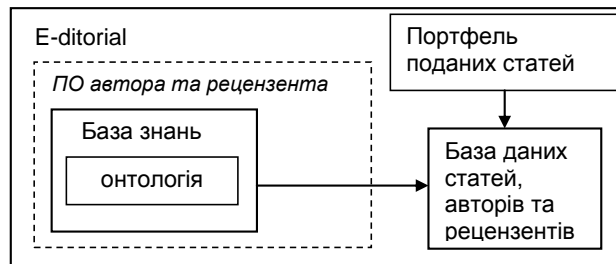


Рис. 5. Взаємозв'язок основних елементів системи E-ditorial через реляційну базу даних статей, авторів та рецензентів

Служба автоматично приймає до «портфелю» подану до публікації наукову статтю, здійснює первинний аналіз її відповідності вимогам до форми представлення, будує концептуальний граф анотації цієї статті у форматі RDF-OWL як онтологію її автора, доповнює підграфами раніше побудованої онтології автора, якщо він раніше публікувався в цьому виданні, доповнює підграфами контексту з онтології проблемної області видання, шукає центр семантичної ваги понять концептуального графа, визначає близькість тематики статті до видання в цілому та кожного з його розділів зокрема, як семантичну відстань між центрами відповідних суміщених ЗКГ, таким чином приймає рішення про віднесення до певного розділу, виділяє з мережі рецензентів тематичної групи вибраного розділу найбільш відповідних кандидатів за раніше встановленими критеріями, які забезпечують компетентність та об'єктивність рецензування та надсилає обраному рецензенту подану до друку статтю з анкетною її оцінювання. За результатами оцінювання рецензентом приймає рішення про повторне рецензування іншим рецензентом, відправлення автору на доопрацювання, включення до найближчого випуску, залишення в резерві публікації, чи відхилення статті (рис. 6).

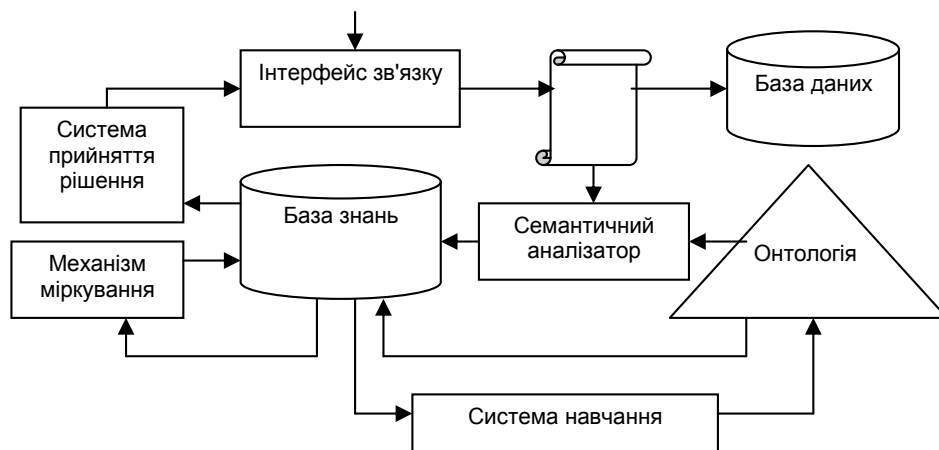


Рис. 6. Архітектура підсистеми залучення компонентів служби до процесу аналізу статті

Основними компонентами системи E-ditorial (рис. 7) є:

- інтелектуальний агент, реалізований на базі Wget, який працює під управлінням ОС Linux і забезпечує видобування публікацій з мережі Інтернет;
- супровідна база даних під управлінням СУБД MySQL, в якій зберігається інформація про користувачів, відповідні предметні області, запити користувачів а також прийняті рішення щодо статей;
- програмний пакет синтаксично-семантичного аналізу на базі LinkParser, який забезпечує побудову семантичних образів статей для їх подальшого опрацювання, а також для автоматичного поповнення онтології системи;
- онтологія, реалізована мовою OWL засобами Protégé API, структура та зміст якої оптимізовані у відповідності до інформаційного профілю фахового видання;
- підсистема класифікації/прийняття рішень щодо статті, доповненої контекстом з онтології.

Система здійснює семантичний аналіз наукової публікації, внаслідок якого будується ЗКГ статті у форматі OWL. Отриманий ЗКГ використовуємо для розв'язування вище описаних задач, та для прийняття відповідних рішень редакційною колегією.

