

ФУНКЦІОНАЛЬНО-ОРІЄНТОВАНЕ ПРОЕКТУВАННЯ ІНФОРМАЦІЙНО-АНАЛІТИЧНИХ СИСТЕМ ОСВІТЬОГО СПРЯМУВАННЯ НА ОСНОВІ ПОПЕРЕДНЬО СФОРМОВАНИХ ПРОФІЛІВ

Одним із шляхів підвищення ефективності освітнього процесу на сьогоднішній день є розробка спеціалізованих програмних засобів, які дозволяють прискорити та спростити отримання найновішої фахової інформації у глобальному інформаційному просторі. В статті пропонується підхід до розробки інформаційно-аналітичних, інформаційно-пошукових, інформаційно-моніторингових та інших систем освітнього призначення, спроможних стати інструментом реалізації сучасної концепції креативного навчання на основі світового електронного інформаційного ресурсу. Запропонований метод проектування передбачає формалізоване компонування спеціалізованої програмної системи з уніфікованих програмних модулів згідно необхідної функціональності, яка визначається пошуково-аналітичними та іншими інформатичними задачами відповідних ланок навчального процесу.

One of the ways to improve educational process today is the development of specialized software that allow to speed up and simplify obtaining latest professional information from global information space. The paper proposes an approach to developing information-analytical, information-retrieval, information-monitoring and other systems for educational purposes, capable of becoming a tool of the modern creative learning concept based on world electronic information resources. The proposed design method provides formalized specialized software system layout using standardized software modules according to the required functionality, which is defined by search, analytical and other information tasks of the educational process appropriate parts.

Вступ

Важливим чинником розвитку освітніх технологій є використання інтелектуалізованих програмних засобів, спрямованих на підвищення ефективності навчання [1]. Одним з ключових напрямків застосування інтелектуальних комп'ютерних систем є забезпечення можливості оперативного отримання науковцями та освітянами потрібної фахової інформації. На сьогоднішній день основним джерелом актуальної інформації стає електронний інформаційний простір [2], тому шлях до розв'язання означеної проблеми лежить через розробку та впровадження різноманітних інформаційних, інформаційно-аналітичних, інформаційно-пошукових, інформаційно-моніторингових тощо систем.

З огляду на актуальність створення та широкого впровадження спеціалізованих систем, спрямованих на підтримку освітнього процесу (Education Intelligence систем [1]) постає задача підвищення ефективності їх розробки, зокрема на етапі виявлення потреб користувачів та формування технічних завдань. Однією з цілей цих етапів є визначення функціональності системи, що розробляється, та переліку компонентів, які цю функціональність забезпечать.

Розглянемо задачу формування структури інформаційних систем для підтримки освітнього процесу. Відомий підхід для рішення аналогічної задачі при розробці систем, що забезпечують певний клас захисту інформації та для систем підтримки державного управління, який використовує ідею попередньо сформованих профілів для систем певної спрямованості [3-6]. Він дозволяє одразу ж отримати перелік послуг, які має реалізовувати відповідна система.

В статті пропонується розвиток згаданого підходу щодо формування структури спеціалізованих інформаційних систем для підтримки навчального процесу. Запропоновано метод проектування, що базується на використанні певного набору компонент повнофункціональної інформаційно-аналітичної системи (ІАС), яка сама по собі забезпечує повний спектр можливостей щодо підтримки освітньої діяльності. Хоча реалізація повнофункціональної ІАС і спроможна вирішити будь-які задачі, що можуть ставитись перед інформаційно-аналітичною системою освітньої спрямованості, наслідком універсальності такої системи є її значна надлишковість у разі необхідності спеціалізованого використання в межах конкретного курсу, дисципліни, навчальної програми, методики навчання тощо.

Метод, що пропонується, дозволяє отримати мінімально достатній склад спеціалізованої системи, спрямованої на реалізацію необхідних функцій.

Повнофункціональна ІАС для підтримки навчального процесу

Функціональність засобів, спрямованих на ефективне отримання інформації, в цілому, відома. Skorистаємось наступною класифікацією типових функціональних послуг ІАС [7]:

- реалізація повного спектру пошукових операцій:

- повнотекстовий пошук,
- пошук за атрибутами,
- семантичний пошук,
- квазісемантичний пошук (із залученням тезаурусів, онтологій тощо),
- асоціативний пошук,
- ранжування пошукового відгуку (як за релевантністю, так і за пертинентністю),
- логічна компенсація дублювань у пошуковому відгуку;

- реалізація функцій змістовного аналізу текстових даних:

- структурний аналіз як структурованих, так і неструктурованих текстових інформаційних об'єктів,
- автоматична класифікація текстових інформаційних об'єктів за відповідними ознаками (тематика, змістова тональність, авторство та інші атрибути),
- автоматична кластеризація текстових інформаційних об'єктів,
- автоматизоване виділення атрибутів текстових інформаційних об'єктів (змістова

тональність, персоналії, асоціативні зв'язки з іншими інформаційними об'єктами тощо),

- автоматизоване анотування текстових інформаційних об'єктів,
- автоматизоване реферування текстових інформаційних об'єктів,
- попереднє виявлення логічних взаємозв'язків та залежностей на змістовному рівні, виявлення причинно-наслідкових логічних зв'язків тощо,
- оцінка рівня оригінальності тексту (виявлення плагіату, тавтології тощо),
- комп'ютерний переклад та інформаційна підтримка навчального перекладу (словники, засоби ідентифікації граматичних конструкцій, ідіоматичних зворотів тощо);
- спрямований моніторинг інформаційного ресурсу зокрема з метою оперативного виявлення оновлень;
- автоматичне виявлення орфографічних помилок у текстових даних;
- автоматична корекція орфографічних помилок у текстових даних;
- можливість логічного упорядкування та агрегування гетерогенного текстомісткого інформаційного ресурсу;
- компенсація на логічному рівні дублювань у текстомістких інформаційних об'єктах.

Якщо мова йде про створення системи, то, окрім функціональних засобів, які безпосередньо реалізують функції системи, також повинні бути засоби забезпечення, котрі здійснюють зберігання, накопичення, модифікацію, перегляд інформації (зокрема для реалізації вітрин даних), засоби адміністрування та інші.

Виходячи з вищезазначеного, можна запропонувати варіант складу повнофункціональної текстоорієнтованої інформаційно-аналітичної системи для підтримки навчального процесу (табл. 1). Текстоорієнтованість передбачає сфокусованість системи на роботі з текстовою інформацією, але, в цілому, можлива робота з довільними інформаційними об'єктами. В першому стовпчику таблиці подано ідентифікатор певного компонента системи, в другому – його назва. В третьому стовпчику знаходиться перелік компонентів, які необхідні для реалізації поточного компонента. Наприклад, при реалізації бази даних доцільно передбачити наявність засобів, які забезпечують взаємодію інших компонентів системи з цією базою даних.

