



УДК 37.013

## MULTIMEDIA DATABASES WITH CUSTOMIZATION FOR USER MODALITY

### МУЛЬТИМЕДІЙНІ БАЗИ ДАНИХ З НАЛАШТУВАННЯМ НА КОРИСТУВАЛЬНУ МОДАЛЬНІСТЬ

**Hordyeyev A.S. / Гордєєв А.С.***d.t.s., prof. / д.т.н., проф.*

ORCID: 0000-0002-9601-6220

*Simon Kuznets Kharkiv National University of Economic,**Kharkiv, Nauky Avenue 9a, 61165**Харківський національний економічний університет імені Семена Кузнеця,**Харків, проспект Науки 9а, 61165*

**Анотація.** В роботі розглядається проблема підвищення ефективності мультимедійних баз даних за рахунок урахування когнітивних особливостей здобувачів вищої освіти, зокрема їх модальності сприйняття інформації. Запропоновано формальну модель користувальницької модальності, представлену у вигляді нормованого вектора переваг, що відображає ступінь домінування візуального, аудіального, текстового та кінестетичного каналів сприйняття. Розроблено метод представлення мультимедійного контенту в тому ж модальному просторі, що забезпечує можливість кількісного зіставлення користувача та даних. На цій основі реалізовано механізм адаптивного ранжування, що враховує як семантичну релевантність, так і відповідність форми подання інформації когнітивним уподобанням користувача. Проведено педагогічний експеримент за участю здобувачів, результати якого показали статистично значуще покращення ключових метрик:  $Precision@10$ ,  $NDCG@10$ , приросту знань та задоволеності користувачів. Встановлено, що облік користувальницької модальності дозволяє підвищити релевантність видачі, знизити когнітивне навантаження та збільшити ефективність освітнього процесу. Отримані результати підтверджують перспективність інтеграції когнітивних характеристик користувача в архітектуру мультимедійних інформаційних систем та формують основу розробки адаптивних навчальних систем нового покоління.

**Ключові слова:** мультимедійні бази даних; користувальницька модальність; когнітивні особливості; мультимодальні дані; ранжування контенту; освітні навчальні системи.

### Вступ.

У разі цифрової трансформації суспільства мультимедійні дані стають домінуючою формою представлення інформації. Текст, зображення, аудіо, відео та інтерактивний контент формують основу сучасних інформаційних систем, освітніх платформ, медіа-сервісів та професійних програм, включаючи сферу поліграфічного виробництва.

Зростання обсягів мультимедійного контенту зумовлене розвитком інтернет-технологій, мобільних пристроїв та хмарних обчислень. Користувачі щодня взаємодіють із величезними масивами різномірних даних, що потребує



ефективних методів їх зберігання, обробки та інтелектуального вилучення. На відміну від традиційних текстових даних, мультимедійна інформація характеризується високою розмірністю, гетерогенністю та складністю семантичної інтерпретації, що суттєво ускладнює завдання організації баз даних.

Особливої актуальності мультимедійні дані набувають в освітньому середовищі. Використання візуальних, аудіальних та інтерактивних матеріалів дозволяє підвищити залученість здобувачів, покращити розуміння складних технологічних процесів та забезпечити більш глибоке засвоєння знань. У дисциплінах технічного профілю, таких як «Технологія поліграфічного виробництва», мультимедійний контент відіграє ключову роль, оскільки дозволяє демонструвати реальні виробничі процеси, обладнання та технологічні схеми.

Разом з тим збільшення обсягу мультимедійних даних призводить до проблеми інформаційного навантаження. Користувачі стикаються з необхідністю швидкого пошуку релевантної інформації, що вимагає розробки інтелектуальних систем, здатних враховувати як зміст даних, а й особливості сприйняття інформації конкретним користувачем.

Сучасні бази даних вирішують завдання зберігання та пошуку, однак у більшості випадків залишаються орієнтованими на універсального користувача. Це обмежує ефективність їх застосування, особливо в освітніх системах, де індивідуальні відмінності у сприйнятті інформації істотно впливають на результати навчання.

Таким чином, актуальність дослідження визначається необхідністю розробки мультимедійних баз даних нового покоління, здатних:

- ефективно обробляти різномірний контент,
- забезпечувати персоналізований доступ до інформації,
- враховувати когнітивні особливості користувачів.

Особливого значення набуває інтеграція моделей користувальницької модальності в архітектуру мультимедійних систем, що дозволяє підвищити релевантність видачі, покращити користувальницький досвід та збільшити



ефективність освітнього процесу.

Користувацька модальність є ключовим когнітивним параметром, що визначає кращий спосіб сприйняття інформації. У контексті мультимедійних баз даних даний фактор набуває особливої значущості, оскільки контент представлений у різних формах — візуальної, аудіальної, текстової та інтерактивної.

З погляду когнітивної науки, модальність відбиває домінуючий канал обробки інформації користувачем. Традиційно виділяють візуальну, аудіальну та текстову (вербальну) модальності, а також кінестетичну, пов'язану з інтерактивною взаємодією. В інформаційних системах ці категорії можуть бути формалізовані та використані для адаптації контенту.

Сучасні бази даних відіграють ключову роль у зберіганні, організації та витягуванні різнорідного цифрового контенту, що включає текст, зображення, аудіо, відео та інтерактивні елементи. Однак, незважаючи на значний прогрес у галузі обробки мультимедійної інформації, існуючі системи в більшості випадків орієнтовані на універсальні механізми пошуку та подання даних, що не враховують індивідуальні особливості користувача.

Традиційні підходи до організації мультимедійних баз даних ґрунтуються на текстових, контентно-орієнтованих та семантичних методах пошуку. Незважаючи на їх ефективність у ряді завдань, вони мають ряд обмежень. Зокрема, текстові методи не здатні адекватно відображати зміст мультимедійних об'єктів, контентно-орієнтовані методи вимагають значних обчислювальних ресурсів і не враховують контекст користувача, а семантичні підходи стикаються з труднощами формалізації та анотування даних. В результаті користувач отримує результати, які формально відповідають запиту, але не завжди оптимальні з погляду сприйняття та зручності використання.

Додатковою проблемою є інформаційне навантаження, обумовлене стрімким зростанням обсягу мультимедійного контенту. Користувачеві необхідно швидко знаходити релевантну інформацію, проте стандартні механізми ранжування не враховують відмінності у когнітивних стилях



сприйняття. У освітньому середовищі це призводить до зниження ефективності навчання, оскільки той самий матеріал може по-різному сприйматися різними здобувачами.

Існуючі рішення в області адаптивних систем в основному зосереджені на рівні інтерфейсу користувача або рекомендаційних алгоритмів, не торкаючись внутрішню структуру мультимедійної бази даних. При цьому користувач зазвичай описується через історію взаємодій або демографічні характеристики, що не відображає його когнітивних переваг, таких як домінуюча модальність сприйняття (візуальна, аудіальна, текстова або кінестетична).

Таким чином, виникає протиріччя між:

- зростаючим обсягом та складністю мультимедійних даних,
- необхідністю персоналізації інформаційної взаємодії,
- та обмеженими можливостями існуючих мультимедійних баз даних.

Ця суперечність визначає основну наукову проблему дослідження, яка полягає у відсутності моделей та методів, що дозволяють інтегрувати когнітивні характеристики користувача безпосередньо у процеси зберігання, пошуку та подання мультимедійних даних.

Вирішення цієї проблеми вимагає розробки адаптивних мультимедійних баз даних нового покоління, в яких:

- користувальницька модель стає частиною структури даних,
- механізми ранжування враховують модальність сприйняття,
- подання контенту динамічно адаптується до користувача.

Особливої актуальності дана проблема набуває в освітніх системах, де облік індивідуальних особливостей учнів безпосередньо впливає якість засвоєння знань. Зокрема, у дисципліні «Технології поліграфічного виробництва» ефективне освоєння матеріалу вимагає використання різних форм подання інформації, що робить завдання адаптації мультимедійного контенту критично важливим.

Таким чином, необхідність створення адаптивних мультимедійних баз даних, що враховують користувальницьку модальність, є ключовим завданням



сучасної інформаційної науки та визначає напрямок цього дослідження.

## **1. Сучасний стан досліджень у галузі мультимедійних баз даних з адаптацією до користувальницької модальності**

Сучасний розвиток мультимедійних інформаційних систем характеризується переходом від традиційних моделей зберігання та обробки даних до інтелектуальних, адаптивних рішень, здатних враховувати особливості користувача та контекст його взаємодії. Важливу роль у цьому процесі відіграли дослідження в області мультимодальних інтерфейсів, які забезпечують взаємодію через різні канали, включаючи голос, жести, візуальні елементи і тактильні сигнали [1]. Такі інтерфейси дозволяють наблизити взаємодію людини із системою до природних форм комунікації, що значно підвищує зручність використання [5].

