



Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Московский государственный технический университет имени  
Н. Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н. Э. Баумана)

---

ФАКУЛЬТЕТ «Информатика и системы управления»  
КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

## Отчет по лабораторной работе №5 по дисциплине «Анализ Алгоритмов»

Тема Организация параллельных вычислений по конвейерному принципу

Студент Нисуев Н.Ф.

Группа ИУ7-52Б

Преподаватель Волкова Л. Л., Строганов Д.В.

2024 г.

# СОДЕРЖАНИЕ

|  |           |
|--|-----------|
| <b>ВВЕДЕНИЕ</b>                          | <b>2</b>  |
| <b>1 Входные и выходные данные</b>       | <b>3</b>  |
| <b>2 Преобразование данных</b>           | <b>4</b>  |
| <b>3 Пример работы программы</b>         | <b>5</b>  |
| <b>4 Тестирование</b>                    | <b>8</b>  |
| <b>5 Исследование</b>                    | <b>9</b>  |
| 5.1 Технические характеристики . . . . . | 9         |
| 5.2 Описание исследования . . . . .      | 9         |
| <b>ЗАКЛЮЧЕНИЕ</b>                        | <b>11</b> |
| <b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b>  | <b>12</b> |

## ВВЕДЕНИЕ

В современном программировании эффективная работа с данными является важным фактором для повышения производительности приложений. Одним из методов оптимизации выполнения задач выступает конвейерная обработка [1], при которой процесс обработки данных разбивается на последовательные этапы. Каждый этап выполняется автономно, что позволяет обрабатывать данные параллельно и рациональнее использовать системные ресурсы.

**Цель лабораторной работы:** разработка ПО, выполняющего скачивание страниц и парсинг рецептов с сайта **gastronom.ru** с помощью параллельных вычислений по конвейерному принципу.

Для достижения поставленной цели необходимо выполнить следующие задачи:

- описать входные и выходные данные;
- разработать ПО, выполняющего парсинг рецептов в конвейере;
- выполнить анализ характеристик разработанного ПО на данных о интервалах времени стадий обработки рецепта.

# 1 Входные и выходные данные

Входные данные: сигнал о завершении конвейерной обработки

Выходные данные: файл *recipes.db* с базой данных *SQLite* считанных рецептов.

## 2 Преобразование данных

В интерфейсе программы ожидается нажатие клавиши *Enter*. При запуске программа считывает *URL* адреса рецептов с базовой страницы `gastronom.ru` и записывает их в очередь которая подается алгоритму на вход. Алгоритм организован в виде конвейера из трёх этапов, каждый из которых выполняется в отдельном потоке.

1. Чтение *HTML* контента с *URL*.
2. Парсинг *HTML* контента:
  - считывание данных о рецепте и запись их в структуру данных;
  - считывание новых *URL* адресов рецептов;
  - проверка новых *URL* адресов рецептов на повторение.
3. Запись рецепта в БД.

### 3 Пример работы программы

На рисунке 3.1 показан пример работы программы.

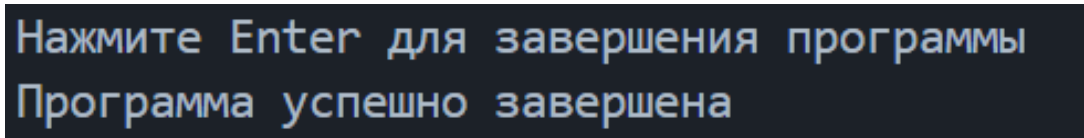


Рисунок 3.1 – Ввод данных в интерфейс приложения

Программа по конвейерному принципу обрабатывает *URL* адреса рецептов и записывает логи в папку *logs*. У каждого этапа свой лог-файл:

- общий лог-файл – *common.log*;
- чтение *HTML* контента с *URL* – *reader.log*;
- парсинг *HTML* контента – *parser.log*;
- запись рецепта в БД – *writer.log*

На рисунках 3.3 – 3.6 представлены примеры лог-файлов.