Таблиця 1 – Складові повнофункціональної ІАС для підтримки навчального процесу

Позначення модуля	Назва компонента	Залежності
БД1	База даних	М1.1
БД1.1	Сегмент бази даних. Інформаційні об'єкти	БД1, М2.1
БД1.2	Сегмент бази даних. Атрибути	БД1, М2.2
БД1.3	Сегмент бази даних. Пошукові індекси	БД1, М2.3
БД1.4	Сегмент бази даних. Морфологічний словник	БД1
БД1.5	Сегмент бази даних. Індекси подібності	БД1, М2.4
БД1.6	Сегмент бази даних. Тезаурус/онтологія	БД1
БД1.7	Сегмент бази даних. Семантичні індекси	БД1, М2.5
БД1.8	Сегмент бази даних. Опис структури інформаційних об'єктів	БД1
БД1.9	Сегмент бази даних. Класифікатор	БД1
БД1.10	Сегмент бази даних. Словник лінгвістичних конструкцій	БД1
БД1.11	Сегмент бази даних. Моніторингові директиви	БД1, М7.1
БД1.12	Сегмент бази даних. Дані для забезпечення адміністрування	БД1, М8.1
М1.1	Засоби взаємодії з базою даних	БД1
М2.1	Засоби додавання інформаційних об'єктів	БД1.1, М1.1
М2.2	Засоби визначення атрибутів інформаційних об'єктів	БД1.2, М1.1, М2.1
М2.3	Засоби індексування для повнотекстового пошуку	БД1.3, М1.1, М2.1, М3.1
М2.4	Засоби індексування за подібністю	БД1.4, М1.1, М2.1, М3.1
М2.5	Засоби семантичного індексування	БД1.5, БД1.7, М1.1, М2.1, М3.1, М3.3
М3.1	Засоби лінгвістичного забезпечення (морфологія)	БД1.4, М1.1
М3.2	Засоби визначення подібності	БД1.5, М1.1
М3.3	Засоби лінгвістичного забезпечення (семантика)	БД1.6, БД1.7, М1.1
М4.1	Засоби атрибутивного пошуку	БД1.2, М1.1, М4.2
М4.2	Засоби ранжування пошукового відгуку	БД1.3, М1.1
М4.3	Засоби компенсації дублювань пошукового відгуку	М3.2, М4.2
М4.4	Засоби повнотекстового пошуку	БД1.3, М1.1, М3.1, М4.2
М4.5	Засоби асоціативного пошуку	М3.2, М4.2
М4.6	Засоби семантичного пошуку	М3.3, М4.2
М4.7	Редактор квазісемантичного запиту	БД1.6, М1.1
М4.8	Засоби квазісемантичного пошуку	БД1.3, М1.1, М4.2, М4.7
М5.1	Засоби перегляду інформаційних об'єктів	БД1.1, М1.1
М5.2	Засоби візуалізації	БД1, М1.1
М5.3	Засоби редагування інформаційних об'єктів	БД1.1, М1.1
М5.4	Засоби редагування атрибутів інформаційних об'єктів	БД1.2, М1.1
М6.1	Засоби структурного аналізу інформаційних об'єктів	БД1.1, БД1.8, М1.1
М6.2	Засоби класифікації інформаційних об'єктів	БД1.5, БД1.9, М1.1
М6.3	Засоби кластеризації інформаційних об'єктів	БД1.5, М1.1
М6.4	Засоби анотування/реферування інформаційних об'єктів	БД1.1, М1.1
М6.5	Засоби виявлення логічних зв'язків в інформаційних об'єктах	БД1.7, М1.1
М6.6	Засоби оцінки оригінальності інформаційних об'єктів	БД1.1, М1.1
М6.7	Засоби комп'ютерного перекладу	БД1.1, БД1.10, М1.1
М7.1	Засоби редагування моніторингових директив	БД1.11, М1.1
М7.2	Засоби сканування інформаційних ресурсів	БД1.11, М1.1
М7.3	Засоби оцінки відповідності інформаційних ресурсів моніторинговим директивам	БД1.11, М1.1, М7.2, М6.1
М7.4	Засоби виявлення орфографічних помилок в текстомістких інформаційних об'єктах	БД1.4, М1.1
М7.5	Засоби корекції орфографічних помилок в текстомістких інформаційних об'єктах	БД1.4, М1.1
М7.6	Засоби логічного упорядкування інформаційних об'єктів	БД1.1, М1.1
М7.7	Засоби статистичної обробки інформації	БД1, М1.1

Продовження Таблиці 1 – Складові повнофункціональної ІАС для підтримки навчального процесу

Позначення модуля	Назва компонента	Залежності
M8.1	Засоби редагування облікових записів користувачів	БД1.12, M1.1
M8.2	Засоби розмежування доступу	БД1.12, M1.1
M8.3	Засоби підтримки багатокористувацької роботи	БД1.12, M1.1, M8.2

Компоненти об'єднані в групи за логічним спрямуванням функцій, котрі вони виконують.

Першою такою групою є модулі, призначені для зберігання даних. Припускається, що в загальному випадку до складу ІАС входить одна чи більше баз даних (БД), в яких знаходяться дані, необхідні для функціонування системи. В залежності від функціональної орієнтованості системи необхідно передбачити відповідні сегменти бази даних.

Сегмент бази даних. Інформаційні об'єкти. Сегмент призначений для зберігання безпосередньо тих об'єктів, які обробляє система (наприклад, електронні версії рефератів, курсових робіт, методичних матеріалів, підручників, конспектів тощо).

Сегмент бази даних. Атрибути. Сегмент зберігає атрибути інформаційних об'єктів (наприклад, назву, авторів, рік створення, кількість сторінок тощо). Перелік атрибутів визначається типом інформаційного об'єкту та задачами, які вирішує ІАС.

Сегмент бази даних. Пошукові індекси. Зберігає інформацію, необхідну для ефективного здійснення повнотекстового пошуку інформаційних об'єктів.

Сегмент бази даних. Морфологічний словник. Цей сегмент необхідний для виконання морфологічного аналізу, наприклад, для визначення базової форми слова.

Сегмент бази даних. Індокси подібності. Зберігає дані індексування інформаційних об'єктів за подібністю, які можуть бути використані декількома аналітичними модулями, зокрема засобами класифікації та кластеризації [8,9].

Сегмент бази даних. Тезаурус/онтологія. Зберігає відповідно тезаурус або онтологію для підтримки аналітичних функцій, пов'язаних з аналізом змісту тексту.

Сегмент бази даних. Опис структури інформаційних об'єктів. Містить метадані, необхідні для структурного аналізу інформаційних об'єктів.

Сегмент бази даних. Класифікатор. Містить один або більше класифікаторів документів, з якими працює система.

Сегмент бази даних. Словник лінгвістичних конструкцій. Зберігає спеціальні дані, необхідні для синтаксичного та семантичного аналізу текстів.

