

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
РІВНЕНСЬКИЙ ДЕРЖАВНИЙ ГУМАНІТАРНИЙ УНІВЕРСИТЕТ
ФІЛОЛОГІЧНИЙ ФАКУЛЬТЕТ
КАФЕДРА РОМАНО-ГЕРМАНСЬКОЇ ФІЛОЛОГІЇ

МОВНА КОНЦЕПТУАЛІЗАЦІЯ АКСІОЛОГІЧНОГО ПРОСТОРУ ЛЮДИНИ

Колективна монографія

ІСТОРІЯ ТА СЬОГОДЕННЯ
ФІЛОЛОГІЧНОЇ НАУКИ

Том 6



РІВНЕНСЬКИЙ ДЕРЖАВНИЙ ГУМАНІТАРНИЙ УНІВЕРСИТЕТ
РІВНЕ 2023

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
РІВНЕНСЬКИЙ ДЕРЖАВНИЙ ГУМАНІТАРНИЙ УНІВЕРСИТЕТ
ФІЛОЛОГІЧНИЙ ФАКУЛЬТЕТ
КАФЕДРА РОМАНО-ГЕРМАНСЬКОЇ ФІЛОЛОГІЇ

МОВНА КОНЦЕПТУАЛІЗАЦІЯ АКСІОЛОГІЧНОГО ПРОСТОРУ ЛЮДИНИ

Колективна монографія

ІСТОРІЯ ТА СЬОГОДЕННЯ
ФІЛОЛОГІЧНОЇ НАУКИ

Том 6



РІВНЕНСЬКИЙ ДЕРЖАВНИЙ ГУМАНІТАРНИЙ УНІВЕРСИТЕТ
РІВНЕ 2023

УДК 81'37; 003; 81'22, 004.83:316.77

М 74

Рецензенти:

Леміш Н.Є. доктор філологічних наук, професор
Український державний університет
імені Михайла Драгоманова (м. Київ)

Толчеева Т.С. доктор філологічних наук, професор
Український державний університет
імені Михайла Драгоманова (м. Київ)

Рекомендовано до друку Вченою радою
Рівненського державного гуманітарного університету
(*протокол № 6 від 25 травня 2023 р.*)

Колектив авторів: Аладько Д.О., Деменчук О.В., Калініченко М.М.,
Константінова О.В., Кучма Т.В., Летюча Л.П., Мізін К.І., Павлова О.І.,
Павловська Л.О., Чеберяк А.М.

Мовна концептуалізація аксіологічного простору людини:
М 74 колективна монографія / кол. авт.; голов. ред. О.В. Деменчук. Рівне:
РДГУ, 2023. 161 с. (Історія та сьогодення філологічної науки; т. 6).
Linguistic conceptualisation of person's axiological space: Collective
work / auth.; O.V. Demenchuk (ed.). Rivne: RSUH, 2023. 161 p. (History and
Contemporaneity of Philological Science; v. 6).

У колективній монографії представлено розвідки, предметом аналізу яких є особливості мовної концептуалізації ціннісної картини світу. На матеріалі романських, германських та слов'янських мов (французька, англійська, німецька, польська, чеська, болгарська, українська) розглянуто низку актуальних питань, присвячених семантико-когнітивному моделюванню аксіологічного простору людини, дослідженню механізмів формування та реалізації ціннісних концептів у різноструктурних мовах, вивченню оцінних ситуацій у структурі тексту та дискурсу, висвітленню проблем методології лінгвістичної оцінки.

The collective work deals with the linguistic conceptualisation of the axiological worldview. Based on the Romance, Germanic and Slavic languages (French, English, German, Polish, Czech, Bulgarian, Ukrainian), the paper focuses on the topical issues related to the semantic and cognitive modelling of person's axiological space, the study of value concepts in contrastive and typological aspects, value-based situations in the text and discourse structures, linguistic evaluation methods.