Традиційні системи управління базами даних (СУБД), переважно реляційного типу, спочатку розроблялися для обробки структурованих даних із чітко заданою схемою. Їхня архітектура оптимізована для роботи з числовими та текстовими даними, проте при переході до мультимедійного контенту (зображення, аудіо, відео, інтерактивні об'єкти) виявляється низка фундаментальних обмежень.

По-перше, суттєвою проблемою є невідповідність моделі даних. Реляційна модель передбачає подання інформації як таблиць з фіксованою структурою, тоді як мультимедійні дані мають складну, часто неструктуровану чи слабо структуровану природу. Наприклад, зображення або відеопотік неможливо адекватно подати у вигляді набору атрибутів без втрати значущої інформації. В результаті мультимедійний контент найчастіше зберігається у вигляді бінарних об'єктів (BLOB), що обмежує можливості його аналізу та обробки на рівні СУБД.

По-друге, традиційні СУБД не забезпечують ефективних механізмів семантичного пошуку мультимедійних даних. Пошук здійснюється переважно за метаданими або ключовими словами, що не дозволяє враховувати зміст самого об'єкта. Наприклад, система не здатна визначити, що зображення містить друкарську машину або що відео демонструє процес офсетного друку, якщо це



явно не зазначено в метаданих. Це призводить до низької релевантності результатів під час роботи з мультимедійним контентом.

Третє обмеження пов'язане із відсутністю вбудованих механізмів роботи з ознаками мультимедійних об'єктів. Сучасні методи обробки мультимедіа передбачають використання ознакового опису (наприклад, візуальні дескриптори, аудіоспектри, ембедінги), однак традиційні СУБД не призначені для зберігання та ефективної обробки таких даних, особливо у високорозмірних просторах.

Крім того, суттєвою проблемою є низька продуктивність при обробці великих обсягів мультимедійних даних. Операції пошуку та аналізу мультимедіа вимагають значних обчислювальних ресурсів, а реляційні СУБД погано масштабуються для завдань, пов'язаних із обробкою потокових даних, відео чи високорозмірних векторів. Це обмежує їх застосування у системах реального часу.

Ще одним важливим обмеженням є відсутність підтримки мультимодальності. Традиційні СУБД не враховують різницю між типами даних (візуальні, аудіальні, текстові) з погляду їх сприйняття користувачем. Всі дані обробляються однаково, без урахування когнітивних особливостей користувача, що унеможливує персоналізацію на рівні бази даних.

Також слід відзначити розрив між рівнем зберігання даних і користувальницькою взаємодією. У більшості систем адаптація реалізується лише на рівні інтерфейсу чи зовнішніх додатків, тоді як сама база даних залишається пасивним сховищем. Це призводить до фрагментації системи та зниження ефективності персоналізації.

Зрештою, традиційні СУБД не підтримують динамічної адаптації контенту. Вони не здатні змінювати спосіб подання або пріоритет видачі даних залежно від контексту користувача, його переваг або поточного завдання. У разі освітніх систем це особливо критично, оскільки різні користувачі вимагають різного формату подачі інформації.

Таким чином, основні обмеження традиційних СУБД при роботі з



мультимедіа можна узагальнити наступним чином: невідповідність моделі даних, обмежені можливості пошуку, відсутність підтримки ознакового представлення, проблеми масштабованості, ігнорування модальності користувача і відсутність адаптивності. Ці обмеження зумовлюють необхідність розробки нових підходів до побудови мультимедійних баз даних, орієнтованих інтелектуальну обробку та персоналізацію контенту.

Проте, незважаючи на значний прогрес, основна увага в існуючих дослідженнях зосереджена переважно на рівні інтерфейсу користувача, тоді як питання адаптації самих мультимедійних даних залишаються значною мірою недостатньо вивченими [7]. Зокрема, концепція пластичності інтерфейсів передбачає здатність системи адаптуватися до різних умов використання, включаючи характеристики користувача, використовуваний пристрій і фізичне середовище взаємодії [20, 18]. Однак дана адаптація, як правило, обмежується зміною форми подання інформації, не торкаючись змісту мультимедійних даних та способи їх організації в базі даних.

Додатковий розвиток отримали розподілені інтерфейси користувача, які дозволяють динамічно перерозподіляти елементи взаємодії між різними пристроями і користувачами [22]. Подібні рішення демонструють високу гнучкість та застосовність у широкому спектрі сценаріїв, включаючи спільну роботу, мультимедійні програми та інтерактивні системи. Істотний внесок у формалізацію таких систем внесла еталонна модель Cameleon, що вводить багаторівневу структуру представлення інтерфейсу користувача - від абстрактних моделей завдань до кінцевого інтерфейсу [2, 8, 9]. Незважаючи на методологічну значущість даної моделі, вона не враховує особливості сприйняття інформації користувачем та не інтегрує когнітивні параметри у процеси обробки даних.

В рамках розвитку адаптивних систем були запропоновані різні підходи до трансформації інтерфейсів користувача, включаючи модельно-орієнтовані методи, адаптивний веб-дизайн та мови опису інтерфейсів, такі як UsiXML [9]. Ці підходи забезпечують гнучкість та масштабованість інтерфейсних рішень,



однак їх застосування в контексті мультимедійних баз даних обмежене, оскільки вони не торкаються механізмів зберігання, індексування та інтелектуального пошуку мультимедійного контенту.

Особливе місце займають мета-інтерфейси, призначені для управління адаптацією інтерфейсу користувача [1, 14]. Вони дозволяють користувачеві контролювати розподіл компонентів системи, відстежувати стан інтерфейсу та налаштовувати параметри взаємодії. Незважаючи на високий рівень керованості, такі рішення вимагають активної участі користувача і не забезпечують автоматичної адаптації на основі прихованих переваг [35]. Зокрема вони не враховують когнітивні характеристики користувача, включаючи його модальність сприйняття інформації.

Аналіз існуючих досліджень показує, більшість сучасних систем не враховують відмінності у способах сприйняття інформації, як-от візуальна, аудіальна чи текстова модальність [15]. У цьому адаптація, зазвичай, реалізується лише на рівні інтерфейсу, тоді як рівень даних залишається статичним. Користувач у таких системах рідко представлений у вигляді формалізованої моделі, придатної для інтеграції у процеси пошуку та ранжирування. Крім того, мультимедійні бази даних часто обмежуються використанням метаданих без урахування мультимодальних ознак та семантичних зв'язків між різними типами контенту [35].

Перетворення можуть відбуватися у різний спосіб. Наприклад, при двосторонньому підході фреймворк злиття [1, 14] обробляє введення користувача та передає його менеджеру діалогів. На відміну від цього, при поділі [1] висновок керується на основі користувачів та їх контексту. У цей час інтерфейс користувача може слідувати закладеній парадигмі [3] або перебудовувати, перерозподіляти [4] і навіть розширювати [24] компоненти на основі певних правил або на основі зіставлення взаємозамінних елементів інтерфейсу користувача [35].

Тим часом, користувачі можуть керувати ефектами перетворень інтерфейсу користувача через спеціальну точку управління, оформлену як мета-



користувацький інтерфейс (Meta-UI) [15]. Мета-UI допомагають користувачам контролювати та оцінювати стан інтерфейсу основної програми, з якою вони взаємодіють [26]. Мета-UI приймали безліч форм і моделей поведінки, від окремих панелей управління для розподілу компонентів по інших платформах (настільні комп'ютери, мобільні пристрої) [4, 17] до складних середовищ веб-розробки (Quill) [8], фреймворків (ReWiRe) [21] або спеціалізованих редакторів завдань [15].

В останні роки розвиток технологій віртуальної та доповненої реальності, а також мультимодальних систем взаємодії, створює передумови для переходу до нового класу систем — мультимедійних баз даних, що адаптуються до модальності користувача [8, 21, 17]. Такі системи орієнтовані як зміну форми подання інформації, а й адаптацію самого змісту відповідно до когнітивними уподобаннями користувача. Це передбачає використання методів машинного навчання, мультимодальних ембеддингів та інтелектуальних алгоритмів ранжирування.

Дослідження персоналізації в мультимедійних системах зосереджені на об'єднанні анотування, пошуку та користувальницьких уподобань. Зокрема, у роботі [40] розглядаються моделі, що поєднують онтології, контекст і користувацькі переваги в єдиний ланцюжок обробки мультимедійних даних. Персоналізація реалізується переважно на рівні представлення та інтерфейсу, в даній роботі користувальницька модальність інтегрується безпосередньо в структуру мультимедійної бази даних та впливає на ранжування.