Сегмент бази даних. Моніторингові директиви. Зберігає дані, що у певному форматі відображають інформаційну потребу користувачів щодо здійснення моніторингу інформаційних ресурсів.

Сегмент бази даних. Дані для забезпечення адміністрування. Містить службову інформацію, що необхідна для функціонування системи (перелік користувачів, їх персональні дані, перелік груп користувачів, доступи до окремих функцій системи тощо).

До другої групи входять засоби взаємодії з базою даних, які надають іншим модулям системи високорівневі функції щодо доступу до окремих сегментів бази даних. Засоби можуть бути реалізованими в складі СУБД або у вигляді окремого програмного модуля.

Оскільки інформація, яка міститься в системі, з великою ймовірністю має динамічний характер, потрібно передбачити засоби для додавання в систему нових даних або знань. Для цього передбачена третя група засобів, до якої входить наступне.

Засоби додавання інформаційних об'єктів. Заносять в систему інформаційні об'єкти. Для підтримки деяких функцій необхідно виконати відповідні процедури індексування (наприклад, індексування для пошукових засобів, індексування за подібністю до інших об'єктів тощо).

Засоби визначення атрибутів інформаційних об'єктів. Виконують визначення атрибутів та наповнення сегмента бази даних, який містить атрибути інформаційних об'єктів.

Засоби індексування для повнотекстового пошуку. Здійснюють процедуру індексування для створення пошукових індексів для засобів повнотекстового пошуку.

Засоби індексування за подібністю. Здійснюють індексування інформаційних об'єктів для реалізації деяких аналітичних функцій (наприклад, кластеризації або класифікації). Подібність об'єктів може визначатись за близькістю їх тематики або за іншими ознаками.

Засоби семантичного індексування. Формують індекси для фіксації змістовних зв'язків в тексті та між документами, які використовуються компонентами, що потребують аналізу змісту документів (таких, як семантичний пошук). Результати аналізу зберігаються в базі даних.

Система, що розглядається, передусім орієнтована на текстові дані, для повноцінної роботи з якими необхідна наявність відповідних засобів обробки текстів, котрі були об'єднані в четверту групу. До неї віднесемо наступне.

Засоби лінгвістичного забезпечення (морфологія). Реалізують сервісні функції для інших модулів, пов'язані з виконанням морфологічного аналізу текстів (наприклад, визначення базової словоформи в ході індексування для повнотекстового пошуку).

Засоби визначення подібності. Призначені для порівняння інформаційних об'єктів та визначення ступеню їх подібності за певним критерієм (тематична спрямованість, схожість значень певних атрибутів тощо).

Засоби лінгвістичного забезпечення (семантика). Реалізують допоміжні функції, необхідні аналітичним модулям, які мають справу зі змістом інформаційних об'єктів (наприклад, визначення взаємозв'язку між документами на семантичному рівні).

П'ята група засобів спрямована на реалізацію пошукових можливостей системи. До неї включено наступне.

Засоби атрибутивного пошуку. Здійснюють пошук інформаційних об'єктів за значеннями сукупності атрибутів.

Засоби ранжування пошукового відгуку. Здійснюють упорядкування інформаційних об'єктів, знайдених пошуковими засобами, за певними ознаками (ступінь релевантності, значення певних атрибутів тощо).

Засоби компенсації дублювань пошукового відгуку. Об'єднують інформаційні об'єкти пошукового відгуку в групи за визначеними критеріями (однакові об'єкти, об'єкти, подібність яких перевищує певний поріг, тощо).

Засоби повнотекстового пошуку. Здійснюють повнотекстовий пошук інформаційних об'єктів.

Засоби асоціативного пошуку. Здійснюють пошук інформаційних об'єктів не тільки за формальною відповідністю пошуковому запиту, а й з урахуванням наявності прямих, опосередкованих чи латентних смислових або логічних зв'язків між об'єктами чи то їх атрибутами (наприклад, можуть бути знайдені інформаційні об'єкти, де згадуються сутності, якимось чином між собою пов'язані).

Засоби семантичного пошуку. Здійснюють пошук інформаційних об'єктів за відповідністю пошуковому запиту на змістовному рівні.

Редактор квазісемантичного запиту. Реалізує формування запиту на квазісемантичний пошук шляхом розширення або звуження вихідного пошукового запиту з залученням тезаурусів або онтологій [10].

Засоби квазісемантичного пошуку. Реалізують пошук за запитом, сформованим редактором квазісемантичного запиту.

Наступна група засобів спрямована на подання інформації для кінцевого користувача (зокрема реалізують вітрини даних) та її модифікацію. До цієї групи віднесемо наступне.

Засоби перегляду інформаційних об'єктів. Об'єднують засоби, що надають можливість перегляду безпосередньо вмісту інформаційних об'єктів, їх атрибутів, зв'язків з іншими об'єктами тощо.

Засоби візуалізації. Служать для відображення нетекстової інформації (наприклад, результатів статистичного аналізу у вигляді таблиць, графіків, діаграм, схем тощо).

Засоби редагування інформаційних об'єктів. Надають користувачу можливість змінювати вміст інформаційних об'єктів, які зберігаються в системі.

Засоби редагування атрибутів інформаційних об'єктів. Здійснюють модифікацію окремих атрибутів інформаційних об'єктів.

Наступна група об'єднує засоби, спрямовані на підвищення ефективності отримання користувачем необхідної інформації за рахунок використання різноманітних аналітичних функцій. До них віднесемо наступне.

Засоби структурного аналізу інформаційних об'єктів. Призначені для визначення структурних елементів текстів та значень деяких атрибутів (наприклад, визначення назви статті, авторів, посилань, відокремлення розділів тощо).

Засоби класифікації інформаційних об'єктів. Здійснюють автоматичну класифікацію тексту, відносячи його до однієї або декількох гілок класифікатора.

Засоби кластеризації інформаційних об'єктів. Формують з множини інформаційних об'єктів групи, що є спорідненими, за визначеними ознаками (наприклад, за тематичною спрямованістю).

Засоби анотування/реферування інформаційних об'єктів. Формують текст, що є анотацією або рефератом до вхідного об'єкту.

Засоби виявлення логічних зв'язків в інформаційних об'єктах. Призначені для ідентифікації зв'язків між сутностями в текстових даних (наприклад, визначення персоналії, назви організації, в якій вона працює, посади тощо).

Засоби оцінки оригінальності інформаційних об'єктів. Визначають обсяг запозичень в текстах.

Засоби комп'ютерного перекладу. Виконують, відповідно, переклад текстів з однієї природної мови на інші.

До наступної групи віднесені додаткові аналітичні засоби, а саме.

Засоби редагування моніторингових директив. Реалізують формалізацію інформаційно-моніторингового інтересу користувачів (для моніторингових систем).

Засоби сканування інформаційних ресурсів. Здійснюють перегляд інформаційних ресурсів, заданих моніторинговими директивами, з встановленою періодичністю для виявлення даних, які відповідають моніторинговому інтересу користувачів.