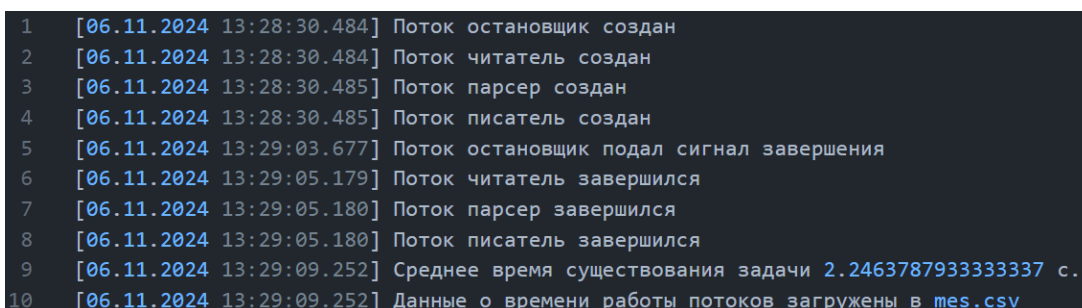


Рисунок 3.2 – Лог-файл *common.log*

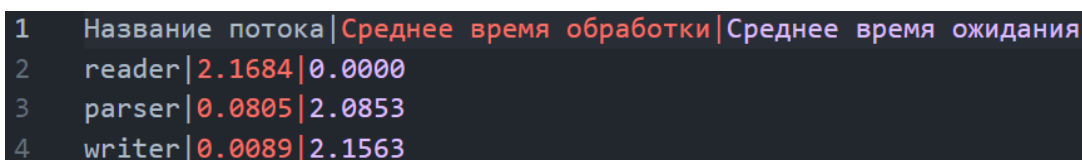


Рисунок 3.3 – Данные из файла *mes.csv*

```

1 [06.11.2024 13:28:36.484] Чтение URL(https://www.gastronom.ru/recipe/45015/kurica-s-cvetnoj-kapustoj-i-gribami)
2 [06.11.2024 13:28:36.485] Загружен HTML для URL(https://www.gastronom.ru/recipe/45015/kurica-s-cvetnoj-kapustoj-i-gribami)
3 [06.11.2024 13:28:32.523] Чтение URL(https://www.gastronom.ru/recipe/38797/salat-iz-grushi-s-granatom)
4 [06.11.2024 13:28:32.523] Загружен HTML для URL(https://www.gastronom.ru/recipe/38797/salat-iz-grushi-s-granatom)
5 [06.11.2024 13:28:34.035] Чтение URL(https://www.gastronom.ru/recipe/54792/zapechennye-mini-frikadelki-s-ajvoj-i-orehami)
6 [06.11.2024 13:28:34.035] Загружен HTML для URL(https://www.gastronom.ru/recipe/54792/zapechennye-mini-frikadelki-s-ajvoj-i-orehami)
7 [06.11.2024 13:28:35.797] Чтение URL(https://www.gastronom.ru/recipe/1859/chizhejk-s-tykvoj)
8 [06.11.2024 13:28:35.797] Загружен HTML для URL(https://www.gastronom.ru/recipe/1859/chizhejk-s-tykvoj)
9 [06.11.2024 13:28:39.193] Чтение URL(https://www.gastronom.ru/recipe/57966/kartoshka-v-duhovke)
10 [06.11.2024 13:28:39.194] Загружен HTML для URL(https://www.gastronom.ru/recipe/57966/kartoshka-v-duhovke)
11 [06.11.2024 13:28:40.865] Чтение URL(https://www.gastronom.ru/recipe/50555/salat-iz-fejhoa-s-kopchenoj-kuricej)
12 [06.11.2024 13:28:40.866] Загружен HTML для URL(https://www.gastronom.ru/recipe/50555/salat-iz-fejhoa-s-kopchenoj-kuricej)
13 [06.11.2024 13:28:43.578] Чтение URL(https://www.gastronom.ru/recipe/56244/pasta-picca-na-shpazhkah)
14 [06.11.2024 13:28:43.578] Загружен HTML для URL(https://www.gastronom.ru/recipe/56244/pasta-picca-na-shpazhkah)
15 [06.11.2024 13:28:46.477] Чтение URL(https://www.gastronom.ru/recipe/56466/brusketty-s-korolevskimi-krevetkami)
16 [06.11.2024 13:28:46.478] Загружен HTML для URL(https://www.gastronom.ru/recipe/56466/brusketty-s-korolevskimi-krevetkami)
17 [06.11.2024 13:28:48.471] Чтение URL(https://www.gastronom.ru/recipe/55491/bystraja-kurinaja-grudka-v-chesnochno-medovom-souse)
18 [06.11.2024 13:28:48.471] Загружен HTML для URL(https://www.gastronom.ru/recipe/55491/bystraja-kurinaja-grudka-v-chesnochno-medovom-souse)
19 [06.11.2024 13:28:50.260] Чтение URL(https://www.gastronom.ru/recipe/31193/tort-tri-moloka)
20 [06.11.2024 13:28:50.260] Загружен HTML для URL(https://www.gastronom.ru/recipe/31193/tort-tri-moloka)
21 [06.11.2024 13:28:52.353] Чтение URL(https://www.gastronom.ru/recipe/56259/midii-na-skovorode-so-smetanoj)