Сучасні дослідження у сфері мультимодальних рекомендаційних систем демонструють перехід до використання гетерогенних даних (текст, зображення, аудіо) які підвищують якість рекомендацій. Огляд [41] показує, що такі системи дозволяють більш точно моделювати переваги користувача і вирішувати проблему інформаційного навантаження. Аналогічно, у роботі [42] автори підкреслюють, що мультимодальні ознаки дозволяють зменшити проблему розрідженості даних та покращити якість рекомендацій. Проте, на відміну багатомодальних рекомендаційних систем, дані підходи не розглядають



модальність як когнітивну характеристику користувача, не інтегрують їх у архітектуру бази даних.

Методи мультимодального подання даних спрямовані на об'єднання різних типів інформації в єдиному латентному просторі. Такі підходи дозволяють зіставляти дані різних модальностей та покращувати завдання класифікації, пошуку та аналізу. Крім того, у роботі [40] автори відзначають відсутність універсальної схеми моделювання мультимодальних систем, незважаючи на активний розвиток цього напрямку. На відміну від вище розглянутих підходів, запропонована модель не тільки об'єднує модальності даних, а й враховує модальність користувача як окремий вимір, пов'язує представлення даних та когнітивні особливості користувача.

Однією з ключових проблем мультимедійних систем є семантичний розрив між низькорівневими ознаками та сприйняттям користувача. Дослідження в галузі крос-модального пошуку спрямовані на подолання цього розриву шляхом зіставлення різних типів даних (наприклад, текст → зображення). Більш сучасні роботи, такі як [41], пропонують методи вирівнювання модальностей та отримання семантики на основі глибокого навчання. Однак, на відміну від методів крос-модального пошуку, вони вирішують проблему відповідності даних, але не враховують відмінності у сприйнятті інформації користувачем.

Сучасні моделі, такі як [42], використовують графові нейронні мережі та механізми уваги для більш точного обліку мультимодальних ознак. Інші роботи [40] демонструють використання трансформерів, LLM та self-supervised learning для підвищення якості рекомендацій. Ці моделі моделюють переваги користувача через взаємодії, не використовують явну когнітивну модель (наприклад модальність), не застосовуються в освітніх мультимедійних БД.

Класичні підходи до отримання мультимедійної інформації ґрунтуються на: вилученні ознак, фільтрації, категоризації даних. Проте такі методи орієнтовані на дані, а ні на користувача і ні враховують когнітивні особливості сприйняття.

На основі аналізу літературних джерел можна виділити ключові обмеження:

1. Відсутність формальної моделі користувальницької модальності.



2. Фокус на даних, а ні на на когнітивних характеристиках.
3. Персоналізація лише на рівні інтерфейсу, а ні на бази даних.
4. Обмежена інтеграція мультимодальних та користувацьких моделей.

Таким чином, проведений аналіз свідчить про наявність суттєвої наукової прогалини, що полягає у відсутності комплексних моделей мультимедійних баз даних, здатних враховувати користувацьку модальність як ключовий параметр при зберіганні, обробці та витягуванні інформації. Заповнення даного пробілу вимагає розробки нових архітектурних та алгоритмічних рішень, орієнтованих на інтеграцію когнітивних характеристик користувача до ядра мультимедійних інформаційних систем.

Метою роботи є розробка та дослідження моделі мультимедійної бази даних з адаптацією до користувацької модальності, що забезпечує підвищення релевантності видачі та ефективності сприйняття інформації за рахунок урахування когнітивних особливостей користувача.

Гіпотези дослідження.

Основна гіпотеза (H1): Адаптація мультимедійного контенту на основі користувацької модальності статистично значно підвищує релевантність видачі.

Приватні гіпотези:

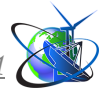
H2: Використання мультимодальних ембеддингів покращує якість персоналізації порівняно з однотипними ознаками.

H3: Динамічне оновлення модальності підвищує задоволеність користувача

H4: Вплив модальності сильніше виявляється в освітніх системах, ніж у розважальних.

## **2. Модель користувацької модальності в мультимедійних базах даних**

У пропонованому підході користувацька модальність розглядається як кількісно вимірний параметр, що відображає кращий канал сприйняття інформації та впливає на процеси пошуку, ранжирування та подання мультимедійного контенту. На відміну від традиційних моделей персоналізації,



де користувач описується через історію запитів чи демографічні характеристики, у роботі акцент робиться на когнітивних особливостях сприйняття.

### **2.1. Простір модальностей та його формалізація**

Теоретичне обґрунтування моделі користувальницької модальності базується на міждисциплінарному синтезі когнітивної психології, теорії мультимедійного навчання та методів представлення даних в інформаційних системах.

Згідно з когнітивними теоріями сприйняття, людина обробляє інформацію через кілька каналів, кожен з яких має обмежену пропускну здатність і специфічні характеристики. У рамках теорії мультимедійного навчання та теорії подвійного кодування показано, що ефективність засвоєння інформації залежить від узгодження форми подання даних з домінуючим каналом сприйняття користувача.

Практичні дослідження в галузі педагогіки та людино-комп'ютерної взаємодії підтверджують, що користувачі демонструють стійкі переваги у способах сприйняття інформації, які можна умовно класифікувати як:

- візуальні (зображення, схеми, відео),
- аудіальні (мова, звук),
- текстові (вербальне уявлення),
- кінестетичні (інтерактивна взаємодія).

Важливо, що дані модальності не є взаємовиключними. Користувач, як правило, використовує кілька каналів одночасно, проте з різним ступенем переваги. Це зумовлює необхідність безперервного, а ні дискретного уявлення модальності.

Для кількісного опису модальності користувача вводиться вектор:

$$U = (w_v, w_a, w_t, w_k), w_j \in [0, 1], \sum_j w_j = 1, \quad (1)$$

де:  $w_v$  - ступінь переваги візуального контенту (зображення, відео);  $w_a$  - аудіального контенту;  $w_t$  - текстової інформації;  $w_k$  - інтерактивної (кінестетичної) взаємодії.

Таке уявлення має кілька теоретично обґрунтованих властивостей:



1. *Нормованість*. Сума ваг дорівнює одиниці, що дозволяє інтерпретувати їх як розподіл ймовірностей уподобань користувача за модальностями.

2. *Безперервність*. Вектор визначає безперервний простір переваг, що більш адекватно відображає реальну поведінку користувача порівняно з жорсткою класифікацією (наприклад, “візуал” або “аудіал”).

3. *Інтерпретованість*. Кожна компонента має ясний зміст і може бути безпосередньо пов'язана з типом контенту.

З математичної точки зору, вектор  $U$  належить 4-вимірному симплексу:

$$\Delta^3 = \{U \in R^4 \mid w_j \geq 0, \sum w_j = 1\}. \quad (2)$$

Цей простір має ряд важливих властивостей:

- кожна точка є унікальною комбінацією модальностей;
- вершини симплексу відповідають "чистим" типам (наприклад, повністю візуальний користувач);
- внутрішні точки відбивають змішані когнітивні переваги.

Таке уявлення широко використовується в теорії ймовірностей, тематичному моделюванні, рекомендаційних системах.

Вибір векторного уявлення обумовлений такими причинами:

1. *Сумісність із методами машинного навчання*. Векторна форма дозволяє застосовувати: кластеризацію (*k-means*), класифікацію, нейронні мережі.

2. *Інтеграція з мультимедійними даними*. Мультимедійний контент може бути представлений у вигляді векторів ознак, що забезпечує єдиний простір для користувача або даних.

3. *Підтримка адаптивних алгоритмів*. Вектор  $U$  може бути безпосередньо включений у функцію ранжування та адаптації.

Кожна компонента вектора  $U$  відповідає певному типу контенту (таблиця 1).

Таким чином, модель забезпечує семантичну відповідність між уподобаннями користувача та характеристиками мультимедійного контенту. З погляду когнітивної теорії, переваги користувача можуть змінюватися в залежності від: контексту (навчання vs пошук), складності матеріалу, рівня



підготовки.

**Таблиця 1 – Семантична відповідність між користувальницькими уподобаннями та характеристиками мультимедійного контенту**

Модальність	Тип контенту	Приклади
$w_v$	візуальний	схеми друку, відео процесів
$w_a$	аудіальний	лекції, коментарі
$w_t$	текстовий	конспекти, інструкції
$w_k$	кінестетичний	симуляції, інтерактив

*Авторська розробка*

Векторна модель припускає природне оновлення:

$$U^{(t+1)} = (1 - \beta)U^{(t)} + \beta B^{(t)}, \quad (3)$$

що робить її придатною для адаптивних систем.

Таким чином, запропонована модель користувальницької модальності має суворе теоретичне обґрунтування:

- спирається на когнітивні теорії сприйняття,
- використовує математично коректне уявлення у вигляді ймовірнісного вектора,
- має геометричну інтерпретацію в просторі симплекса,
- сумісна з методами машинного навчання та мультимедійними даними.