Засоби оцінки відповідності інформаційних ресурсів моніторинговим директивам. Визначають релевантність даних інформаційних ресурсів моніторинговим директивам.

Засоби виявлення орфографічних помилок в інформаційних об'єктах та засоби корекції орфографічних помилок в інформаційних об'єктах. Можуть працювати на етапі додавання інформації в систему, на етапі індексування, на етапі здійснення повнотекстового пошуку тощо.

Засоби логічного упорядкування інформаційних об'єктів. Забезпечують формування динамічних каталогів за рахунок упорядкування об'єктів на логічному рівні за певними ознаками (наприклад, об'єднання об'єктів за значеннями певних атрибутів).

Засоби статистичної обробки інформації. Виконують статистичний аналіз як даних, що безпосередньо присутні в інформаційних об'єктах, так і метаданих, таких як індексні масиви, історія дій відповідних користувачів та інші.

Нарешті, до останньої групи віднесені засоби, спрямовані на забезпечення функціонування системи в цілому, та засоби адміністрування. До них віднесемо наступне.

Засоби редагування облікових записів користувачів. Надають адміністраторові системи можливість створення, модифікації, видалення тощо облікових записів користувачів для багатокористувацьких систем.

Засоби розмежування доступу. Реалізують логічне відокремлення сеансів роботи різних користувачів.

Засоби підтримки багатокористувацької роботи. Реалізують механізми підтримки паралельної роботи з системою багатьох користувачів.

Звісно, можна запропонувати альтернативні варіанти переліку компонентів, що можуть входити до повнофункціональної ІАС для підтримки навчання. Наприклад, для рішення окремих задач можуть бути введені нові або деталізовані компоненти, що вже існують.

Проектування структури спеціалізованої ІАС освітнього призначення

При створенні програмних систем, призначених для підвищення ефективності окремих ланок навчання, важливою задачею є визначення структури такої системи. В попередньому розділі наведено перелік засобів, які входять до повнофункціональної ІАС, спрямованої на всебічну інформаційну підтримку освітньої діяльності. Задача реалізації ІАС для підтримки конкретних форм навчальної діяльності, на рішення якої спрямований метод, що пропонується, полягає у визначенні мінімального переліку компонентів, які забезпечать повноцінну інформаційну та інструментальну підтримку відповідних форм навчальної роботи. Очевидно, даний перелік буде являти собою підмножину компонентів повнофункціональної ІАС.

З таблиці 1 видно, що повнофункціональна ІАС S може бути подана у вигляді двох множин: множини компонентів (позначимо її як M) та множини зв'язків між ними (позначимо цю множину як E), тобто $S=S(M,E)$. Таке подання можна розглядати як граф [11], в якому M – множина вершин, а E – множина ребер. Елементи E – це двохелементні підмножини множини M . Якщо в E існує елемент (m_1, m_2) , це означає, що компонент m_1 залежить від m_2 , тобто для реалізації компонента m_1 необхідна також реалізація компонента m_2 . Очевидно, що елементи E є впорядкованими, тобто S – орієнтований граф.

Граф S може бути заданий таблицею 1. Елементи множини вершин M перелічені в першій колонці. Перша та третя колонки таблиці формують перелік дуг E , початкам яких відповідає компонент в першому стовпчику.

До спеціалізованої ІАС повинні увійти лише ті компоненти, які необхідні для реалізації функцій, що відповідають її спеціалізації, тому формування структури такої ІАС зводиться до задачі формування графа $S'(M',E')$, що буде підграфом S . Даний підграф буде мати наступну властивість:

$$\forall m \in M' \Rightarrow T^+(m) \subset M' \quad (1)$$

Це означає, що якщо до M' входить деякий компонент m_i , то до M' також входять всі компоненти, від яких залежить m_i , і які повинні бути розроблені для реалізації m_i , тобто разом з будь-якою вершиною до M' також входить її пряме транзитивне замкнення.

Множина дуг E' визначається наступним чином:

$$(m_i, m_j) \in E': m_j \in M' \wedge m_i \in M' \wedge (m_i, m_j) \in E$$

Вхідними даними для формування S' є відповідний функціональний профіль [5,6], котрий визначає перелік послуг, що їх має надавати спеціалізована ІАС. На його основі формується початковий перелік компонентів M'_1 . Множина M' буде включати в себе M'_1 , розширену вершинами, що відповідають умові (1).

Для системи, котра розробляється, може існувати стандартний функціональний профіль [5], який дозволяє отримати перелік послуг, що його має реалізувати система.

Кожній з цих послуг можна поставити у відповідність компонент з множини M , що дозволить сформувати множину M'_1 . У випадку, коли стандартного профілю, що відповідає системі, не існує, необхідно сформувати індивідуальний функціональний профіль, який дозволить сформувати M'_1 .

Для формування M'_1 можна скористатися таблицею 2, за допомогою якої формується характеристична функція $\chi_{M'_1}$ [12]. В таблиці 2 наведено множину логічних тверджень $T(t_1, t_2, \dots, t_{|T|})$. Кожне з тверджень t_i має значення 1 або 0 в залежності від вимог до конкретної системи, що розробляється. Також кожному твердженню t_i відповідає певний перелік компонентів, які мають бути включеними до складу системи у разі, якщо твердження вірне. Цей перелік можна розглядати як підмножину M , що має характеристичну функцію χ_{t_i} .

Наприклад, χ_{t_1} можна записати у вигляді

$$\chi_{t_1}(m) = \begin{cases} 1, m = M \text{ 2.1} \\ 0, m \neq M \text{ 2.1} \end{cases}$$

У цьому разі $\chi_{M'_1}$ визначається наступним чином:

$$\chi_{M'_1}(m_j) = \sum_{i=1}^{|T|} t_i \wedge \chi_{t_i}(m_j) \quad (2)$$

Узагальнений алгоритм формування системи S' можна подати як послідовність трьох основних кроків:

1) Визначення початкового переліку компонентів M'_1 з таблиці 1, спрямованих на реалізацію відповідних функцій. M'_1 формується на основі профіля системи та з використанням таблиці 2 і виразу (2).

2) Визначення переліку засобів M' шляхом розширення початкового переліку M'_1 за рахунок вершин з його прямого транзитивного замкнення.

$$M' = M'_{nom} \cup T^+(M'_{nom})$$

Таким чином до переліку засобів, які необхідно реалізувати у складі системи, включаються ті компоненти, від яких залежать засоби, які вже включені в початковий перелік.

Для визначення прямого транзитивного замкнення можна скористатися відомими алгоритмами [11,13,14].

