

Виконав: студент групи ІКІ-18м
Тягун Д. Т.
Керівник: к. т. н., доц. каф. ОТ
Савицька Л. А.

МЕТОД ТА ПРОГРАМНИЙ ЗАСІБ ЗАСТОСУВАННЯ МЕТАДАНИХ В ПРОЦЕСАХ ПОШУКУ

Актуальність дослідження

- ▶ Майже експоненційний ріст потоку нових наукових розробок означає, що в адекватні часові терміни практично мало хто може встигнути ознайомитись із матеріалами, що й призводить до того, що гіпотетично перспективні технології навіть не були помічені ні науковим суспільством в цілому, ні гіпотетичним ринком користувачів.
- ▶ Це формує вже досить давно **актуальну задачу** ефективного пошуку інформації в спеціалізованих системах, архівах, і в тому числі, і в мережі Інтернет. Звісно, потужні пошукові системи Google, або CiteSeer дозволяють значно полегшити людині пошук і оцінку релевантності знайдених результатів. Однак, на сучасному етапі розвитку досі необхідний науковий та творчий пошук якщо не цілковито нових підходів, то можливо нетривіального використання уже існуючих, що дозволили б покращити ситуацію.

Мета, об'єкт, предмет дослідження

- ▶ **Метою** дослідження магістерської кваліфікаційної роботи є збільшення результатів пошуку документів, що задовольняють запиту, в рамках деякої статичної колекції документів.
- ▶ **Об'єкт** дослідження - процеси роботи з метаданими.
- ▶ **Предмет** дослідження - методи та програмні засоби застосування метаданих в процесах пошуку.

Наукова новизна одержаних результатів

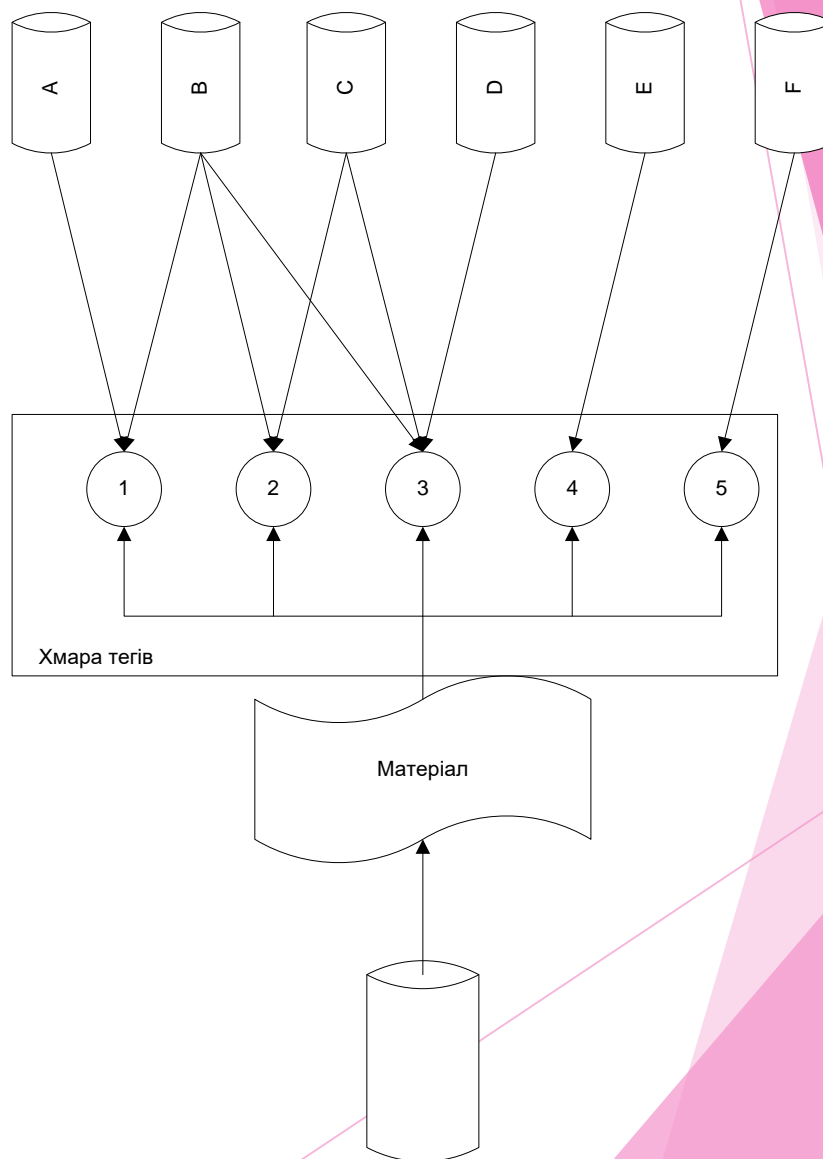
- ▶ вперше запропоновано метод застосування метаданих в процесах пошуку. Запропонований у даній роботі метод застосування метаданих в процесах пошуку дозволяє збільшити результатів пошуку документів, що задовольняють запиту, в рамках деякої статичної колекції документів;
- ▶ вдосконалено модель застосування метаданих в процесах пошуку, що дозволяє створення користувачем власних мета-сутностей;
- ▶ вдосконалено процес роботи з мета-даними, який дозволяє створення семантичних ланцюжків між поняттями;
- ▶ вдосконалено процес створення мета-сутностей, які дозволяють зменшувати кількість тегів у хмарі тегів.