Це дозволяє розглядати користувальницьку модальність не як допоміжну характеристику, а як формальний параметр, що інтегрується в архітектуру мультимедійних баз даних та алгоритми обробки інформації.

## **2.2. Подання мультимедійного контенту у просторі модальностей.**

Формалізація мультимедійного контенту в просторі модальностей є логічним продовженням моделі користувальницької модальності та забезпечує єдине уявлення як користувача, так і даних у загальному ознаковому просторі. Це створює основу для побудови адаптивних алгоритмів пошуку та ранжування.



З погляду когнітивної теорії мультимедійного навчання, ефективність сприйняття інформації визначається як її змістом, а й формою представлення. Будь-який мультимедійний об'єкт, незалежно від його типу, може бути розглянутий як комбінація різних каналів передачі:

- візуального (зображення, відео),
- аудіального (звук, мова),
- текстового (символьна інформація),
- інтерактивного (дії користувача).

Таким чином, мультимедійний контент за своєю природою є мультимодальним, і його подання у вигляді єдиного об'єкта потребує декомпозиції по цих каналах.

Кожен об'єкт мультимедійної бази даних  $d_i$  також представляється як модального вектора:

$$d_i = (f_v, f_a, f_t, f_k), \quad (4)$$

де  $f_j$  відображає ступінь присутності відповідної модальності в об'єкті.

Наприклад:

- відео-лекція  $\rightarrow$  високий  $f_v$  і  $f_a$ ;
- текстова стаття  $\rightarrow$  високий  $f_t$ ;
- VR-симуляція  $\rightarrow$  високий  $f_k$ .

Для забезпечення сумісності об'єктів використовується нормалізація:

$$f_j = \frac{\text{об'єм контенту типу } j}{\sum_j \text{об'єм контенту}}. \quad (5)$$

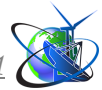
Дане перетворення забезпечує:

1. *Приведення до єдиного масштабу.* Усі компоненти перебувають у діапазоні  $[0, 1]$ .

2. *Інтерпретацію як розподілу.* Вектор  $d_i$  можна трактувати як імовірнісний розподіл модальностей.

3. *Інваріантність до розміру об'єкта.* Важливою є структура контенту, а не його абсолютний обсяг.

Приклади інтерпретації.



Відео-лекція:  $d_i \approx (0.5, 0.4, 0.1, 0.0)$ – переважають візуальний та аудіальний канали.

Текстова стаття:  $d_i \approx (0.1, 0.0, 0.9, 0.0)$ .

VR- симуляція:  $d_i \approx (0.4, 0.1, 0.0, 0.5)$ .

Таким чином, кожен об'єкт отримує кількісний опис, що відбиває його структуру.

Аналогічно для користувача моделі, вектор  $d_i$  належить симплексу:

$$d_i \in \Delta^3. \quad (6)$$

Це означає, що це об'єкти перебувають у єдиному просторі, відстань з-поміж них відбиває різницю у структурі контенту, можлива кластеризація мультимедійних об'єктів.

Подання контенту у вигляді модального вектора має такі переваги:

1. Уніфікація різних типів даних. Дозволяє уявити текст, зображення, відео, інтерактивні системи в єдиному форматі.

2. Сумісність із моделлю користувача. Оскільки користувач описується вектором:

$$U = (w_v, w_a, w_t, w_k) \quad (7)$$

стає можливим обчислення узгодженості:

$$match(U, d_i) \quad (8)$$

що є основою адаптивного ранжування.

3. Підтримка машинного навчання. Векторна форма дозволяє застосовувати кластеризацію, класифікацію та нейронні моделі.

4. Інтеграція до СУБД. Така модель може бути реалізована в базі даних як: додаткові атрибути, ознаки, що індексуються, векторні поля.

Введене уявлення дозволяє формалізувати відповідність між користувачем та контентом:

$$M(U, d_i) = w_v f_v + w_a f_a + w_t f_t + w_k f_k. \quad (9)$$

Таким чином, чим вище значення, тим краще вміст відповідає користувачу, а також з'являється можливість адаптивного ранжування.



У контексті дисципліни "Технології поліграфічного виробництва" це означає:

- візуалам → схеми друкарських машин;
- аудіалів → пояснення процесів;
- кінестетикам → симуляції;
- текстовим → інструкції.

Подання мультимедійного контенту у вигляді модального вектора: відповідає когнітивній природі сприйняття інформації; забезпечує математично коректний опис структури даних; дозволяє об'єднати користувача та контент у єдиному просторі; створює основу адаптивних мультимедійних баз даних. Таким чином, ця модель є ключовим елементом переходу від традиційних систем зберігання до інтелектуальних систем персоналізованої взаємодії.

### ***2.3. Збір та формування ознакового простору користувача***

Формування вектора модальності користувача  $U$  вимагає побудови адекватного ознакового простору, що відображає реальні когнітивні переваги користувача. У пропонованому підході використовується комбінована модель, заснована на інтеграції явних та неявних даних, що відповідає сучасним підходам у рекомендаційних системах та адаптивних інформаційних середовищах.

З погляду когнітивної науки, переваги користувача не можуть бути надійно визначені виключно на основі самооцінки, оскільки користувач не завжди усвідомлює свої когнітивні особливості, відповіді анкет можуть бути суб'єктивними, переваги можуть змінюватися в процесі взаємодії. У той же час поведінкові дані відображають реальні патерни взаємодії, проте вони схильні до шуму і залежать від контексту та інтерфейсу.

Отже, найбільш обґрунтованим є гібридний підхід, що поєднує декларативні дані (анкети) та емпіричні дані (поведінка).

Джерела даних.

1. *Явні дані*. Включають: анкети (наприклад, визначення соціотипу), тести когнітивного стилю, уподобання користувача. Переваги явних даних у їх



інтерпретованості та швидкій ініціалізації моделі. До недоліків можна віднести суб'єктивність та статичність.

2. *Неявні дані*. Формуються на основі логування дій користувача: час перегляду відео, читання текстів, прослуховування аудіо, взаємодія із симуляціями. Переваги: відображають реальну поведінку, динамічно оновлюються. Недоліки: вимагають обробки, містять шум.

На основі зібраних даних формується агрегований вектор спостережень:

$$B = (t_v, t_a, t_t, t_k), \quad (10)$$

де  $t_j$  - сумарний час або інтенсивність взаємодії з контентом типу  $j$ .

Величини  $t_j$  можуть інтерпретуватися як час, так й як: кількість дій, інтенсивність взаємодії, частота вибору контенту.

Для переходу до порівняльного простору використовується нормалізація:

$$b_j = \frac{t_j}{\sum_j t_j}. \quad (11)$$

Результатом є нормований вектор:

$$B = (b_v, b_a, b_t, b_k). \quad (12)$$

Ця нормалізація забезпечує:

1. Перехід до імовірнісної інтерпретації. Вектор можна трактувати як розподіл переваг.

2. Інваріантність до тривалості сесії. Користувачі з різним часом активності стають порівнянними.

3. Стійкість до масштабування даних. Виключається вплив абсолютних величин.

Отриманий вектор  $B$  використовується для ініціалізації моделі

$$U^{(0)} = B. \quad (13)$$

або оновлення

$$U^{(t+1)} = (1 - \beta)U^{(t)} + \beta B^{(t)} \quad (14)$$

де:  $\beta$  - коефіцієнт навчання;  $B^{(t)}$  - спостереження на кроці  $t$ .

Як і вектор  $U$ , вектор спостережень  $B$ : належить симплексу  $\Delta^3$ ; відображає



положення користувача у просторі модальностей; може використовуватись для кластеризації.

Використання часу  $t_j$  як базової ознаки обумовлено тим, що час є універсальною мірою залученості, відбиває глибину взаємодії та стійке до випадкових кліків.

Практична інтерпретація в контексті дисципліни «Технології поліграфічного виробництва» :

- високий  $b_v \rightarrow$  студент віддає перевагу схемам і відео процесам друку;
- високий  $b_a \rightarrow$  віддає перевагу поясненням;
- високий  $b_k \rightarrow$  працює із симуляціями;
- високий  $b_t \rightarrow$  читає теоретичні матеріали.

Запропонований метод формування ознакового простору користувача заснований на когнітивно обґрунтованій моделі сприйняття, об'єднує суб'єктивні та об'єктивні дані, забезпечує математично коректне уявлення у вигляді ймовірнісного вектора, дозволяє динамічно адаптувати модель користувача.

Таким чином, вектор спостережень  $B$  є ключовим елементом переходу від статичної персоналізації до адаптивних мультимедійних систем, що враховує реальні поведінкові патерни користувача.

#### **2.4. Кластеризація користувачів**

Кластеризація користувачів є ключовим етапом у побудові моделі користувальницької модальності, оскільки дозволяє виявити стійкі групи користувачів з близькими когнітивними уподобаннями. Як базовий метод вибрано алгоритм k-середніх (k-means), що обумовлено його інтерпретованістю, обчислювальною ефективністю та відповідністю геометричній природі даних.