Таблиця 2 – Логічні твердження для визначення початкового переліку компонентів спеціалізованої ІАС

Позначення твердження	Логічне твердження	Компоненти, які відповідають твердженню
Засоби для додавання інформаційних об'єктів		
t_1	Інформаційна складова системи є динамічною і передбачається введення нової інформації	M2.1
t_2	Необхідне автоматичне визначення атрибутів інформаційних об'єктів, що додаються	M2.2
Засоби пошуку		
t_3	Необхідна реалізація атрибутивного пошуку	M4.1
t_4	Необхідна реалізація повнотекстового пошуку за фразою	M4.4
t_5	Необхідна реалізація асоціативного пошуку	M4.5
t_6	Необхідна реалізація семантичного пошуку	M4.6
t_7	Необхідна реалізація квазісемантичного пошуку	M4.8
Засоби перегляду та модифікації інформаційних об'єктів		
t_8	Необхідна реалізація перегляду вмісту інформаційних об'єктів	M5.1
t_9	Необхідна реалізація перегляду нетекстових даних	M5.2
t_{10}	Необхідно виконувати модифікацію інформаційних об'єктів	M5.3
t_{11}	Необхідно редагувати значення атрибутів інформаційних об'єктів	M5.4
Засоби аналітики		
t_{12}	Необхідно реалізовувати класифікацію об'єктів	M6.2
t_{13}	Необхідно виконувати кластеризацію об'єктів	M6.3
t_{14}	Необхідно автоматично формувати анотації та/або реферати інформаційних об'єктів	M6.4
t_{15}	Необхідно виявляти фактографічну інформацію в текстах	M6.5
t_{16}	Необхідно оцінювати оригінальність інформаційних об'єктів	M6.6
t_{17}	Необхідний комп'ютерний переклад текстомістких об'єктів	M6.7
t_{18}	Необхідно здійснювати моніторинг певних інформаційних ресурсів	M7.2
t_{19}	Необхідне виявлення орфографічних помилок в інформаційних об'єктах	M7.4
t_{20}	Необхідна корекція орфографічних помилок в інформаційних об'єктах	M7.5
t_{21}	Необхідна побудова динамічних каталогів за значеннями певних атрибутів	M7.6
t_{22}	Необхідно виконувати статистичний аналіз певних даних, з якими має справу система?	M7.7, M5.2
Засоби забезпечення роботи системи		
t_{23}	Необхідне обмеження доступу до системи в цілому або до певних її функцій та/або ресурсів	M8.2
t_{24}	Система має бути багатокористувацькою	M8.1, M8.3

Наприклад, алгоритм визначення прямого транзитивного замкнення M' початкового переліку компонентів M'_1 може бути наступним [13]:

2а) Визначимо поточне значення M' як

$$M'_{nom} := M'_1$$

2б) До переліку засобів, які необхідно реалізувати, включаються ті компоненти, від яких безпосередньо залежать засоби, що вже ввійшли у перелік.

$$M'_{nom} := M'_{nom} \cup \bigcup_{i,j} (m_j : \exists(m_i, m_j) \in E \wedge$$

$$\wedge m_i \in M'_{nom} \wedge m_j \notin M'_{nom})$$

Даний крок виконується поки є компоненти, які мають бути включені в перелік. Якщо на черговій ітерації до переліку не включено жодного компонента, відбувається перехід на наступний крок алгоритму.

2в) Результуючим переліком компонентів буде

$$M' := M'_{nom}$$

3) Визначення множини дуг E' , яка буде включати в себе дуги з множини дуг E між вершинами, які входять до M' :

$$E' := \bigcup_{i,j} ((m_i, m_j) : (m_i, m_j) \in E \wedge \\ \wedge m_i \in M'_{nom} \wedge m_j \in M'_{nom})$$

Результатом роботи алгоритму стане підграф $S'(M', E')$, який відповідає структурі спеціалізованої ІАС.

Приклад 1. Формування структури спеціалізованої інформаційно-пошукової системи.

Розглянемо наступний приклад. Нехай необхідно розробити інформаційно-пошукову систему з наступними базовими можливостями:

- інформація в систему додається в ручному, автоматизованому та автоматичному режимах;
- система має надавати користувачеві можливість перегляду необхідної йому інформації;
- система має надавати можливість здійснювати повнотекстовий, атрибутивний, асоціативний, квазісемантичний та семантичний пошук;
- знайдені однакові інформаційні об'єкти у пошуковому відгуку мають об'єднуватись в групи;
- система має надавати можливість корекції орфографічних помилок в текстових даних, що вводяться до інформаційної бази системи та в рядок пошуку.

Перерахований перелік послуг є, по суті, профілем спеціалізованої системи. Для формування переліку компонентів, що мають входити до складу системи, та зв'язків між ними, скористаємось наведеним вище алгоритмом.

Крок 1

Сформуємо початковий перелік необхідних компонентів M'_1 .

Для введення інформації використаємо M2.1.

Для перегляду інформації необхідні M5.1.

Для надання відповідних режимів пошуку необхідні M4.1, M4.4, M4.5, M4.6, M4.8.

Для відображення результатів пошуку необхідні M4.2.

Для об'єднання однакових знайдених об'єктів у групи необхідні M4.3.

Для корекції орфографічних помилок необхідні M7.5.

Відповідна характеристична функція χ_{M1} наведена в таблиці 3.

Таким чином, в результаті виконання цього кроку отримаємо початковий перелік компонентів

$M'_1 = \{M2.1, M5.1, M4.1, M4.4, M4.5, M4.6, M4.8, M4.2, M4.3, M7.5\}$;

Перейдемо до наступного кроку алгоритму.

Крок 2а.

Встановлюємо $M'_{nom} = M'_1$

Крок 2б. Ітерація 1.

Згідно таблиці 1 для реалізації M2.1 та M5.1 необхідна наявність BD1.1 та M1.1 (для зберігання основної інформації для пошуку).

Для реалізації M4.1 необхідна наявність BD1.2, M1.1, M4.2 (для виконання атрибутивного пошуку необхідна наявність сегмента бази даних, в якому зберігаються значення атрибутів; для відображення результатів пошуку необхідні засоби ранжування).

Для реалізації M4.4 необхідна наявність BD1.3, M1.1, M3.1, M4.2 (повнотекстовий пошук здійснюється з залученням індексної інформації, яка зберігається у відповідному сегменті BD; для виконання пошуку необхідно використовувати засоби, які забезпечують морфологічний аналіз пошукового запиту; для упорядкування результатів пошуку необхідні засоби ранжування).

Для реалізації M4.5 необхідна наявність M3.2 та M4.2 (асоціативний пошук здійснюється з використанням визначення подібності інформаційних об'єктів; для упорядкування результатів пошуку необхідні засоби ранжування).

Для реалізації M4.6 необхідна наявність M3.3 та M4.2 (для реалізації семантичного пошуку необхідний відповідний аналіз інформаційних об'єктів та пошукового запиту; для упорядкування результатів пошуку необхідні засоби ранжування).