Практичне значення одержаних результатів

- ▶ Розроблено новий метод застосування метаданих в процесах пошуку.
- ▶ Вдосконалено процеси роботи з мета-даними та мета-сутностями, які дозволяють створення семантичних ланцюжків між поняттями та водночас зменшувати кількість тегів у хмарі тегів.
- ▶ Вдосконалено алгоритми роботи з мета-даними та мета-сутностями, які дозволяють збільшити результати пошуку.
- ▶ Розроблено програмний засіб застосування метаданих в процесах пошуку.

Широке фолксономічне тегування

- ▶ Це потік синонімів для одного і того самого поняття.
- ▶ Скажімо, синонімічний ряд тегів для ОС Linux виглядає так: Linux, GNU/Linux, Линакс, Лінукс, Лінуха, Лялікс, Лунікс.



Модель застосування метаданих в процесах пошуку

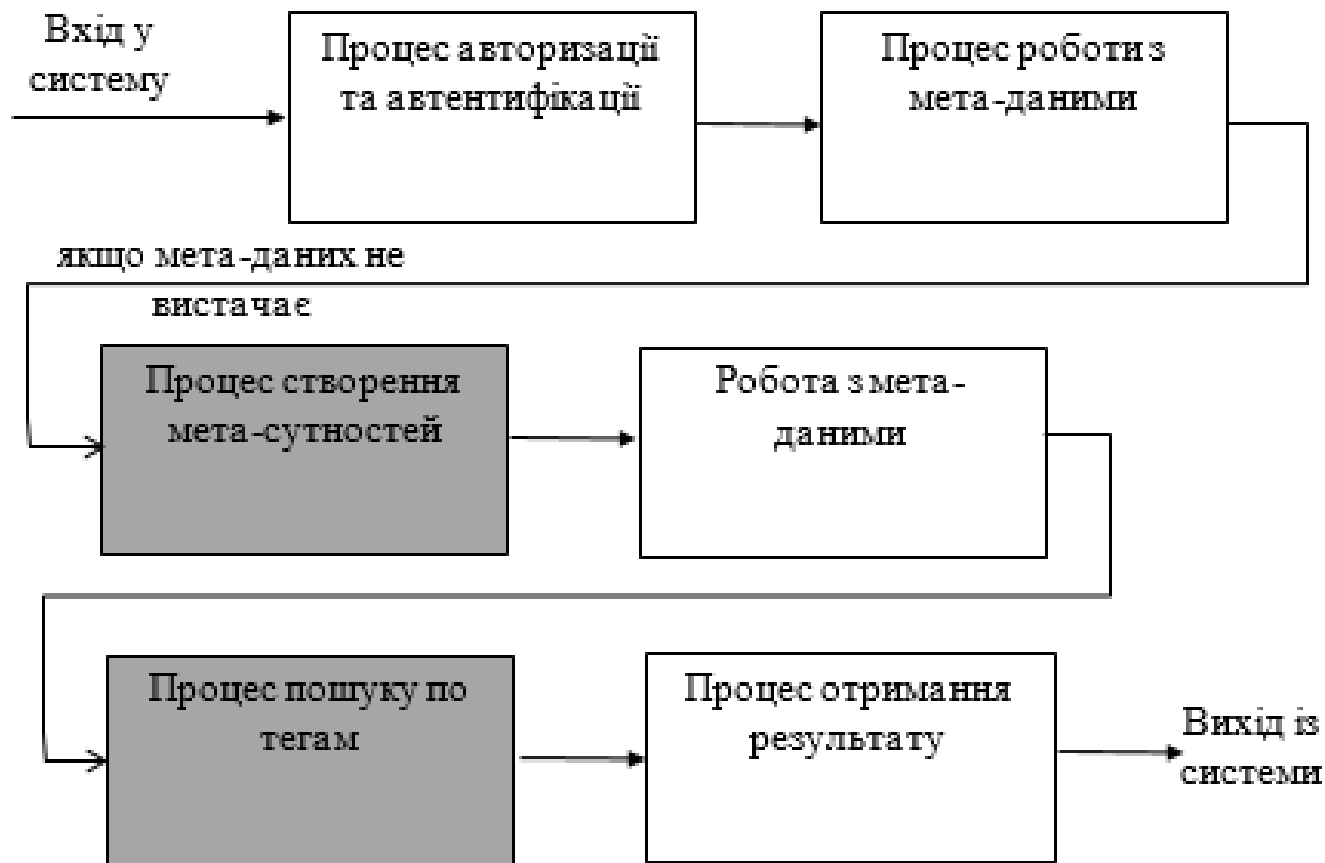
$$O = \{I_1, I_2, \dots, I_n\},$$

$$I_0 = \{T_1, T_2, \dots, T_j; I_1, I_2, \dots, I_k\}, \text{ причому } I_k \neq I_0$$

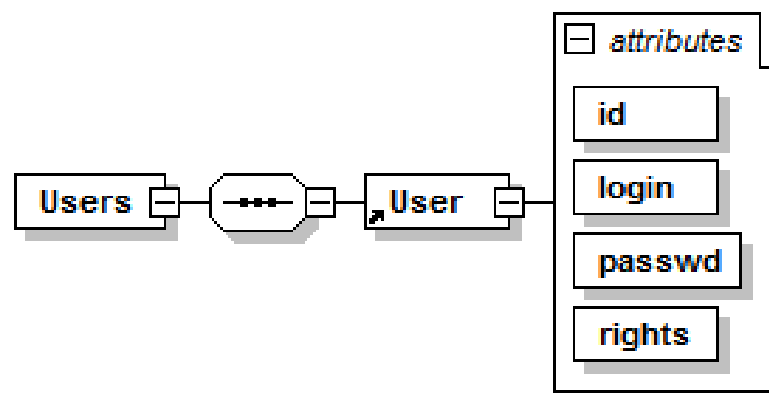
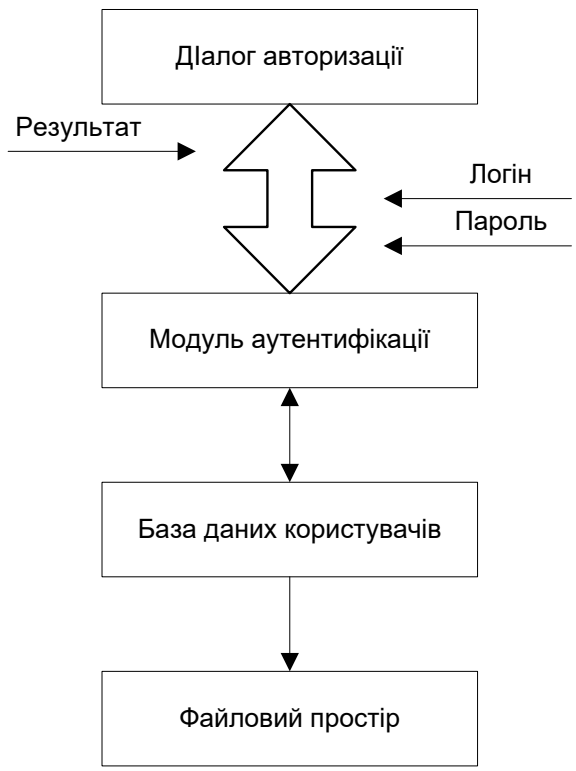
$$I^1 \sim I^2, \text{ за умови } I^1 \cap I^2 = \emptyset$$

$$\begin{cases} F(0) = I^1 \cap I^2 \neq \emptyset \\ F(1) = I^1 \cap (I_1^2, I_2^2, \dots, I_k^2) \neq \emptyset \\ \dots \\ F(n) = I^1 \cap (I_1^n, I_2^n, \dots, I_k^n) \neq \emptyset \end{cases}$$