УДК 81'37; 003; 81'22, 004.83:316.77

©Автори публікацій, 2023

©Рівненський державний гуманітарний університет, 2023

ЗМІСТ

**КОГНІТИВНІ МОДЕЛІ
В СЕМАНТИЦІ ОЦІННОЇ ЛЕКСИКИ**

<i>Деменчук О.В.</i> ОЦІННИЙ ФОКУС ПЕРЦЕПТИВНИХ НОМІНАЦІЙ	5
<i>Кучма Т.В.</i> ЗАСОБИ ВЕРБАЛІЗАЦІЇ КАТЕГОРІЇ ОЦІНКИ (НА МАТЕРІАЛІ АСОЦІАТИВНОГО ЕКСПЕРИМЕНТУ)	22
<i>Мізін К.І., Летюча Л.П.</i> ЦІННІСНІ ПРІОРИТЕТИ КРІЗЬ ПРИЗМУ ПСИХОЛІНГВІСТИЧНОГО ЕКСПЕРИМЕНТУ ТА СОЦІОЛІНГВІСТИЧНОГО МОНІТОРИНГУ: ЗІСТАВНО-ЛІНГВОКУЛЬТУРОЛОГІЧНИЙ АНАЛІЗ .	31

**ЦІННІСНІ КОНЦЕПТИ
В ЗІСТАВНО-ТИПОЛОГІЧНОМУ ВИСВІТЛЕННІ**

<i>Аладько Д.О.</i> ОЦІННА ХАРАКТЕРИСТИКА ЛЮДИНИ ЗАСОБАМИ ФРАЗЕОЛОГІЗМІВ З КОМПОНЕНТОМ-ЗООНІМОМ В АНГЛІЙСЬКІЙ ТА УКРАЇНСЬКІЙ МОВАХ	58
<i>Павловська Л.О.</i> ЕКСПРЕСИВНО-ОЦІННА СЕМАНТИКА ВЕРБАЛЬНИХ ФОРМУЛ ПОБАЖАНЬ У РІЗНОСТРУКТУРНИХ МОВАХ	66

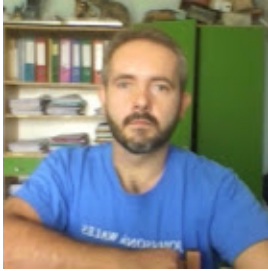
ОЦІННІ СИТУАЦІЇ В СТРУКТУРІ ТЕКСТУ ТА ДИСКУРСУ

<i>Константинова О.В.</i> АКСІОЛОГІЧНИЙ АСПЕКТ ПЕРЕКЛАДУ ІРОНІЇ В РОМАНІ ШОНА БАЙЗЕЛЛА «ЩОДЕННИК КНИГАРЯ»	99
<i>Чеберяк А.М.</i> КОНФРОНТАТИВНА СТРАТЕГІЯ ДИСКРЕДИТАЦІЇ В АСПЕКТІ РЕАЛІЗАЦІЇ НЕГАТИВНО-ОЦІННОГО ВПЛИВУ	127

МЕТОДОЛОГІЯ ЛІНГВІСТИЧНОЇ ОЦІНКИ

<i>Калініченко М.М.</i> СЕМАНТИЧНІ МОДЕЛІ В ЛІТЕРАТУРНОМУ АНАЛІЗІ ТА АВТОРСЬКОМУ ПРАВІ	138
<i>Павлова О.І.</i> ЕПОНІМИ ЯК МЕТАОЦІННІ НОМІНАЦІЇ	146

МЕТОДОЛОГІЯ ЛІНГВІСТИЧНОЇ ОЦІНКИ



Калініченко М.М.

кандидат філологічних наук, доцент

Рівненський державний гуманітарний університет

СЕМАНТИЧНІ МОДЕЛІ В ЛІТЕРАТУРНОМУ АНАЛІЗІ ТА АВТОРСЬКОМУ ПРАВІ

Семантичні моделі, що увійшли до арсеналу світової гуманітаристики ще у 1970-ті роки та набули загального розповсюдження наприкінці минулого сторіччя, революціонізували магістральні напрямки літературознавства та сферу комп'ютерної обробки текстів, дозволивши машинам розуміти людську мову й виконувати такі завдання як переклад і статистичні дослідження текстів. Розглянемо особливості практичного застосування семантичних моделей у двох споріднених галузях, де науково-методичні засоби, які пропонуються фахівцям семантичні моделі, виявилися особливо корисними – мова йде про експертизу авторського права та літературний аналіз.