Для реалізації M4.8 необхідна наявність BD1.3, M1.1, M4.2 та M4.7 (пошук виконується з використанням пошукових індексів; для упорядкування результатів пошуку необхідні засоби ранжування; для виконання квазісемантичного пошуку необхідний відповідний редактор запиту).

Для реалізації M4.2 необхідна наявність BD1.3 та M1.1 (при ранжуванні звичайно використовується індексна інформація).

Таблиця 3. Значення характеристичної функції χ_{M1} для прикладу розробки інформаційно-пошукової системи

Компонент	Значення χ_{M1}	Компонент	Значення χ_{M1}	Компонент	Значення χ_{M1}
БД1	0	M2.4	0	M6.1	0
БД1.1	0	M2.5	0	M6.2	0
БД1.2	0	M3.1	0	M6.3	0
БД1.3	0	M3.2	0	M6.4	0
БД1.4	0	M3.3	0	M6.5	0
БД1.5	0	M4.1	1	M6.6	0
БД1.6	0	M4.2	1	M6.7	0
БД1.7	0	M4.3	1	M7.1	0
БД1.8	0	M4.4	1	M7.2	0
БД1.9	0	M4.5	1	M7.3	0
БД1.10	0	M4.6	1	M7.4	0
БД1.11	0	M4.7	0	M7.5	1
БД1.12	0	M4.8	1	M7.6	0
M1.1	0	M5.1	1	M7.7	0
M2.1	1	M5.2	0	M8.1	0
M2.2	0	M5.3	0	M8.2	0
M2.3	0	M5.4	0	M8.3	0

Для реалізації M4.3 необхідна наявність M3.2 та M4.2 (для виявлення дублювань необхідний аналіз подібності; виявлені дублікати повинні певним чином відобразитися в результатах пошуку).

Для реалізації M7.5 необхідна наявність БД1.4 та M1.1 (засоби орфокоорекції потребують наявності спеціального словника).

Таким чином, після виконання першої ітерації отримаємо наступну множину компонентів:

$M'_{nom} = \{M2.1, M5.1, M4.1, M4.4, M4.5, M4.6, M4.8, M4.2, M4.3, M7.5, БД1.1, M1.1, БД1.2, БД1.3, M3.1, M3.2, M3.3, M4.7, БД1.4\}$;

Крок 2б. Ітерація 2.

Для реалізації БД1.1 необхідна наявність БД1 та M2.1 (необхідні засоби для наповнення відповідного сегменту БД).

Для реалізації M1.1 необхідна наявність БД1 (вже включена).

Для реалізації БД1.2 необхідна наявність БД1 та M2.2 (необхідні засоби для наповнення відповідного сегменту БД).

Для реалізації БД1.3 необхідна наявність БД1 та M2.3 (необхідні засоби для наповнення відповідного сегменту БД).

Для реалізації M3.1 необхідна наявність БД1.4 та M1.1 (вже включені).

Для реалізації M3.2 необхідна наявність БД1.5 та M1.1 (для визначення подібності використовуються відповідні індекси).

Для реалізації M3.3 необхідна наявність БД1.6, БД1.7 та M1.1 (для семантичної підтримки необхідний тезаурус або онтологія, потрібне використання відповідних індексів).

Для реалізації M4.7 необхідна наявність БД1.6 та M1.1 (вже включені).

Для реалізації БД1.4 необхідна наявність БД1 (вже включена).

Результат другої ітерації наступний:

$M'_{nom} = \{M2.1, M5.1, M4.1, M4.4, M4.5, M4.6, M4.8, M4.2, M4.3, M7.5, БД1.1, M1.1, БД1.2, БД1.3, M3.1, M3.2, M3.3, M4.7, БД1.4, БД1, M2.2, M2.3, БД1.5, БД1.6, БД1.7\}$

Крок 2б. Ітерація 3.

Для реалізації БД1 необхідна наявність M1.1 (вже включені).

Для реалізації M2.2 необхідна наявність БД1.2, M1.1 та M2.1 (вже включені).

Для реалізації M2.3 необхідна наявність БД1.3, M1.1, M2.1 та M3.1 (вже включені).

Для реалізації БД1.5 необхідна наявність БД1 (вже включена) та M2.4 (для наповнення відповідного сегменту БД).

Для реалізації БД1.6 необхідна наявність БД1 (вже включена).

Для реалізації БД1.7 необхідна наявність БД1 (вже включена) та M2.5 (для наповнення відповідного сегменту БД).

Результат цієї ітерації наступний:

$M'_{nom} = \{M2.1, M5.1, M4.1, M4.4, M4.5, M4.6, M4.8, M4.2, M4.3, M7.5, BD1.1, M1.1, BD1.2, BD1.3, M3.1, M3.2, M3.3, M4.7, BD1.4, BD1, M2.2, M2.3, BD1.5, BD1.6, BD1.7, M2.4, M2.5\}$.

Крок 2б. Ітерація 4.

Для реалізації M2.4 необхідна наявність BD1.4, M1.1, M2.1 та M3.1 (вже включені).

Для реалізації M2.5 необхідна наявність BD1.5, M1.1, M2.1, M3.1 та M3.3 (вже включені).

Крок 2в.

Оскільки перелік засобів, отриманий на попередній ітерації, не змінився, то він і є результатом, тому $M \leftarrow M'_{nom} = \{M2.1, M5.1, M4.1, M4.4, M4.5, M4.6, M4.8, M4.2, M4.3, M7.5, BD1.1, M1.1, BD1.2, BD1.3, M3.1, M3.2, M3.3, M4.7, BD1.4, BD1, M2.2, M2.3, BD1.5, BD1.6, BD1.7, M2.4, M2.5\}$.

Крок 3.

E' буде підмножиною E , яку легко отримати з таблиці 1, вибравши ті її рядки, перший елемент яких відповідає елементам M' .

Результат роботи алгоритму – граф $S'(M', E')$, який відповідає структурі інформаційно-пошукової системи. Наведемо графічне зображення S' (рис. 1). Матриця суміжності S' наведена в таблиці 4.

Приклад 2. Формування структури спеціалізованої інформаційно-моніторингової системи.

Нехай необхідно розробити інформаційно-моніторингову систему, функціональність якої передбачає наявність наступних послуг:

- користувач повинен мати змогу формулювати свій моніторинговий інтерес (у вигляді переліку сайтів, що підлягають моніторингу, періодичності моніторингу та вказавши тематичну направленість інформації, яка його цікавить);

- система повинна з заданою періодичністю сканувати ресурси, які були вказані користувачем, та перевіряти їх оновлення на наявність необхідної інформації;

- у разі знаходження інформації, яка відповідає моніторинговому інтересу, зацікавлені користувачі мають отримати відповідні повідомлення;

- система має зберігати результати моніторингу, забезпечувати можливість їх перег-

ляду та повнотекстового пошуку в масивах результатів моніторингу;

- система має працювати в мережному середовищі та бути придатною для використання багатьма користувачами.

Крок 1.