Метод застосування метаданих в процесах пошуку



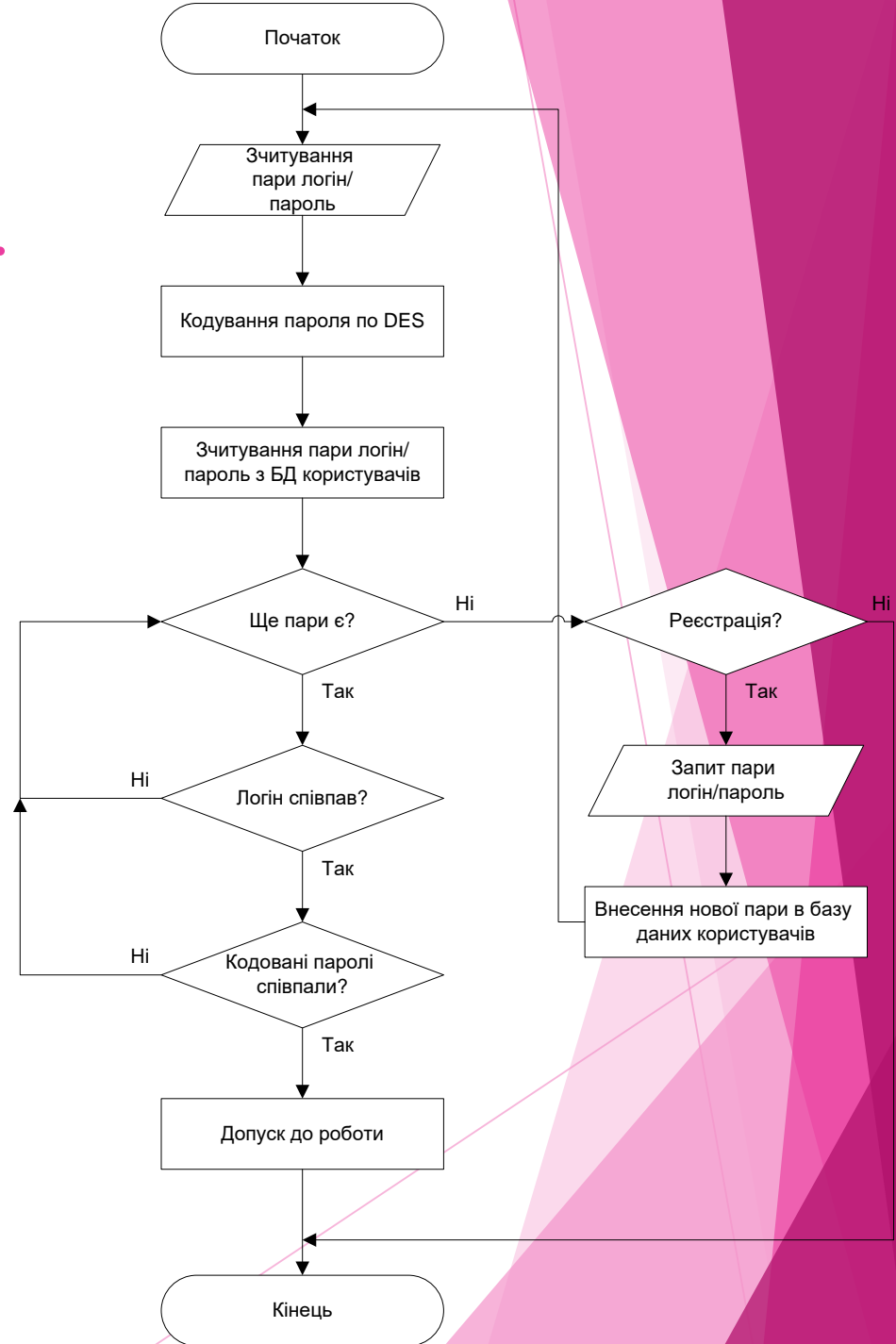
Процес авторизації та автентифікації



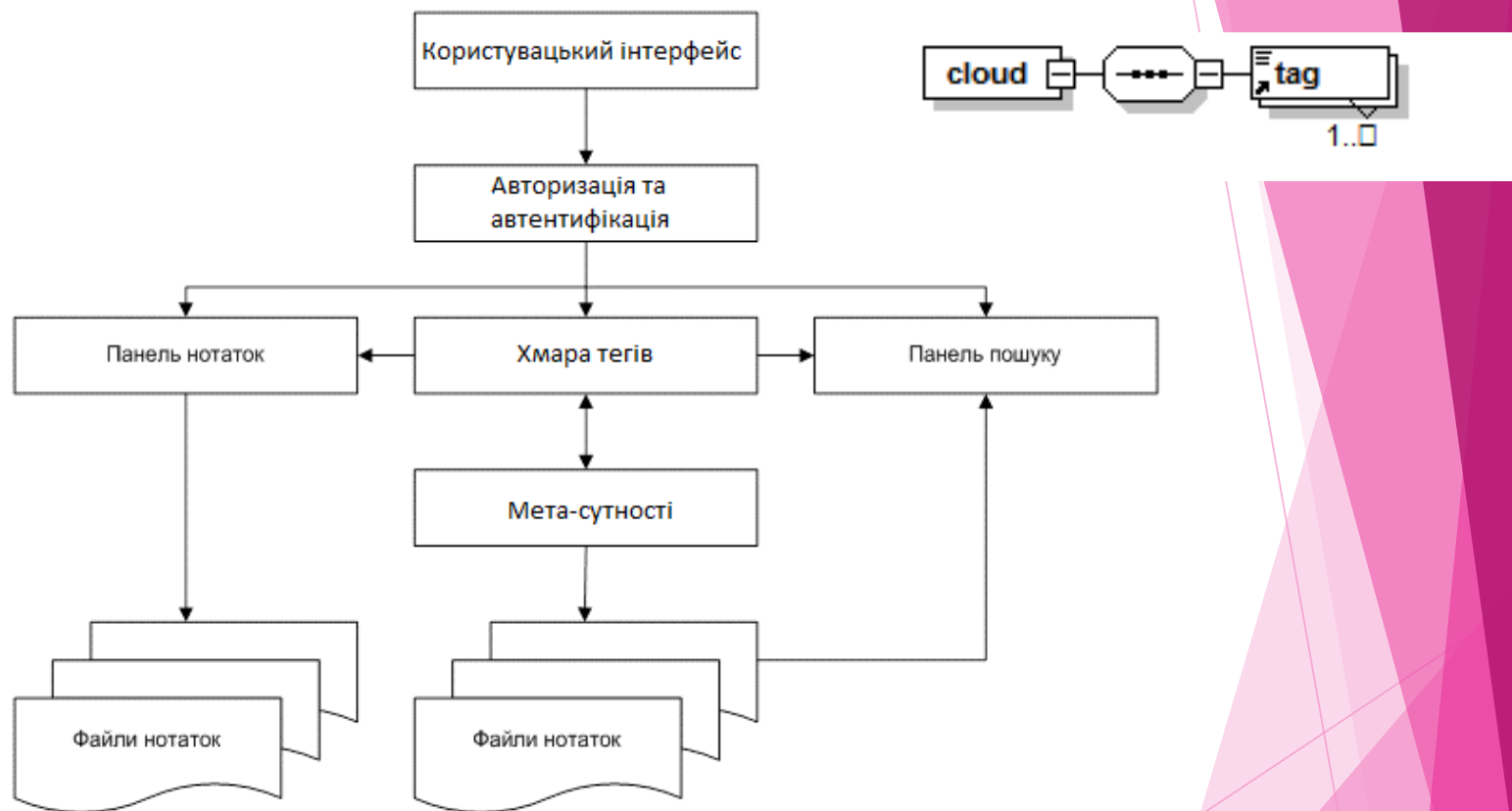
Generated by XmlSpy

www.altova.com

Алгоритм роботи процесу авторизації та автентифікації

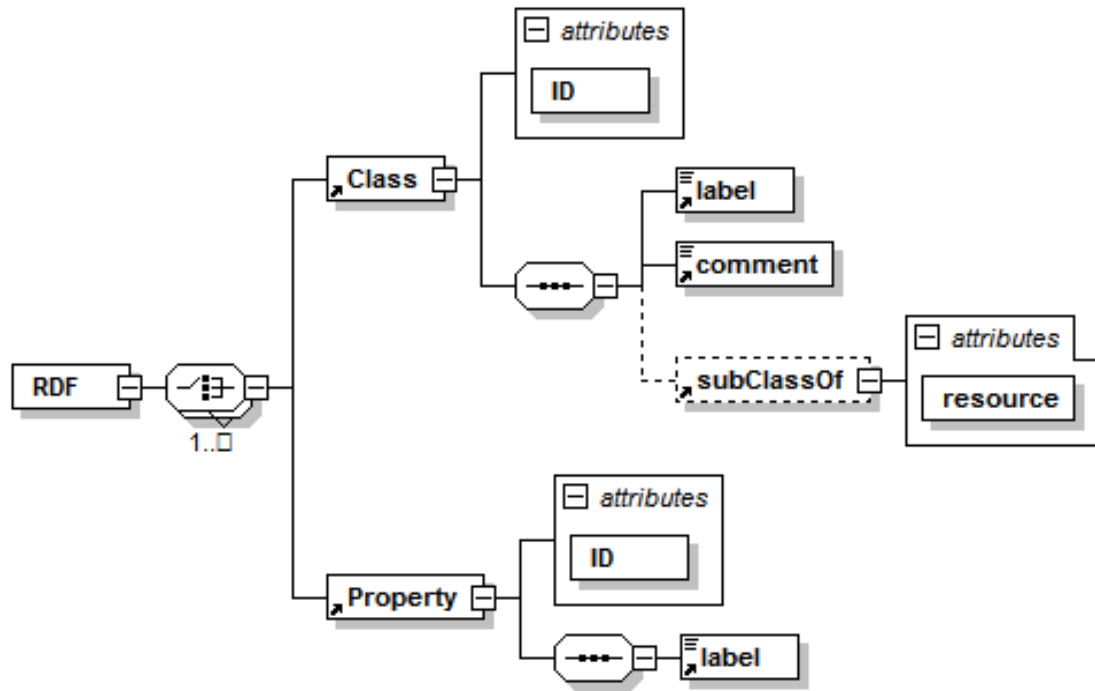


Процес роботи з мета-даними