У галузі експертизи авторського права, семантичні моделі використовуються для аналізу значних обсягів тексту та ідентифікації можливих порушень авторських прав. Ці моделі здатні виявляти схожості між текстами, навіть якщо у них наявні радикальні перефразування, аналізуючи імпліцитний семантичний субстрат змісту слів та їхніх комбінацій. Застосування семантичних моделей у експертизі авторських прав та літературному аналізі істотно покращує ефективність і точність текстуального аналізу. Використання семантичних моделей у цих областях має величезний потенціал на майбутнє, зокрема, у програмах виявлення плагіату, або в розробці цифрових інструментів для допомоги експертам з авторських прав у їх роботі. Однак важливо усвідомлювати, що ці моделі також мають свої обмеження, і використовувати їхній інструментарій у поєднанні з іншими методами аналізу, щоб забезпечити більш цілісне і об'єктивне розуміння тексту. Зрештою, як це відбувається із будь-якою іншою новою технологією, ефективність

семантичних моделей для експертизи авторських прав та літературного аналізу немалою мірою залежить від фахових компетентностей експертів і науковців.

Latent Semantic Analysis (LSA) одна з найперших і найпопулярніших семантичних моделей літературного аналізу, що активно використовується у експертизі авторського права. Дана модель була представлена в роботах Томаса К. Ландавера, Пітера В. Фольца і Даррелла Лахама ще наприкінці 1990-х років. LSA – це математична, кількісна техніка, яка використовує декомпозицію з метою дослідження латентної семантичної структури тексту (Landauer, Foltz, Laham., 1998, p. 259-284.). Її використовують у різноманітних комп'ютерних застосунках для виявлення випадків плагіату та порушення авторського права.

Іншою популярною семантичною моделлю, яку використовують у витонченості авторського права, є тематичне моделювання, яку вперше представили Девід Блей, Ендрю Ні та Майкл І. Джордан у 2003-му році (Blei, Jordan, 2003, p. 993-1022). Тематичне моделювання – це техніка машинного навчання, яка ідентифікує латентні теми в тексті на основі співвідношення слів. Вона використовується в сфері авторського права для визначення теми твору та характеристики відмінностей досліджуваного тексту у порівнянні із оригінальним твором з метою визначення можливості добросовісного використання творів згідно з доктриною «fair use».

Загалом, як свідчить історія використання семантичних моделей літературного аналізу у сфері авторського права, цей науковий напрямок має істотний вплив на галузь, дозволяючи більш ефективно і точно виявляти випадки плагіату та порушення авторського права.

Розглянемо деякі з магістральних напрямків застосування семантичних моделей у літературному аналізі та експертизах текстів як об'єктів авторського права.

Одна з популярних семантичних моделей, яка використовується у експертизі авторського права, – це латентний семантичний аналіз (LSA). Його сутність полягає у аналізі взаємозв'язків між словами та виявленні їхньої імпліцитної семантики. Ця модель може бути використана для порівняння тексту двох документів та виявлення подібності змісту, навіть якщо використані відмінні слова або фрази. Ідентифікуючи ці схожості, експерти із авторського права можуть визначити, чи є один документ несанкціонованою копією іншого.

Для прикладу, з метою дослідження певного документа, що містить ознаки академічного плагіату, ми можемо використовувати семантичні моделі для аналізу цього документа та порівняння його з відомими джерелами інформації, щоб визначити, чи плагіат дійсно може бути підтверджений. З цією метою LSA застосовується для генерації так званої «векторно-просторової моделі» для документа, яка представляє досліджуваний об'єкт у вигляді набору векторів у багатовимірному просторі, де кожен вектор представляє різні слова або фрази у тексті. Наступним кроком є порівняння векторно-просторової моделі із аналогічними моделями інших відомих документів, таких як раніше опубліковані роботи або інші джерела інформації, щоб виявити ознаки схожості (Manning, Schütze, 1999, p. 123).

Подібним чином латентний семантичний аналіз у літературознавстві може бути використаний для визначення теми, аналізу розвитку персонажів та розуміння багаторівневого змісту тексту. Ця модель допомагає літературним критикам і науковцям відшукати такі приховані елементи тексту, які залишаються непоміченими під час першого «поверхневого читання» (що цілком відповідає основоположним принципам «close reading» Нової Критики).