Виходячи з послуг системи та використавши таблицю 2 сформуємо початковий перелік необхідних засобів:

$M'_j = \{M7.1, M7.2, M7.3, M5.2, BD1.1, BD1.2, M5.1, M4.4, M8.3\}$.

Крок 2а, 2б.

Процес виконання кроків алгоритму подамо у вигляді таблиці 5.

Крок 2в.

Оскільки в результаті виконання третьої ітерації перелік компонентів M' пот не змінився, він і є результатом:

$M \leftarrow M'_{nom} = \{M7.1, M7.2, M7.3, M5.2, BD1.1, BD1.2, M5.1, M4.4, M8.3, BD1.11, M1.1, M6.1, BD1, M2.1, M2.2, BD1.3, M3.1, M4.2, BD1.12, M8.2, M2.3, BD1.4, M8.1, BD1.8\}$.

Крок 3.

За даним переліком неважко сформувати множини зв'язків E' використовуючи таблицю 1. Результуючий граф S' відповідає структурі інформаційно-моніторингової системи, яку зобразимо в графічному вигляді (рис. 2). Матриця суміжності S' наведена в таблиці 6.

Висновки

В статті запропоновано метод формування структури спеціалізованих аналітичних програмних систем для інформаційної підтримки певних ланок навчальної діяльності, який базується на виділенні відповідної підмножини компонентів з повнофункціональної ІАС. Розглянуто варіант повнофункціональної ІАС, спрямованої на всебічну інформаційну підтримку сучасного навчального процесу.

Розроблений метод дозволяє формувати структури спеціалізованих ІАС, що розглядаються як деякі підмножини повнофункціональної ІАС. Виходячи з вимог до спеціалізованої ІАС обирається стандартний або створюється індивідуальний функціональний профіль, який включає множини сервісів, що має забезпечувати ІАС. На основі профілю формується початковий перелік компонентів повнофункціональної ІАС, які відповідають спеціалізації системи та мають обов'язково входити до її складу.