Кожен користувач описується вектором модальності:

$$x_i = (w_v, w_a, w_t, w_k), x_i \in R^4, \sum_j w_j = 1. \quad (15)$$

Таким чином, безліч користувачів:

$$X = \{x_1, x_2, \dots, x_n\}, x_i \in R^4 \quad (16)$$

розташоване в 4-мірному симплексі, що задає природний метричний простір для аналізу подібності.



Завдання кластеризації полягає в розбитті множини  $X$  на  $K$  кластерів:

$$C = \{C_1, C_2, \dots, C_K\}. \quad (17)$$

таким чином, щоб користувачі всередині одного кластера були максимально схожі, а між кластерами максимально різні.

Алгоритм  $k$ -середніх мінімізує внутрішньокластерну дисперсію:

$$J = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2, \quad (18)$$

де  $\mu_k$  - центроїд (середній вектор) кластера  $C_k$ ,  $\|\cdot\|$  - евклідова норма.

Мінімізація  $J$  означає зменшення розкиду користувачів усередині кластера та формування компактних та інтерпретованих груп.

З когнітивної точки зору це відповідає виділенню:

- "візуально-орієнтованих" користувачів,
- "аудіально-орієнтованих",
- змішаних типів.

Алгоритм кластеризації. Алгоритм  $k$ -means реалізується ітераційно:

Крок 1. Призначення кластерів. Кожен користувач  $x_i$  відноситься до найближчого центроїду:

$$C_k = \{x_i : \|x_i - \mu_k\| \leq \|x_i - \mu_j\|, \forall j\}. \quad (19)$$

Крок 2. Перерахунок центроїдів:

$$\mu_k = \frac{1}{|C_k|} \sum_{x_i \in C_k} x_i. \quad (20)$$

Крок 3. Повторення. Процес повторюється до збіжності: центроїди перестають змінюватися чи досягається максимум ітерацій.

Вибір алгоритму  $k$ -means обумовлений такими факторами:

1. Відповідність векторної моделі. Користувач вже представлений у вигляді вектора, що робить  $k$ -means природним вибором.
2. Геометрична інтерпретованість. Кожен кластер має центроїд, який можна трактувати як типовий профіль користувача.
3. Сумісність із адаптивною моделлю. Центроїди можуть використовуватися як початкові значення  $U$  та шаблони модальностей.

Практична інтерпретація. У дисципліні «Технології поліграфічного



виробництва» кластери можуть інтерпретуватися так:

- кластер 1 → візуали (схеми, відео друку)
- кластер 2 → аудіали (пояснення процесів)
- кластер 3 → текстові (інструкції)
- кластер 4 → кінестетики (симуляції)

Кластеризація користувачів методом k- means: забезпечує виділення стійких типів модальності, узгоджується з геометрією простору ознак, дозволяє перейти від індивідуальних моделей до узагальнених профілів, служить основою адаптивних мультимедійних систем.

### **2.5. Інтеграція моделі в мультимедійну базу даних**

Інтеграція моделі модальності користувача в мультимедійну базу даних є ключовим етапом переходу від традиційних систем зберігання до інтелектуальних адаптивних систем. У пропонованому підході модальність розглядається не як зовнішній параметр інтерфейсу, а як вбудований компонент моделі даних, що впливає на процеси пошуку, ранжирування та подання контенту.

Традиційні мультимедійні СУБД поділяють рівень зберігання даних, рівень обробки, рівень інтерфейсу користувача. У цьому персоналізація, зазвичай, реалізується лише на рівні додатків. Це призводить до архітектурної фрагментації та знижує ефективність адаптації.

У пропонованій моделі реалізується наскрізна інтеграція користувацьких характеристик, при якій користувальницька модель включається до схеми БД, механізми пошуку враховують модальність, адаптація відбувається на рівні ядра системи.

Задля реалізації цього підходу пропонується розширення схеми БД.

Таблиця користувачів:

$$User(id, w_v, w_a, w_t, w_k). \quad (21)$$

де  $w_j$  - ваги користувальницької модальності.

Таблиця контенту:

$$Content(id, f_v, f_a, f_t, f_k, embedding). \quad (22)$$



де  $f_j$  — модальні характеристики контенту, *embedding* — семантичний простір (наприклад, вектор ознак).

Запропонована структура реалізує двоплоскову модель даних:

1. Семантичний простір: відбиває зміст даних; реалізується через *embedding*.
2. Модальний простір: відбиває форму уявлення; визначається векторами  $f_j$  і  $w_j$ .

Такий поділ відповідає фундаментальній різниці між *що представлено* (семантика) і *як представлено* (модальність).

Для забезпечення ефективності пошуку використовується комбінована індексація:

1. Семантична індексація: векторні індекси (*FAISS*, *ANN*), пошук по *embedding*.
2. Модальна індексація: індексування за ознаками  $f_v, f_a, f_b, f_k$ ; фільтрація та ранжування за структурою контенту.

Комбінований скоринг. Під час виконання запиту використовується інтегральна функція:

$$R^*(q, d_i, U) = \alpha \cdot R_{semantic}(q, d_i) + (1 - \alpha) \cdot M(U, d_i), \quad (23)$$

де  $R_{semantic}$  - семантична релевантність;  $M(U, d_i) = \sum w_j f_j$  - модальна відповідність;  $\alpha \in [0, 1]$  - коефіцієнт балансу.

Інтерпретація моделі: семантична частина відповідає за зміст, модальна - за зручність сприйняття. Система шукає не просто релевантний контент, а найбільш вдалий спосіб його представлення для конкретного користувача.

Запропонована архітектура реалізує адаптацію на двох рівнях:

1. Адаптація на рівні пошуку: зміна ранжування результатів; облік користувальницької модальності під час видачі.

Приклад:

- візуалу → вище піднімаються відео та схеми,
- текстового користувача → статті.

2. Адаптація на рівні подання: зміна форми відображення; вибір формату



контенту.

Приклад:

- додавання озвучки,
- заміна тексту на схему,
- активація інтерактивних елементів

Таким чином, запропонований підхід формує основу нового класу мультимедійних баз даних — адаптивних, персоналізованих і когнітивно-орієнтованих систем.

## ***2.6. Інтерпретація моделі***

Запропонована модель має низку переваг: формалізує когнітивні особливості користувача; інтегрується у існуючі СУБД; підтримує машинне навчання; масштабується великі дані. Головна відмінність від традиційних підходів полягає в тому, що модальність користувача стає не зовнішнім параметром інтерфейсу, а внутрішнім елементом моделі даних, що впливає на всі етапи обробки мультимедійної інформації. Таким чином, розроблена модель користувальницької модальності забезпечує перехід від статичних мультимедійних баз даних до адаптивних систем, здатних враховувати когнітивні особливості користувача. Це створює підґрунтя для побудови інтелектуальних мультимедійних систем нового покоління.

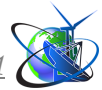
## **3. Постановка експерименту**

### ***3.1. Умови проведення експерименту***

Для перевірки висунутої гіпотези (H1) було проведено контрольований педагогічний експеримент у межах дисципліни «*Технології поліграфічного виробництва*». Експеримент реалізований за схемою із незалежними групами.

У дослідженні взяли участь 50 студентів першого курсу. Учасники були випадково поділені на дві групи:

- контрольна група (CG) – 25 студентів, які використовують мультимедійну базу даних без адаптації;
- експериментальна група (EG) - 25 студентів, які використовують мультимедійну базу даних з адаптацією на основі користувальницької



модальності.

Перед початком експерименту провели вхідний тест для підтвердження сумісності вихідного рівня знань в обох групах.

Навчальний матеріал охоплював основні теми дисципліни, включаючи технології друку (офсетна, цифрова); матеріали (папір, фарби); технологічні процеси; поліграфічне обладнання.

Контент було представлено у мультимедійній формі: текстові матеріали; візуальні схеми та ілюстрації; відеодемонстрації; інтерактивні елементи (симуляція).

В експериментальній системі реалізована інтеграція моделі користувальницької модальності. Кожен користувач описується вектором (1), який впливає як ранжування контенту, і форму його представлення.

Експеримент включав п'ять послідовних етапів:

1. Попереднє тестування: оцінка базового рівня знань студентів. Визначення користувальницької модальності: виконувалося з використанням анкетування, аналізу поведінки користувачів та кластеризації методом k- means.

2. Навчальний етап. Контрольна група одержувала однаковий контент без адаптації. Експериментальна група отримувала контент адаптований з урахуванням модальності користувача.

3. Виконання навчальних завдань. Студенти виконували: пошук інформації; вивчення матеріалів; практичні завдання.