Окрім латентного семантичного аналізу, у літературознавстві використовується так зване тематичне моделювання (topic modeling). Тематичне моделювання працює шляхом аналізу слів, котрі використовуються в тексті та групує їх за темами, заснованими на їх латентному змісті. Ця модель може допомогти літературним аналітикам визначити повторювані теми в тексті, а також простежити взаємозв'язки між різними персонажами і сюжетними точками. Втім, цей науково-методичний засіб використовується й у авторському праві. У галузі експертизи таких об'єктів авторського права як літературні тексти, згадані семантичні моделі можуть бути використані для аналізу значних обсягів тексту та виявлення можливих порушень авторських прав. Ці моделі здатні виявляти схожості між різними текстами, навіть якщо використані слова різні, аналізуючи латентний зміст цих слів (Aggarwal, Zhai, 2012; 2014, p. 208).

Topic modeling або ж тематичне моделювання виявилось цінним інструментом для експертизи об'єктів авторського права у кількох напрямках. Ось кілька конкретних прикладів:

- виявлення плагіату: topic modeling може використовуватись для виявлення випадків плагіату, порівнюючи теми документа із тими, що відомі з інших джерел. Наприклад, якщо два документи мають високу ступінь схожості в розподілі й розкритті змісту теми, це може свідчити про те, що один документ було скопійовано з іншого;

- визначення оригінальності: тематичне моделювання також може використовуватись для визначення оригінальності документа, порівнюючи його теми з тими, що відомі з раніше опублікованих робіт. Якщо документ має унікальний набір тем, які не дуже схожі з темами будь-яких інших відомих робіт, це може свідчити про те, що документ є оригінальним;

- аналіз порушення авторських прав: тематичне моделювання може використовуватись для аналізу порушення авторських прав, ідентифікуючи теми, які найбільш пов'язані з певним автором чи роботою. Порівнюючи ці теми з тими, що відомі з інших робіт, експерти можуть ідентифікувати можливі випадки порушення авторських прав.

Аналіз допустимого (несанкціонованого) використання: тематичне моделювання також може використовуватись для аналізу допустимого використання, ідентифікуючи теми, які найбільш пов'язані з певним використанням роботи. Порівнюючи ці теми з тими, що відомі з інших робіт, експерти можуть визначити, чи є певне використання трансформативним (тобто набуває ознак оригінального твору) чи похідним щодо оригінального тексту (Rosenberg, 2015, p. 98).

Хоча LSA й тематичне моделювання широко застосовуються у різних царинах аналізу тексту, вони мають деякі ключові відмінності, на які варто звернути увагу науковцям і фахівцям судової експертизи. LSA більше спрямоване на виявлення подібності семантичного субстрату між порівнюваними документами, в той час як тематичне моделювання більше спрямоване на виявлення спільних тем у межах одного документа. LSA також дещо краще підходить для виявлення академічного плагіату та порушення авторських прав.

Крім LSA й тематичного моделювання, існують й інші семантичні моделі, котрі набули розповсюдження у науковій практиці фахівців експертизи авторського права та літературного аналізу.

Векторне моделювання слів або ж word embedding – це ще один тип семантичної моделі, яка представляє слова у багатовимірному просторі їхніх лексичних зав'язків на основі контексту в семантичній структурі текстів (McMenamin, 2017, p. 49). У випадку експертизи

авторського права, векторне моделювання може бути використане для виявлення випадків плагіату, порівнюючи вектори різних документів. У літературному аналізі вони можуть бути використані для аналізу використання специфічної мови, еволюції персонажів, теми твору та ін.

Розпізнавання сутностей, що мають назви та імена (Named entity recognition, або ж NER): NER – це ще один тип семантичної моделі, яка ідентифікує та класифікує номіновані сутності в тексті, такі як люди, організації та місця. У випадку експертизи авторського права, NER може бути використане для виявлення посилань на захищені авторські права роботи в документі. У літературному аналізі воно може бути використане для аналізу відносин між різними персонажами, сюжетними елементами тощо.