Центральною точкою програмної моделі є «хмара тегів», до якої звертаються як панель нотаток, так і панель мета-сутностей. Хмара тегів містить в собі всі семантичні мітки - теги, які використовуються користувачами прпід часи створенні нотаток. На панелі мета-сутностей на їх базі вибудовуються зв'язки, які, потім будуть використані в панелі пошуку.

Процес створення мета-сутностей



Generated by XmlSpy

www.altova.com

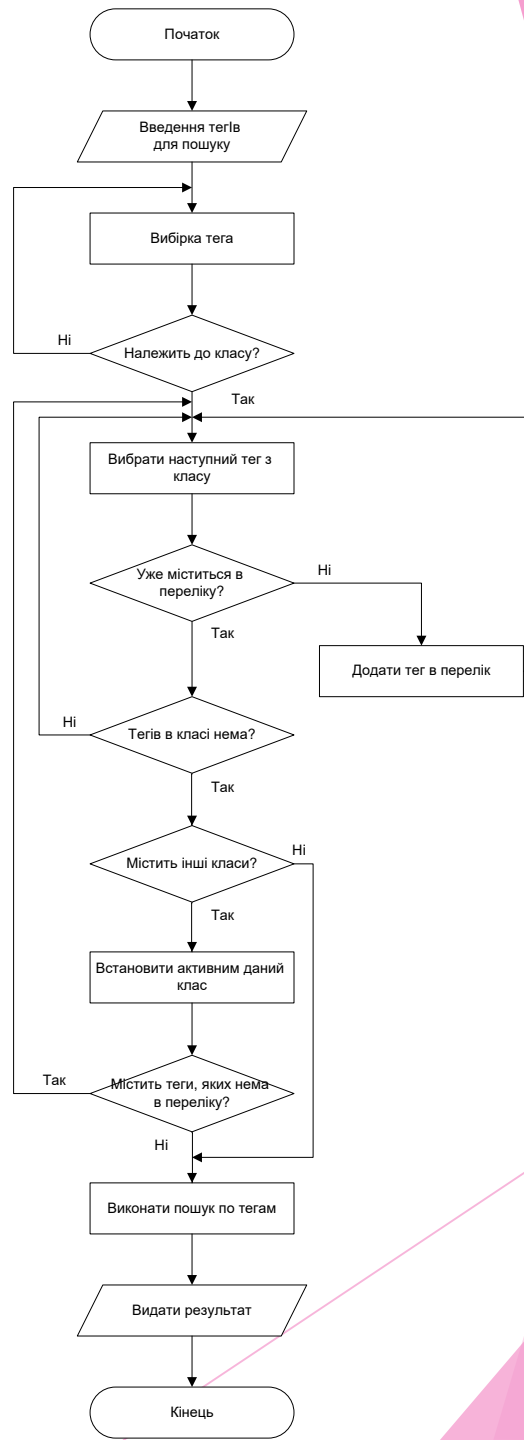
Скористаємось для побудови схеми створення мета-сутностей форматом RDF.

Підмножина RDF обмежує побудову мета-сутностей одним рівнем логічних структур. Таким чином, елементи хмари тегів характеризується як властивості (Property), що створюють спеціалізовані класи (Class), які можуть складатись лише із властивостей, або містити якісь інші класи.