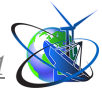
4. Підсумкове тестування та анкетування: контроль знань (post-test); оцінка задоволеності (шкала Лайкерта 1-5).

### **3.2. Метрики оцінювання**

Для комплексної оцінки ефективності запропонованої системи використовуються три групи метрик: інформаційного пошуку, освітні та поведінкові.

1. *Метрики інформаційного пошуку.* Ці метрики оцінюють, наскільки добре система знаходить та ранжує релевантний контент.

1.1. Метрика Precision@10 (точність на перших 10 результатах) знаходиться



за формулою:

$$Precision@10 = \frac{\text{Количество релевантных документов в топ-10}}{10}. \quad (24)$$

Дана метрика показує, яка частка з перших 10 результатів справді корисна користувачеві.

Наприклад. Студент шукає: "процес офсетного друку". Система видала 10 матеріалів: 8 – за темою; 2 - нерелевантні.

$$Precision@10 = \frac{8}{10} = 0.8 \quad (25)$$

Інтерпретація результатів: 0.8-1.0 → висока точність; 0.5–0.7 → середня точність; <0.5 → слабка точність. Під час навчання користувач рідко дивиться далі за перші результати і намагається прийняти рішення швидко. Тому Precision@10 – критично важлива метрика UX.

1.2. Метрика NDCG @10 (нормалізований дисконтований кумулятивний приріст). Це складніша метрика, яка враховує порядок результатів і рівень релевантності. Ця метрика розраховується у два етапи. На першому етапі визначаються числові значення кумулятивного приросту

$$DCG@10 = \sum_{i=1}^{10} \frac{rel_i}{\log_2(i+1)} \quad (26)$$

де:  $rel_i$  – релевантність документа (наприклад, 0-3);  $i$  – позиція.

На другому етапі проводиться нормалізація

$$NDCG@10 = \frac{DCG@10}{IDCG@10}. \quad (27)$$

де  $IDCG$  – ідеальний порядок.

Сенс метрики полягає в тому, що вона враховує, що верхні позиції важливіші за нижні, також розрізняє: "хороший результат на 1 місці" або "той же результат на 10 місці".

Наприклад, якщо найкорисніший матеріал: на 1 місці → високий NDCG; на 10 місці → низький NDCG

Інтерпретація цієї метрики:

- 0.8–1.0 → відмінне ранжування
- 0.6–0.8 → гарне
- <0.6 → слабке



Метрика Precision не враховує порядок результатів та ступінь релевантності, а NDCG – враховує. Для системи, що розробляється, це ключова метрика.

2. *Освітні метрики.* Ці метрики оцінюють реальний результат навчання, а не лише пошук.

### 2.1. Метрика Learning Gain (приріст знань).

$$Gain = Score_{post} - Score_{pre}. \quad (28)$$

Метрика показує, наскільки студент покращив знання після роботи із системою. Наприклад, при вивченні матеріалу до застосування навчальної системи студент набрав 40 балів, а при навчанні із застосуванням запропонованої системи – 70 балів. Тоді  $Gain = 70 - 40 = 30$ .

Інтерпретація цієї метрики така: високий Gain → система ефективно навчає; низький → контент не засвоюється.

Це найголовніша метрика для освітньої системи. Навіть якщо пошук хороший, але знань немає – система марна.

### 3. *Поведінкові метрики.* Оцінюють як користувач взаємодіє із системою.

3.1. Час виконання завдання (Time). Метрика дозволяє оцінити скільки часу студент витрачає на: пошук, вивчення та виконання завдання. Менше часу → система зручніша. Занадто мало → можливе поверхове навчання. Ця метрика враховує симбіоз ефективності навчання та когнітивного навантаження.

### 3.2. CTR (Коефіцієнт клікабельності).

$$CTR = \frac{\text{Кількість кліків}}{\text{Кількість показів}} \quad (29)$$

Коефіцієнт клікабельності показує, наскільки результати привабливі та релевантні. Наприклад: 10 результатів → 5 кліків.  $CTR = 0,5$ .

Інтерпретація цієї метрики: високий CTR → користувач "довіряє" системі; низький → результати нецікаві. Зростання CTR означає найкраще потрапляння під модальність користувача.

### 4. *Суб'єктивні метрики.*

4.1. Задоволеність користувача. Вимірюється за шкалою Лайкерта (1-5): 1 - дуже погано; 5 - відмінно. Ця метрика оцінює зручність, зрозумілість та комфорт. Навіть за хороших метриків користувач може не приймати систему. Це є



критично для впровадження.

Жодна метрика сама по собі недостатня. Тільки їх сукупність дозволяє оцінити систему як пошукову, навчальну та користувальницьку.

### 3.3. Статистичний аналіз

Для перевірки статистичної значущості відмінностей між контрольною та експериментальною групами використовувався  $t$ -критерій Стюдента для незалежних вибірок.

Формалізація критерію

$$t = \frac{\bar{x}_{EG} - \bar{x}_{CG}}{\sqrt{\frac{s_{EG}^2}{n_{EG}} + \frac{s_{CG}^2}{n_{CG}}}}, \quad (30)$$

де:  $\bar{x}_{EG}, \bar{x}_{CG}$  - Середні значення показників в експериментальній та контрольній групах;  $s_{EG}^2, s_{CG}^2$  - Вибіркові дисперсії;  $n_{EG}, n_{CG}$  - Обсяги вибірок

Поріг статистичної значимості прийнято рівним:  $p < 0.05$ .

Обґрунтування застосування  $t$ -критерію.

1. Тип дослідницького підходу. У роботі реалізовано експеримент із незалежними групами: контрольна група (без адаптації); експериментальна група (з адаптацією). Кожен учасник входить лише до однієї групи, що повністю відповідає умовам застосування  $t$ -критерію для незалежних вибірок.

2. Тип даних. Аналізовані показники (Precision, NDCG, Learning Gain, Time) є кількісними та вимірюваними в інтервальній шкалі. Це відповідає базовим вимогам застосування  $t$ -критерію.

3. Обсяг вибірки. Розмір вибірок у експерименті становить  $n_{EG} = n_{CG} = 25$  учасників. При таких значеннях виконується умова застосування критерію, а також завдяки центральній граничній теоремі розподіл середніх наближається до нормального.

4. Перевірена гіпотеза. Формулюються гіпотези:

Нульова гіпотеза ( $H_0$ ):  $\mu_{EG} = \mu_{CG}$

Альтернативна гіпотеза ( $H_1$ ):  $\mu_{EG} \neq \mu_{CG}$

$t$ -критерій дозволяє перевірити, чи є різницю між середніми статистично значущими, а чи не випадковими.



5. Інтерпретація результату. Якщо:  $p < 0.05$ , то: різницю між групами вважаються статистично значущими, нульова гіпотеза відкидається, підтверджується гіпотеза  $H_1$  про вплив модальної адаптації.

Вибір цього критерію обумовлений такими факторами:

1. Простота та інтерпретованість. Дозволяє безпосередньо порівняти середні значення двох груп.

2. Широке застосування у освітніх дослідженнях. Є стандартом для аналізу педагогічних експериментів.

3. Відповідність структурі даних. Незалежні вибірки та кількісні показники.

4. Достатня статистична потужність. При розмірі вибірки ( $n=25$ ) забезпечує надійні результати.

Таким чином, застосування  $t$ -критерію Стьюдента для незалежних вибірок є обґрунтованим, оскільки відповідає плану експерименту, адекватний типу даних, дозволяє коректно перевірити гіпотезу про вплив користувальницької модальності. Використання даного критерію забезпечує статистичну строгість та достовірність отриманих результатів.

#### 4. Результати

У таблиці 2 представлені результати порівняння контрольної та експериментальної груп.

**Таблиця 2 – Порівняння показників ефективності**

Метрика	Контрольна група	Експериментальна група	Приріст
Precision@10	0.63	0.79	+25.4%
NDCG@10	0.66	0.84	+27.3%
Приріст знань	12.5	18.9	+51.2%
Час виконання (хв)	18.2	14.1	-22.5%
CTR	0.43	0.58	+34.9%
Задоволеність	3.9	4.7	+20.5%

*Авторська розробка*



Результати  $t$ -тесту показали статистично значущі різницю між групами:

- Precision@10:  $p = 0.012$ .
- NDCG @ 10:  $p = 0.009$ .
- Приріст знань:  $p = 0.004$

У всіх випадках виконується умова ( $p < 0.05$ ), що підтверджує значущість отриманих результатів.

Результати експерименту демонструють, що облік модальності користувача в мультимедійних базах даних призводить до стійкого поліпшення всіх ключових показників. Найбільш значні покращення спостерігаються: у метриці NDCG - підвищення якості ранжування; у прирості знань - поліпшення засвоєння матеріалу; в CTR - підвищення залучення користувачів.

