

На дисципліні «Інфопошук» створювалася **ІПС** (інформаційно-пошукова система), що дозволяє проводити процес пошуку в великій колекції деякого неструктурованого матеріала, що задовольняє інформаційній потребі.

**Задачі, що реалізує система:**

1. Словник термінів, що зберігається на диск
  2. Оцінка розміру колекції, загальної кількості слів в колекції та розміру словника.
  3. Побудова матриці інцидентності «термін-документ»; інвертованого індексу (пари термін — *docID*)
  4. Булевий пошук в інвертованому індексі по запиту користувача
  5. Двослівний індекс та координатний інвертований індекс по колекції документів; фразовий пошук і пошук з урахуванням відстані для кожного з цих індексів
  6. Побудова прямого дерева термінів словника B-Tree
  7. Перестановочний індекс та 3-грамний індекс
  8. Реалізація запитів з джокерами по 3-грамному індексу та постфільтрація результатів
  9. Стискання словника та інвертованого індексу колекції як стрічки зі вказівниками; бінарний пошук слова у стиснутому словнику по стрічці
  10. Реалізація пошуку за запитом по колекції документів з використанням зонного ранжування
- ІПС не опрацьовує великі колекції документів, а саме ті, процес обробки яких перевищує оперативну пам'ять. ІПС також не реалізує кластеризацію (пошук з розділенням на кластери).

- Пошукова система працює з такими форматами текстових документів: .txt, .fb2.

*Попова Аліса,*

*ФІ, ІПЗ – 2, група 4*