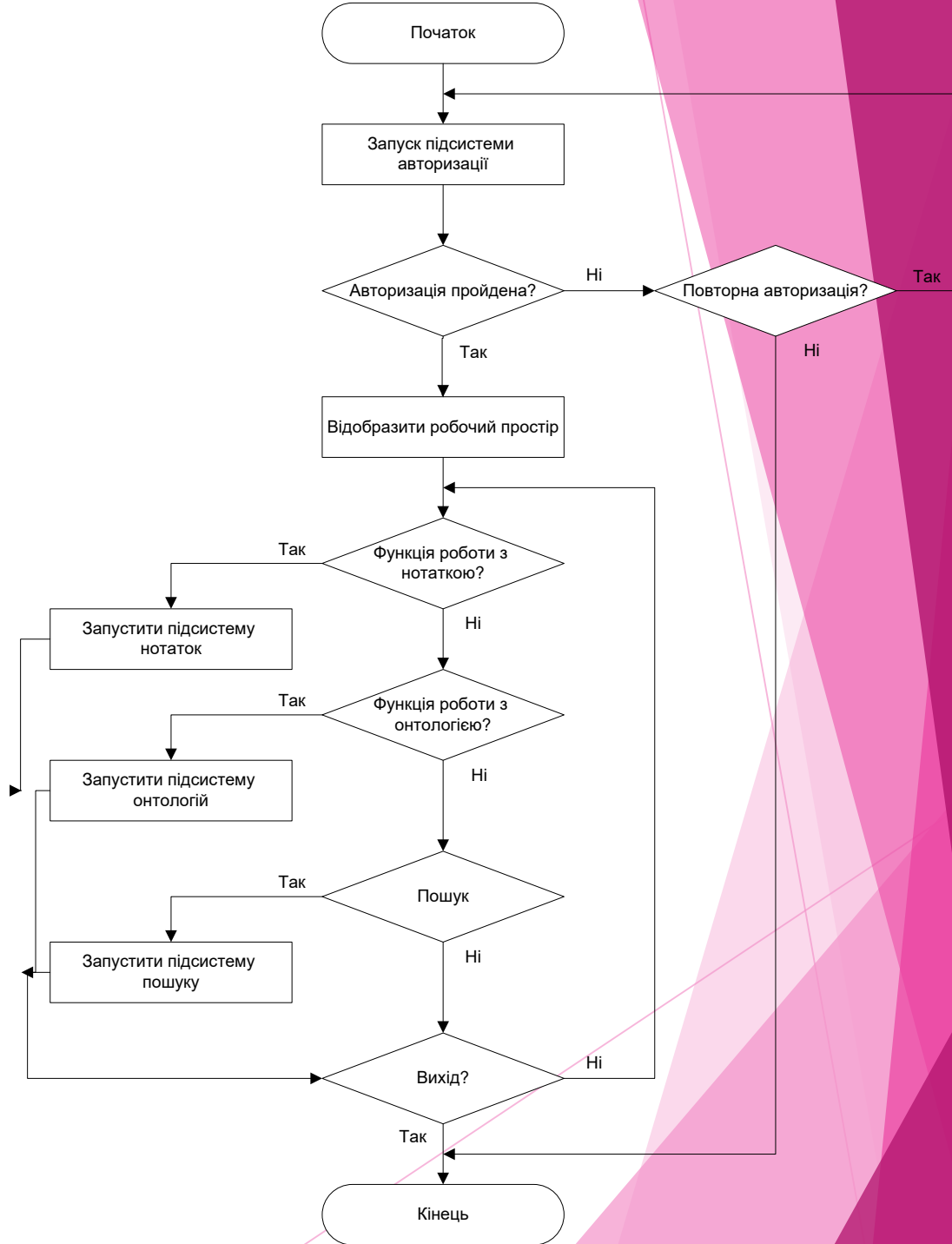
Процес та алгоритм пошуку по тегам

Процес пошуку в даній програмній моделі реалізований за допомогою трьох стратегій:

- 1) пошук по заголовкам;
- 2) пошук за змістом;
- 3) нечіткий пошук за даною мета-сутністю.