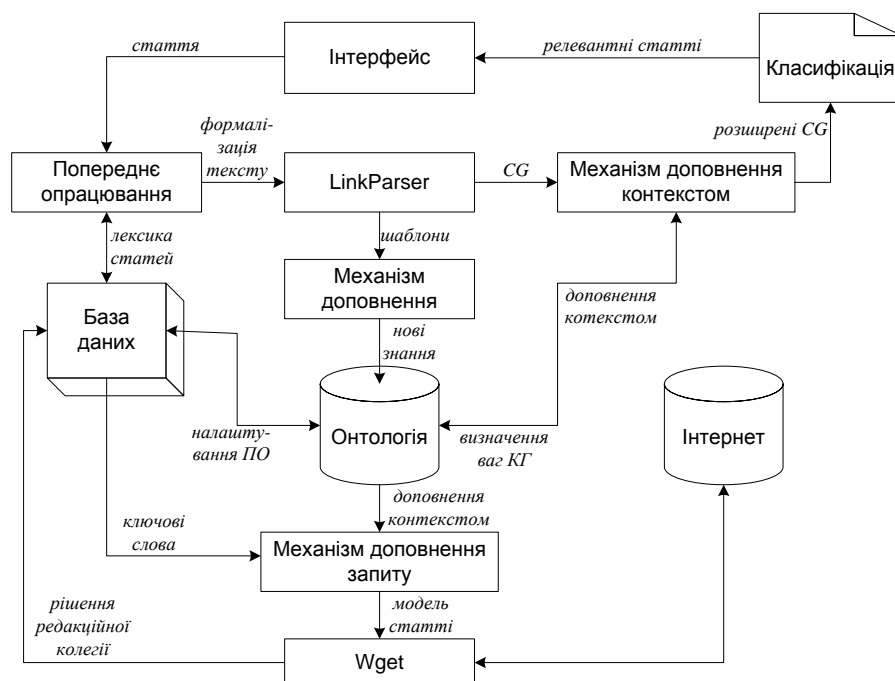


Рис. 7. Архітектура підсистеми семантичного аналізу статті

## Висновки

У роботі розглянуто архітектуру віртуальної самокерованої редакційної колегії фахового періодичного видання E-ditorial. В основі функціонування такої системи лежить інтелектуальний агент, який містить методи та засоби автоматичної підготовки до електронного друку поданих статей і знаходить найближчого за тематикою рецензента. Для реалізації цих функцій побудовано метрику між концептуальними графами, які є моделями поданих до друку статей. Вводиться поняття центру семантичної ваги концептуального графа, а відстань від сфери компетентності рецензента до статті визначається як відстань між центрами семантичних ваг відповідних концептуальних графів. Для кластеризації статей фахового видання використано теорію складних мереж. Побудована система E-ditorial дозволяє автоматизувати діяльність редакційної колегії фахового періодичного видання та приймати релевантні рішення редакційною колегією.

1. Добров Г.М. Наука о науке. – К. : Наук. думка, 1989. – 304 с..
2. Граничин О.Н. Информационные технологии в управлении. – Киев. – БИНОМ. Лаборатория знаний. : ИНТУИТ, 2008. – 336 с.
3. Мриглод О. І. Компоненти інформаційної системи підтримки рішень для наукових періодичних видань. Автореф. дис. ... зступеня кандидата наук. / НУ «Львівська політехніка», 2009.
4. Leitner F. A text-mining perspective on the requirements for electronically annotated abstracts. – FEBS. – Madrid, Spain. – 2008. – P. 1178–1181.
5. Хорошевский В.Ф. Пространства знаний в сети Интернет и Semantic Web // Искусственный интеллект и принятие решений. – Институт системного анализа РАН, № 1, 2008. – С. 80–97.
6. Седжвик Р. Фундаментальные алгоритмы на C++. Алгоритмы на графах: Пер. с англ./Роберт Седжвик. – СПб: ООО "ДиаСофтЮП", 2002. – 496 с.
7. Досин Д.Г. Інтелектуальні системи, базовані на онтологіях / Д.Г.Досин, В.В.Литвин, Ю.В.Нікольський, В.В.Пасічник / Львів: Цивілізація, 2009. – 414 с.
8. Newman M. E. J. The Structure and Function of Complex Networks. SIAM Review. – 2003. – Vol. 45. – P. 167–256.
9. Головач Ю. Складні мережі. / Головач Ю., фон Фербер К., Олемської О., Головач Т., Мриглод О., Олемської І., Пальчиков В. / Журнал фізичних досліджень. – 2006. – Том 10. – С. 247–291.
10. Girvan M. Community structure in social and biological networks. / Girvan M., Newman M.E.J. / Proc. Natl. Acad. Sci. USA. – 2002. – Vol. 99. – P. 7821–7826.