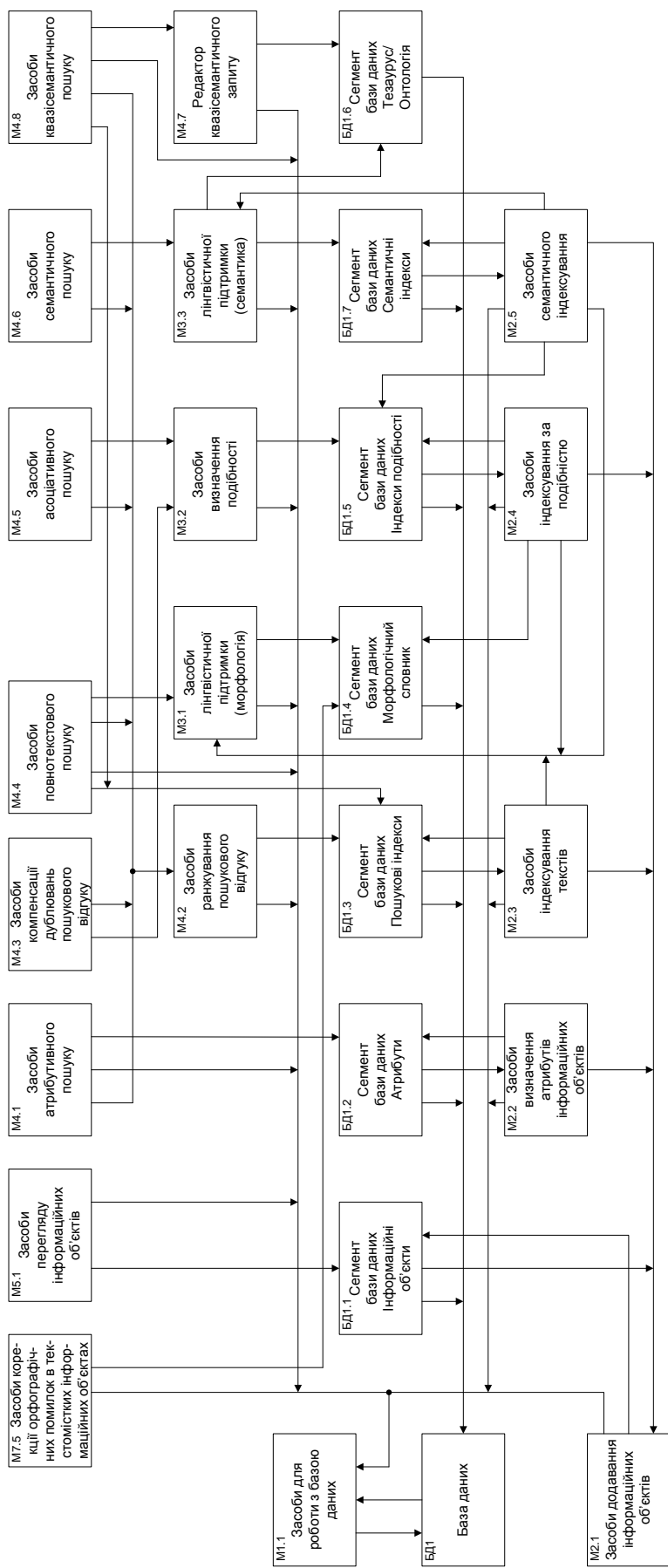


Рис. 1 Приклад структури інформаційно-пошукової системи

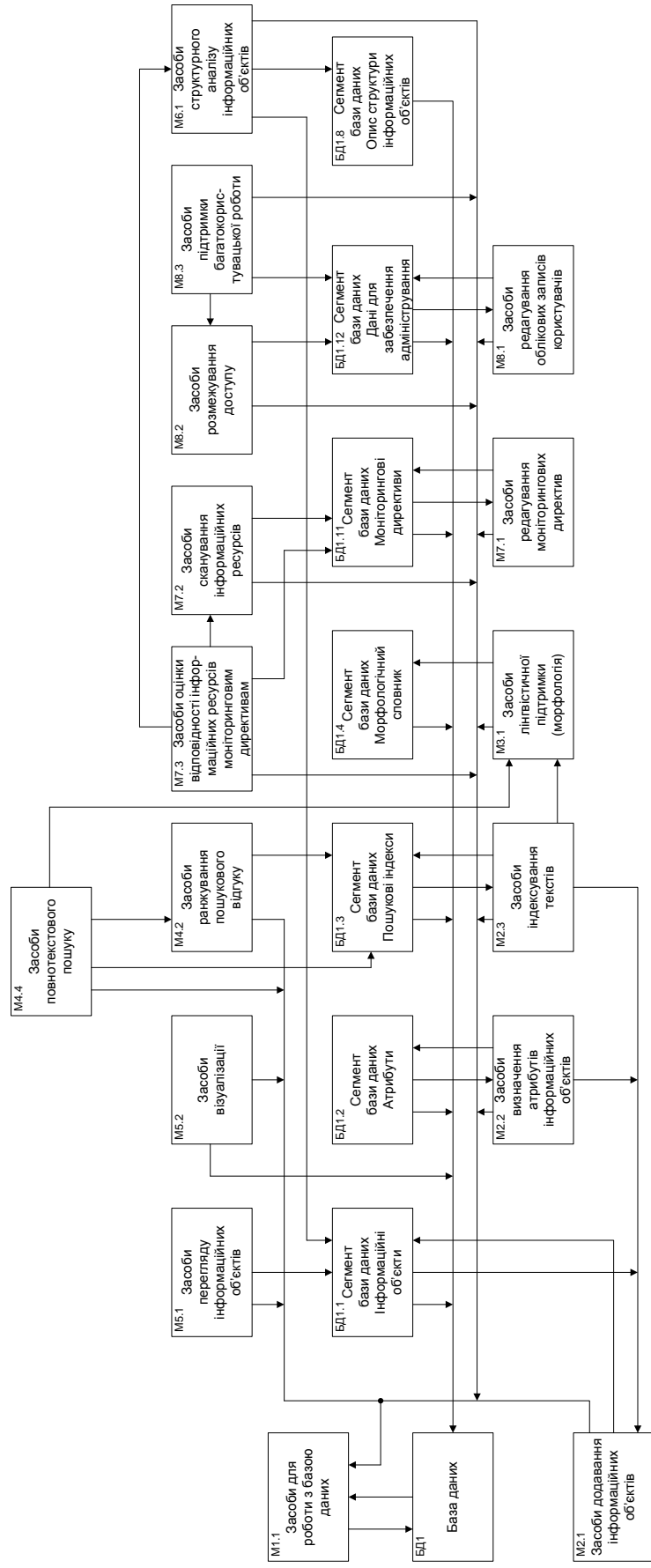


Рис.2 Приклад структури інформаційно-моніторингової системи

Таблиця 5. Зміна поточного переліку компонентів при формуванні структури моніторингової системи

№ ітерації	Результат ітерації
0	$M'_{nom}=M'_j=\{M7.1, M7.2, M7.3, M5.2, БД1.1, БД1.2, M5.1, M4.4, M8.3\}$
1	$M'_{nom}=\{M7.1, M7.2, M7.3, M5.2, БД1.1, БД1.2, M5.1, M4.4, M8.3, БД1.11, M1.1, M6.1, БД1, M2.1, M2.2, БД1.3, M3.1, M4.2, БД1.12, M8.2\}$
2	$M'_{nom}=\{M7.1, M7.2, M7.3, M5.2, БД1.1, БД1.2, M5.1, M4.4, M8.3, БД1.11, M1.1, M6.1, БД1, M2.1, M2.2, БД1.3, M3.1, M4.2, БД1.12, M8.2, M2.3, БД1.4, M8.1, БД1.8\}$
3	$M'_{nom}=\{M7.1, M7.2, M7.3, M5.2, БД1.1, БД1.2, M5.1, M4.4, M8.3, БД1.11, M1.1, M6.1, БД1, M2.1, M2.2, БД1.3, M3.1, M4.2, БД1.12, M8.2, M2.3, БД1.4, M8.1, БД1.8\}$

Для отримання структури повноцінної системи виконується розширення початкового переліку її елементів за рахунок компонентів, від яких залежать ті компоненти, які вже було включено в перелік. Отримана множина компонентів разом з множиною зв'язків між ними описують структуру ІАС необхідної спеціалізації.

Використання наведеного методу дозволяє суттєво прискорити процес створення або ІАС або її модифікації. Метод орієнтований, перш за все, на роботу з системами інформаційної підтримки навчального процесу, хоча його можна застосувати і для ІАС іншого спрямування, наприклад, для підтримки наукової діяльності.

СПИСОК ЛІТЕРАТУРИ

1. В. Гришачов, Д. Замятін, О. Кебало, А. Михайлюк, Л. Огнівчук, В. Тарасенко. Програмний інструментарій Education Intelligence як засіб підвищення ефективності креативного навчання // Міжнародний науковий журнал "Комп'ютеринг". Т. 10., Випуск 2. – Тернопіль. – 2011. – с. 114-132.
2. Ланде Д.В. Інформаційні потоки в глобальних комп'ютерних мережах: монографія / Д.В. Ланде. – К.: Наукова думка, 2009, - 295 с.
3. НД ТЗІ 2.2-001-98. Критерії оцінки захищеності інформації в комп'ютерних системах від несанкціонованого доступу / ДСТСЗІ СБ України. - Введ. 1998.
4. НД ТЗІ 2.2-002-98. Класифікація автоматизованих систем і стандартні функціональні профілі захищеності оброблюваної інформації від несанкціонованого доступу / ДСТСЗІ СБ України. - Введ. 1998.
5. Кебало О.С., Михайлюк А.Ю., Тарасенко В.П. Функціональні профілі спеціалізованих інформаційно-аналітичних систем // Науковий вісник Чернівецького університету: Збірник наук. праць. 2009. Вип. 423: Фізика. Електроніка.: Тематичний випуск «Комп'ютерні системи та компоненти». Частина I. – Чернівці: ЧНУ – С. 117-123.
6. Л.Ф.Василевич, А.Ю.Михайлюк, В.П.Тарасенко, О.К.Тесленко. Функціонально-орієнтований підхід до проектування інтелектуальних інформаційно-аналітичних систем // Реєстрація, зберігання і обробка даних. – 2010. Т.12, №2. С. 128-142.
7. Тарасенко В.П., Михайлюк А.Ю., Сніжко М.В., Бігун Л.М. Функціональність спеціалізованих інформаційно-аналітичних систем для підтримки інформаційно-навчальної діяльності // Проблеми інформатизації та управління. – 36. наук. праць. – К.: НАУ, 2009. – № 3 (27). – С. 123-130.
8. Ким Д.О., Мьюллер Ч.У., Клекка У.Р. Факторный, дискриминантный и кластерный анализ. –М: Мир, 1980. –389с.
9. F. Sebastiani. Machine learning in automated text categorization // ACM Computing Surveys (2002). – ACM, 2002. – Vol. 34/1. – P.1-47.
10. А.Ю. Михайлюк, О.В. Пилипчук, М.В. Сніжко, В.П. Тарасенко. Квазісемантичний пошук текстових даних в електронному інформаційному ресурсі // Радиоелектроника и информатика. – Харьков: ХНУРЕ. – 2009. – №3. – С.61-67.
11. Р. Уилсон. Введение в теорию графов // М.:Мир, 1977.
12. Н. Бурбаки. Теория множеств // М.:Мир, 1965.
13. Т.В. Волченская, В.С. Князьков. Компьютерная математика: часть 2. Теория графов. Учебное пособие // Пенза, 2002.
14. Н. Кристофидес. Теория графов. Алгоритмический подход // М.: Мир, 1978.