Отримані результати підтверджують гіпотезу H1: Адаптація мультимедійного контенту на основі користувальницької модальності статистично значуще підвищує релевантність видачі та ефективність навчання.

Аналіз вкладу окремих компонентів показав, що: найбільше впливає модально-орієнтоване ранжування; додаткове покращення досягається за рахунок динамічного оновлення профілю користувача.

Проведений експеримент підтверджує, що запропонована модель мультимедійної бази даних з урахуванням користувальницької модальності істотно перевершує традиційні підходи як за якістю пошуку, так і за освітніми результатами. Отримані дані демонструють перспективність інтеграції когнітивних характеристик користувача до архітектури мультимедійних інформаційних систем.

**Висновки.** У роботі розглянуто проблему адаптації мультимедійних баз даних до користувальницької модальності як одного з ключових факторів підвищення ефективності взаємодії користувача з інформаційною системою. Аналіз існуючих підходів показав, що сучасні мультимедійні системи переважно орієнтовані універсальні механізми обробки даних і враховують когнітивні особливості сприйняття інформації, що знижує релевантність видачі та ефективність навчання.



В рамках дослідження запропонована формальна модель користувальницької модальності у вигляді нормованого вектора переваг, що дозволяє представити когнітивні характеристики користувача математично коректної формі. Ця модель забезпечує перехід від дискретної класифікації користувачів до безперервного уявлення, що адекватно відображає реальні поведінкові патерни.

Розроблено підхід до подання мультимедійного контенту в тому самому ознаковому просторі, що і модель користувача, що дозволяє формалізувати завдання узгодження між користувачем та контентом. Це створює підґрунтя для побудови адаптивних алгоритмів ранжирування.

Запропоновано метод формування ознакового простору користувача, заснований на інтеграції явних та неявних даних, що забезпечує більш точний та динамічний опис переваг. Застосування алгоритму кластеризації (k-means) дозволило виявити стійкі групи користувачів з різними домінуючими модальностями та підвищити стійкість моделі до шуму.

Ключовим результатом роботи є розробка архітектури мультимедійної бази даних, в якій модальність користувача інтегрована безпосередньо в структуру даних і процеси обробки. Введено дворівневу модель адаптації: на рівні пошуку (ранжування); на рівні подання (форма подання контенту).

Таким чином, пропонований підхід забезпечує перехід від традиційних мультимедійних баз даних до інтелектуальних, адаптивних та когнітивно-орієнтованих систем.

Незважаючи на отримані результати, цей напрямок має значний потенціал для подальшого розвитку. Перспективне розширення моделі користувальницької модальності за рахунок урахування додаткових факторів, включаючи контекст використання (пристрій, середовище), рівень підготовки користувача, когнітивне навантаження.

Важливим напрямом є застосування методів глибокого навчання та мультимодальних трансформерів для автоматичного вилучення ознак та побудови більш точних моделей відповідності користувачеві та контенту. Це



дозволить зменшити залежність від ручного налаштування параметрів та підвищити масштабованість системи.

Додаткові дослідження можуть бути спрямовані на розробку ймовірнісних моделей (наприклад, на основі байєсівських підходів або розподілів Діріхле), що більш адекватно відображають природу розподілу модальностей у симплексному просторі.

Загалом, результати роботи підтверджують, що інтеграція користувальницької модальності до мультимедійних баз даних є перспективним напрямом розвитку інтелектуальних інформаційних систем і відкриває нові можливості для створення персоналізованих, ефективних та адаптивних цифрових середовищ.

## **Література**

1. Bernacki, Jarosław, Ida Błazejczyk, Agnieszka Indyka-Piasecka, Marek Kopel, Elzbieta Kukla, and Bogdan Trawinski. "Responsive web design: testing usability of mobile web applications." in Da Nang (ed.), 'Intelligent Information and Database Systems: 8th Asian Conference – ACIIDS 2016, pp. Springer Berlin Heidelberg, 257–269.
2. Blair MacIntyre and Trevor F. Smith. (2018) "Thoughts on the Future of WebXR and the Immersive Web." 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany: 338–342.
3. Calvary, Gaelle, Joelle Coutaz, David Thevenin, Quentin Limbourg, Laurent Bouillon, and Jean Vanderdonckt. (2003) "A Unifying Reference Framework for multi-target user interfaces." *Interacting with Computers* 15 (3): 289–308.
4. Calvary, Gaelle, Joelle Coutaz, David Thevenin, Quentin Limbourg, Nathalie Souchon, Laurent Bouillon, Murielle Florins, and Jean Vanderdonckt. (2002) "Plasticity of User Interfaces: A Revised Reference Framework." in Costin Pribeanu and Jean Vanderdonckt (eds), *Task Models and Diagrams for User Interface Design: Proceedings of the First International Workshop on Task Models and Diagrams for User Interface Design - TAMODIA 2002, 18-19 July 2002, Bucharest, Romania,*



INFOREC Publishing House Bucharest, Bucharest, Romania: 127–134.

5. Chow Kevin, Coyiuto Caitlin, Nguyen Cuong, and Yoon, Dongwook. (2019). “Challenges and Design Considerations for Multimodal Asyn-chronous Collaboration in VR.” *Proceedings of the ACM on Human-Computer Interaction* 3: 1–24.

6. Coutaz, J. and Caelen, J. (1991) “A Taxonomy for Multimedia and Multimodal User Interfaces.” *Proc. of the ERCIM Workshop on Multimedia Technology*, Lisbonne: ERCIM, November 1991.

7. Dumas, Bruno, Denis Lalanne, and Sharon Oviatt. (2009) “Multimodal Interfaces: A Survey of Principles, Models and Frameworks.” *Human Machine Interaction: Research Results of the MMI Program*, Springer Berlin Heidelberg, Berlin, Heidelberg: 3–26.

8. Gallud, Jose A., Ricardo Tesoriero, Jean Vanderdonckt, Maria Lozano, Victor Penichet, and Federico Botella. (2011) “Distributed user interfaces.” *CHI '11 Extended Abstracts on Human Factors in Computing Systems, CHI EA '11*, Association for Computing Machinery, New York, NY, USA: 2429–2432.

9. Geiger Christian, Paelke Volker, Reimann Christian and Rosenbach, W.. (2000). *A framework for the structured design of VR/AR content*: 75–82.

10. Ghiani, Giuseppe, Fabio Paterno, and Carmen Santoro. (2010) “On-demand cross-device interface components migration.” *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '10*, Association for Computing Machinery, New York, NY, USA: 299–308.

11. Gordyeyev, A., Trishch, R., Khomiak, E., Cherniak, O., Khomenko, V., Antonenko, N. (2026). *The Quality of Creating Multimedia Learning Systems Taking into Account the Cognitive Characteristics of Students*. In: Choudrie, J., Mahalle, P.N., Perumal, T., Joshi, A. (eds) *ICT for Intelligent Systems. ICTIS 2025. Lecture Notes in Networks and Systems*, vol 1509. Springer, Singapore. [https://doi.org/10.1007/978-981-96-8793-0\\_23](https://doi.org/10.1007/978-981-96-8793-0_23)

12. Grolaux, Donatien, Jean Vanderdonckt, and Peter Van Roy. (2005) “Attach Me, Detach Me, Assemble Me Like You Work.” In: Maria Francesca Costabile and



Fabio Paterno (eds), *Human-Computer Interaction - INTERACT 2005*, Springer Berlin Heidelberg, Berlin, Heidelberg: 198–212.

13. Grundy, J., and Hosking, J. (2000). “Developing adaptable user interfaces for component-based systems.” *Proceedings - 1st Australasian User Interface Conference, AUIC 2000*, IEEE, Institute of Electrical and Electronics Engineers, United States of America: 17–25.

14. Jakub Szwarc, Anna Rasmus, Joanna Koziara, and Jacek Matulewski. (2023) “What is supposed to help, actually hinders: about the impact of the GUI dynamics on the efficiency of gaze text entry.” *Procedia Computer Science* 225: 2292–2301.

15. Kim, Y., Kim, T., Shin, W. Y., & Kim, S. W. (2024, March). MONET: Modality-embracing graph convolutional network and target-aware attention for multimedia recommendation. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining* (pp. 332-340). <https://doi.org/10.48550/arXiv.2312.09511>

16. Kovachev, Dejan, Dominik Renzel, Petru Nicolaescu, Istvan Koren, and Ralf Klamka. (2014) “Direwolf: A framework for widget-based distributed user interfaces.” *Journal of Web Engineering* 13(07): 203–222.

17. Lacoche, Jeremy and Villain, Eric. (2022). “Evaluating Usability and User Experience of AR Applications in VR Simulation.” *Frontiers in Virtual Reality* 3: 881318.