Аналіз настроїв (sentiment analysis) – це специфічний різновид семантичної моделі, яка акцентує емоційний тон певного тексту. У випадку експертизи авторського права, аналіз настроїв може бути використаний для визначення того, чи є документ пародією, або прямим копіюванням вже існуючого твору (Jurafsky, Martin, 2014, р. 206-209). У літературному аналізі воно може бути використане для аналізу настрою і тону тексту, а також того, як вони змінюються із розвитком сюжету.

Синтаксичний аналіз збігів – це тип семантичної моделі, яка аналізує граматичну структуру речення. У випадку експертизи авторського права, цей тип може бути використаний для виявлення випадків парафразування в документах. У літературному аналізі використовується для аналізу синтаксису і структури тексту та його рецепції.

Загалом, існує багато різних типів семантичних моделей, які можуть бути використані у питаннях експертизи авторського права та літературного аналізу, й кожна з них має свої переваги і обмеження. Отже, вибір моделі буде залежати від конкретних потреб аналізу.

Хоча семантичні моделі довели свою ефективність як інструменти у сферах експертизи авторських прав та літературного аналізу, втім, вони мають кілька помітних обмежень та потенційних методичних недоліків. Назвемо деякі із основних проблем, що виникають при використанні семантичних моделей.

У першу чергу, мова йде про надмірну залежність від кількісного математичного аналізу. Семантичні моделі часто використовуються для обробки великого обсягу текстів, що може

призвести до акцентування кількісних показників та мінімізації значущості якісного експертного аналізу. Це, у свою чергу, неодмінно призводить до спотворення результатів й надавання хибної переваги автоматизованим засобам підрахунку.

Кожна семантична модель має власні переваги та обмеження. Отже, не існує жодної ідеальної моделі для усіх типів аналізу. Наприклад, застосування LSA призводить до помітних труднощів з ідентифікацією плагіату у тих випадках, коли порівнювані документи дуже відрізняються за змістом і структурою, тоді як тематичне моделювання потерпає від проблеми з ідентифікацією тонких семантичних структур в художніх текстах.

До числа недоліків належить й упередженість комп'ютерних моделей аналізу: семантичні моделі штучного інтелекту навчаються на великих обсягах текстової інформації, що може призвести до виникнення своєрідної упередженості у написанні алгоритму. Наприклад, модель, тривалий час навчена на наборі даних, в якому переважають автори-чоловіки, матиме певні складнощі з точним аналізом текстів, написаних жінками.

Відсутність контексту: Семантичні моделі, застосовані для комп'ютерного аналізу, нерідко потерпають від проблем з урахуванням широкого семантичного контексту досліджуваного текстуального об'єкту. Зокрема, вони не здатні точно ідентифікувати культуральне або історичне значення певного тексту.

Неправильна інтерпретація значення: семантичні моделі ґрунтуються на статистичних структурах даних для ідентифікації підтексту, що може призводити до хибної інтерпретації семантичного значення. Наприклад, модель може виявити схожість між двома текстами, що базується на поверхневих лінгвістичних ознаках, а не на подібності визначального змістового підтексту.

Підводячи підсумки необхідно акцентувати, що попри те, що семантичні моделі мають багато переваг у сферах експертизи авторських прав та літературному аналізі, важливо усвідомлювати їхні обмеження та використовувати їх у поєднанні з іншими методами, щоб забезпечити більш комплексне розуміння тексту, що розглядається.

Ось декілька способів вирішення потенційних проблем з використанням семантичних моделей у галузі авторського права та літературному аналізі:

- комбінування кількісного та якісного аналізу: Щоб уникнути надмірного впливу автоматизованого кількісного аналізу, важливо

також включати в процес якісний аналіз, який здійснює людина-експерт. Це передбачає уважне прочитання текстів, врахування ширшого культурного та історичного контексту, фінальну експертну оцінку;

- використання відразу декількох моделей: замість того, що застосовувати лише одну семантичну модель, необхідно використовувати відразу декілька моделей для аналізу певного тексту. Це допоможе подолати обмеження кожної моделі та забезпечити більш детальний та комплексний підхід;

- вирішення проблеми надмірного впливу окремого комплексу даних: важливо переконатися, що набір даних, який використовується для навчання семантичної комп'ютерної моделі, є різноманітним та репрезентативним. Це передбачає збирання даних із різних джерел та використання варіативних методів для забезпечення рівноваги репрезентації різних груп у наборі даних алгоритму комп'ютерної моделі.

