

## Практическая работа №4. Обработка файлов с помощью технологии LINQ

### Задание:

Реализовать обработку файлов в заданной пользователем папке. Задачи поиска и анализа определяются вариантом. Варианты В1.1 – В1.5 предполагают анализ текстовых файлов (слова, символы), варианты В2.1 - В2.5 предполагают анализ информации о файлах (размер, дата создания и т.п.).

Каждая вычислительная задача должна быть оформлена изолированно (нет зависимости от других задач). Работа каждой вычислительной функции определяется входными аргументами (целевая папка; параметры решения задачи N, K, M, если требуется). Результат решения задачи является возвращаемым значением функции.

Каждую задачу необходимо реализовать двумя способами: с применением технологии LINQ и без использования технологии LINQ. Оценить время выполнения каждой задачи с помощью объекта Stopwatch.

В структуризации рекомендуется задействовать паттерн «Стратегия» - ввести интерфейс с 4-мя вычислительными функциями и классы-решатели (n-p, LinqSolver и NoLinqSolver), которые предоставляют реализации вычислительных функций разными способами (с Linq и без Linq).

Продумать и реализовать пользовательский интерфейс с Windows Forms для установки параметров поиска и вывода результатов. Задействовать элементы DataGridView и Chart для отображения результатов.

Вариант	Задачи поиска
V1.1	<ol style="list-style-type: none"> <li>1. Самое часто встречающееся слово по всем файлам</li> <li>2. N часто встречающихся «общеупотребимых» слов, которые встречаются в каждом файле</li> <li>3. Файл с наибольшим количеством уникальных слов (слова, которые не повторяются в рамках одного файла)</li> <li>4. Распределение всех слов по длине</li> </ol>
V1.2	<ol style="list-style-type: none"> <li>1. Самая часто встречающаяся пара подряд идущих символов</li> <li>2. N самых длинных слов</li> <li>3. Файл с наибольшим количеством «общих» слов (слова, которые присутствуют во всех файлах)</li> <li>4. Распределение знаков пунктуации по частоте</li> </ol>
V1.3	<ol style="list-style-type: none"> <li>1. Самая часто встречающаяся пара подряд идущих слов</li> <li>2. N часто встречающихся слов по всем файлам</li> <li>3. Файл с наибольшим количеством специфических слов (слова, которые отсутствуют в других файлах)</li> <li>4. Распределение длинных слов (более M символов) по частоте</li> </ol>
V1.4	<ol style="list-style-type: none"> <li>1. Самое длинное слово по всем файлам</li> <li>2. N пар символов, которые наиболее часто встречаются</li> <li>3. Пара файлов с наибольшим количеством одинаковых слов (“дубликаты файлов”)</li> <li>4. Распределение коротких слов (не более L символов) по частоте</li> </ol>
V1.5	<ol style="list-style-type: none"> <li>1. Самое часто встречающееся длинное слово (длина не меньше M символов)</li> <li>2. N наиболее часто встречающихся специфических слов (присутствуют только в каком-то одном файле с повторами)</li> <li>3. Файл с наибольшим количеством повторяющихся слов (слова, которые присутствуют в файле не менее двух раз)</li> <li>4. Распределение гласных букв по частоте</li> </ol>

Вариант	Задачи поиска
V2.1	<ol style="list-style-type: none"> <li>1. Самое длинное название файла (без учёта пути к файлу)</li> <li>2. N файлов наибольшего размера</li> <li>3. Наибольшая группа файлов, созданных в фиксированный интервал времени - K минут (н-р, файлы f1, f2, f3 созданы в интервал 13.50-14.00 10.02.2023)</li> <li>4. Распределение файлов по расширениям</li> </ol>
V2.2	<ol style="list-style-type: none"> <li>1. Самый большой по размеру файл</li> <li>2. N самых длинных названий файлов (без учёта пути к файлу)</li> <li>3. Файлы-псевдодубликаты, у которых совпадает только размер или только имя</li> <li>4. Распределение вложенных папок по размеру</li> </ol>
V2.3	<ol style="list-style-type: none"> <li>1. Самое часто встречающееся расширение по всем файлам</li> <li>2. N самых последних файлов (по дате создания)</li> <li>3. Папки-дубликаты, у которых совпадает размер, количество и имена файлов</li> <li>4. Распределение файлов по длине имени (без учёта пути к файлу)</li> </ol>
V2.4	<ol style="list-style-type: none"> <li>1. Самая большая по размеру вложенная папка</li> <li>2. N самых старых файлов (по дате создания)</li> <li>3. Файлы-дубликаты, у которых совпадает имя и размер</li> <li>4. Распределение вложенных папок по количеству файлов</li> </ol>
V2.5	<ol style="list-style-type: none"> <li>1. Самая большая по количеству файлов вложенная папка</li> <li>2. N файлов наименьшего размера</li> <li>3. Специализированные вложенные папки, которые содержат более M% файлов с одним расширением</li> <li>4. Распределение файлов по временным интервалам фиксированной длины (сколько файлов создано в каждый временной интервал длины K минут, начиная от текущего момента времени).</li> </ol>