18. Lin Pan, Zhiqiang Pan, Fei Cai, Honghui Chen, Multimodal recommender systems: A survey of representation, modeling, and optimization, *Information Fusion*, Volume 128, 2026, 103991, ISSN 1566-2535. <https://doi.org/10.1016/j.inffus.2025.103991>

19. Lin Pan, Zhiqiang Pan, Fei Cai, Honghui Chen, Multimodal recommender systems: A survey of representation, modeling, and optimization, *Information Fusion*, Volume 128, 2026, 103991, ISSN 1566-2535. <https://doi.org/10.1016/j.inffus.2025.103991> .

20. Liu, Q., Hu, J., Xiao, Y., Zhao, X., Gao, J., Wang, W., ... & Tang, J. (2024). Multimodal recommender systems: A survey. *ACM Computing Surveys* , 57 (2), 1-17.



<https://doi.org/10.48550/arXiv.2302.03883>

21. Lu, Y., Hu, Y., Shen, X., and Chen, Z. (2023). “An Immersive Layout Framework for Web Design in Virtual Reality.” *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*: 36.

22. Lu, Yijuan & Sebe, Nicu & Hytinen, Ross & Tian, Qi. (2011). Personalization in multimedia retrieval: A survey. *Multimedia Tools Appl.* 51. 247-277. 10.1007/s11042-010-0621-0.

23. Malitesta, D., Cornacchia, G., Pomo, C., Merra, F. A., Di Noia, T., & Di Sciascio, E. (2025). Formalizing multimedia recommendation through multimodal deep learning. *ACM Transactions on Recommender Systems*, 3 (3), 1-33. <https://doi.org/10.48550/arXiv.2309.05273>

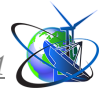
24. Martinez-Ruiz, Francisco, Jean Vanderdonckt, and Jaime Munoz-Arteaga. (2008) “Web User Interface Generation for Multiple Platforms.” *CEUR Workshop Proceedings* 445(1).

25. Melchior, Jeremie, Donatien Grolaux, Jean Vanderdonckt, and Peter Van Roy. (2009) “A toolkit for peer-to-peer distributed user interfaces: concepts, implementation, and applications.” *Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems, EICS '09*, Association for Computing Machinery, New York, NY, USA: 69–78.

26. Mendonca, Hildeberto, Jean-Yves Lionel Lawson, Olga Vybornova, Benoit Macq, and Jean Vanderdonckt. (2009) “A fusion framework for multimodal interactive applications.” *Proceedings of the 2009 International Conference on Multimodal Interfaces, ICMI-MLMI '09*, Association for Computing Machinery, New York, NY, USA: 161–168.

27. Motti, Vivian Genaro and Jean Vanderdonckt. (2013) “A computational framework for context-aware adaptation of user interfaces.” *IEEE 7th International Conference on Research Challenges in Information Science (RCIS)*, IEEE, Paris, France: 1–12.

28. Motti, Vivian Genaro, Dave Raggett, Sascha Van Cauwelaert, and Jean Vanderdonckt. (2013) “Simplifying the development of cross-platform web user



interfaces by collaborative model-based design.” Proceedings of the 31st ACM International Conference on Design of Communication, SIGDOC '13, Association for Computing Machinery, New York, NY, USA: 55–64.

29. Reeves, Leah, Jennifer Lai, James Larson, Sharon Oviatt, T. Balaji, Stephanie Buisine, Penny Collings, Philip Cohen, Ben Kraal, Jean-Claude Martin, Michael Mctear, T. Raman, Kay Stanney, Hui Su, and Qian Wang. (2004) “Guidelines for multimodal user interface design.” *Commun. ACM* 47(1):57–59.

30. Sanctorum, Audrey, and Beat Signer. (2019) “Towards End-User Development of Distributed User Interfaces.” *Universal Access in the Information Society (UAIS)* 18 (11).

31. Severi Peltonen, Luca Mezzalira, and Davide Taibi. (2021) “Motivations, benefits, and issues for adopting Micro-Frontends: A Multivocal Literature Review.” *Information and Software Technology* 136:0950–5849.

32. Sharp, H., Rogers, Y., and Preece, J. (2019). *Interaction Design, beyond Human-computer Interaction* (5th ed.). Wiley. Indianapolis, IN, USA. [26] Coutaz, J. (2007). “Meta-User Interfaces for Ambient Spaces.” in K. Coninx, K. Luyten, and K. A. Schneider (eds.), *Task Models and Diagrams for Users Interface Design* Springer Berlin Heidelberg. Berlin, Heidelberg: 1–15.

33. Smorgun, Ilya, David Lamas, and Eduardo Mercer. (2016) “Towards a Pattern Language for Distributed User Interfaces.” Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '16, Association for Computing Machinery, New York, NY, USA: 2712–2718.

34. Sottet, Jean-Sebastien, Gaelle Calvary, Joelle Coutaz, Jean-Marie Favre, Jean Vanderdonckt, Adrian Stanciulescu, and Sophie Lepreux. (2007) “A language perspective on the development of plastic multimodal user interfaces.” *Journal of Multimodal User Interfaces* 1 (6): 1–12.

35. Stanciulescu, Adrian and Jean Vanderdonckt. (2007) “Design Options for Multimodal Web Applications.” in Gaelle Calvary, Costin Pribeanu, Giuseppe Santucci, and Jean Vanderdonckt (eds), *Computer-Aided Design of User Interfaces V*, Springer Netherlands, Dordrecht: 41–56.



36. Stanciulescu, Adrian, Quentin Limbourg, Jean Vanderdonckt, Benjamin Michotte, and Francisco Montero. (2005) “A transformational approach for multimodal web user interfaces based on UsiXML.” Proceedings of the 7th International Conference on Multimodal Interfaces, ICMI '05, Association for Computing Machinery, New York, NY, USA: 259–266.

37. Stanton, M., Hartley, T., Loizides, F., and Worrallo, A. (2017). “Dual-Mode User Interfaces for Web Based Interactive 3D Virtual Environments Using Three.js.” in Proceedings of the International Conference on Theory and Practice of Digital Libraries: 441–444

38. Vanderdonckt, Jean, Gaelle Calvary, Joelle Coutaz, and Adrian Stanciulescu. (2008) “Multimodality for Plastic User Interfaces: Models, Methods, and Principles.” Multimodal User Interfaces: From Signals to Interaction, Springer Berlin Heidelberg, Berlin, Heidelberg: 61–84.

39. Vanderdonckt, Jean. (2010) “Distributed User Interfaces: How to Distribute User Interface Elements across Users, Platforms, and Environments.” Proc. of XIth Congreso Internacional de Interacción Persona-Ordenador Interacción'2010, XIth Congreso Internacional de Interacción Persona-Ordenador Interacción'2010, edited by J. L. Garrido, F. Paterno, J. Panach, K. Benghazi, and N. Aquino, AIPO, Valencia, September 7–10.

40. Vanderhulst, Geert, Daniel Schreiber, Kris Luyten, Max Muhlhauser, and Karin Coninx. (2009) “Edit, inspect and connect your surroundings: a reference framework for Meta-UIs.” Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems, EICS '09, Association for Computing Machinery, New York, NY, USA: 167–176.

41. Wu Jason, Amanda Swearngin, Xiaoyi Zhang, Jeffrey Nichols, and Jeffrey P. Bigham. (2023). “Screen correspondence: Mapping interchangeable elements between UIs.” in arXiv preprint arXiv:2301.08372

42. Wu, X., Huang, A., Yang, H., He, H., Tai, Y., & Zhang, W. (2024). Towards bridging the cross-modal semantic gap for multi-modal recommendation. arXiv preprint arXiv:2407.05420. <https://doi.org/10.48550/arXiv.2407.05420>



**Abstract.** *The article considers the problem of increasing the efficiency of multimedia databases by taking into account the cognitive characteristics of higher education students, in particular their modality of information perception. A formal model of user modality is proposed, presented in the form of a normalized vector of preferences, reflecting the degree of dominance of visual, auditory, textual and kinesthetic channels of perception. A method of presenting multimedia content in the same modal space is developed, which provides the possibility of quantitative comparison of the user and data. On this basis, an adaptive ranking mechanism is implemented, which takes into account both semantic relevance and the correspondence of the form of information presentation to the user's cognitive preferences. A pedagogical experiment was conducted with the participation of students, the results of which showed a statistically significant improvement in key metrics: Precision@10, NDCG@10, knowledge gain and user satisfaction. It has been established that taking into account user modality allows to increase the relevance of the output, reduce the cognitive load and increase the efficiency of the educational process. The obtained results confirm the prospects of integrating the cognitive characteristics of higher education students into the architecture of multimedia information systems and form the basis for the development of adaptive educational systems of the new generation.*

**Key words:** *multimedia databases; user modality; cognitive features; multimodal data; content ranking; educational information systems.*

Статтю надіслано: 06.04.2026 г.

© Гордєєв А.С.