- студіювання ширшого контексту: хоча семантичні моделі здатні ефективно ідентифікувати подібні фрагменти в тексті, вони не завжди можуть врахувати культурний або історичний контекст, в якому розташований досліджуваний текст в цілому. Щоб вирішити цю проблему, важливо враховувати контекст, в якому було створено текст, та обов'язково залучати до формування кінцевих висновків проміжні умовиводи людини-експерта.

- верифікація результатів експертизи комісією спеціалістів: для уникнення хибної інтерпретації, результати семантичного аналізу можуть бути розглянуті групою спеціалістів, котрі добре обізнані із контекстом досліджуваних об'єктів. Це може допомогти забезпечити особливу точність аналізу та виявити важливі нюанси розуміння тексту.

В цілому, застосування вищеназваних підходів уможливило ефективне використання семантичних моделей в експертизі авторського права та літературному аналізі, а також допомагає ліквідувати чи гранично послабити їхні функціональні обмеження.

Представлений у публікації огляд сутності та особливостей використання семантичних моделей у експертизі авторського права та літературному аналізі дозволяє зробити наступні висновки. Семантичні моделі значно поліпшують ефективність аналізу текстів. Ці моделі використовують передові технології для ідентифікації структурних патернів, тем і схожих елементів у текстах, а також

дозволяють фахівцям отримати глибше розуміння творів, які вони аналізують.

Хоча семантичні моделі мають багато переваг, вони також мають свої функціональні обмеження. Важливо бути обізнаним із цими обмеженнями та використовувати аналітичні моделі у поєднанні з іншими методами для забезпечення більш комплексного розуміння тексту.

Є кілька способів вирішення можливих проблем з використанням семантичних моделей, зокрема поєднання кількісного та якісного аналізу, використання відразу декількох моделей, врахування упереджень селекції аналітичних відомостей, студіювання ширшого контексту, додаткова перевірка результатів експертами-людьми.

Зрештою, використання семантичних моделей в експертизі авторського права та літературному аналізі має величезну практичну цінність, але важливо бути обізнаним з їх обмеженнями та використовувати їх разом з іншими методами аналізу для забезпечення максимального об'єктивного і всебічного дослідження. Із розвитком технологій будуть розроблені нові й досконаліші семантичні моделі, які дозволять проводити ще більш детальний аналіз текстів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- Aggarwal, C.C., & Zhai, C. (2012). *Mining text data*. Berlin: Springer Science & Business Media.
- Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003). Latent Dirichlet allocation. *Journal of machine learning research*, 3(Jan), 993-1022.
- Jurafsky, D., & Martin, J.H. (2014). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. London: Pearson Education.
- Landauer, T.K., Foltz, P.W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, 25(2-3), 259-284.
- Manning, C.D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: The MIT Press.
- McMenamin, G.R. (2017). *Basic text mining: classify, cluster, visualize data*. CRC Press.
- Rosenberg, M. (2015). *Text mining with R: A tidy approach*. Sebastopol: O'Reilly Media, Inc.

Наукове видання

МОВНА КОНЦЕПТУАЛІЗАЦІЯ
АКСІОЛОГІЧНОГО ПРОСТОРУ ЛЮДИНИ

Колективна монографія

Історія та сьогодення філологічної науки

Том 6

Academic Edition

LINGUISTIC CONCEPTUALISATION
OF PERSON'S AXIOLOGICAL SPACE

Collective work

History and Contemporaneity of Philological Science

Volume 6

Формат 60x84¹/₁₆.

Папір офсетний. Друк ризографічний.

Гарнітура Century Schoolbook.

Авт. арк. 8,8. Наклад 100.

Редакційно-видавничий відділ

Рівненського державного гуманітарного університету

м. Рівне, 33000, вул. Ст. Бандери, 12

тел.: (0362) 26-48-83

Друк: «ПараГраф+»

м. Рівне, 33000, вул. Гетьмана Мазеши, 6/8

тел.: (0362) 26-49-27