Алгоритм функціонування головного модуля програми



Робота програми

Логін

Логін:

Пароль:

Зареєструватись

✓ Прийняти ✗ Скасувати

Реєстрація

Новий логін:

Новий пароль:

✓ Прийняти ✗ Скасувати

Memorabilia - test

Записи Онтологія Мова

вересня 2013

пн	вт	ср	чт	пт	сб	нд
35						1
36	2	3	4	5	6	7 8
37	9	10	11	12	13	14 15
38	16	17	18	19	20	21 22
39	23	24	25	26	27	28 29
40	30					

Панель редагування

Запис: Noname

Заголовок
Тест один

Зміст
Тестовий запис

Теги

Онтологія: Noname

Класи

Додати з хмари тегів

Склад класу

Назва класу

Опис класу

Діалогове вікно пошуку


Пошук [X]

Пошук


Шукати в: ▼


Що шукати:


RDF-онтологія



Період

Початкова дата: 

Кінцева дата: 

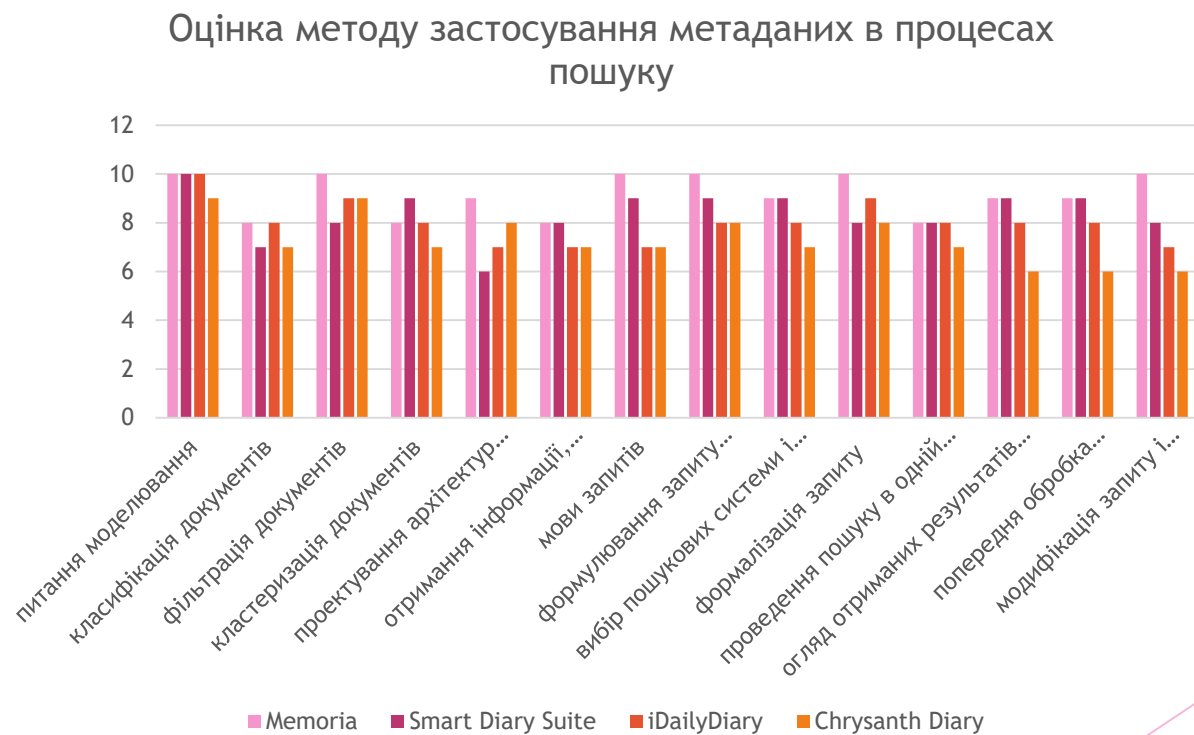


Результати пошуку

Дата	Заголовок

Прийняти **Скасувати**

Оцінка роботи методу на прикладі демонстраційної моделі



ВИСНОВОК

- ▶ З метою полегшити процеси пошуку і виключити аналіз всього тексту в пошуках ключових слів та виразів, в даній роботі пропонується використовувати «мета-теги», які в стислій формі описують вміст ресурсу, які в кінцевому підсумку, супроводжуються «хмарою тематичних тегів».
- ▶ Результатом виконання магістерської кваліфікаційної роботи є збільшення результатів пошуку документів, що задовольняють запиту, в рамках деякої статичної колекції документів. Було спроектовано демонстраційний приклад, який дозволяє показати, яким чином можна здійснювати нечіткий пошук на основі хмари тегів та семантичних метасутностей.
- ▶ Запропонований метод застосування метаданих в процесах пошуку, що реалізований у демонстраційній моделі в інтегральному показнику дозволяє збільшити результатів пошук документів, що задовольняють запиту, в рамках деякої статичної колекції документів на 12,7% у порівнянні із аналогами.

Дякую за увагу!
Доповідь закінчено.