22 [06.11.2024 13:28:52.353] Загружен HTML для URL(https://www.gastronom.ru/recipe/56259/midii-na-skovorode-so-smetanoj)
23 [06.11.2024 13:28:53.978] Чтение URL(https://www.gastronom.ru/recipe/56638/makarony-s-kurinyum-farshem)
24 [06.11.2024 13:28:53.979] Загружен HTML для URL(https://www.gastronom.ru/recipe/56638/makarony-s-kurinyum-farshem)
25 [06.11.2024 13:28:55.868] Чтение URL(https://www.gastronom.ru/recipe/59267/salat-muzhskoj-kapriz-so-svininoj)
26 [06.11.2024 13:28:55.868] Загружен HTML для URL(https://www.gastronom.ru/recipe/59267/salat-muzhskoj-kapriz-so-svininoj)
27 [06.11.2024 13:28:58.216] Чтение URL(https://www.gastronom.ru/recipe/55405/bifshteks-s-jajcom)
28 [06.11.2024 13:28:58.217] Загружен HTML для URL(https://www.gastronom.ru/recipe/55405/bifshteks-s-jajcom)
29 [06.11.2024 13:29:00.491] Чтение URL(https://www.gastronom.ru/recipe/21868/phali-iz-svezhej-zeleni)
30 [06.11.2024 13:29:00.491] Загружен HTML для URL(https://www.gastronom.ru/recipe/21868/phali-iz-svezhej-zeleni)
31 [06.11.2024 13:29:02.838] Чтение URL(https://www.gastronom.ru/recipe/22546/salat-cesar-originalnyj-recept)
32 [06.11.2024 13:29:02.839] Загружен HTML для URL(https://www.gastronom.ru/recipe/22546/salat-cesar-originalnyj-recept)

```

Рисунок 3.4 – Лог-файл *reader.log*

```

1 [06.11.2024 13:28:32.566] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/45015/kurica-s-cvetnoj-kapustoj-i-gribami)
2 [06.11.2024 13:28:34.072] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/38797/salat-iz-grushi-s-granatom)
3 [06.11.2024 13:28:35.831] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/54792/zapechennye-mini-frikadelki-s-ajvoj-i-orehami)
4 [06.11.2024 13:28:39.243] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/1859/chizhejk-s-tykvoj)
5 [06.11.2024 13:28:40.899] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/57966/kartoshka-v-duhovke)
6 [06.11.2024 13:28:43.620] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/50555/salat-iz-fejhoa-s-kopchenoj-kuricej)
7 [06.11.2024 13:28:46.511] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/56244/pasta-picca-na-shpazhkah)
8 [06.11.2024 13:28:48.508] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/56466/brusketty-s-korolevskimi-krevetkami)
9 [06.11.2024 13:28:50.302] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/55491/bystraja-kurinaja-grudka-v-chesnochno-medovom-souse)
10 [06.11.2024 13:28:52.395] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/31193/tort-tri-moloka)
11 [06.11.2024 13:28:54.018] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/56259/midii-na-skovorode-so-smetanoj)
12 [06.11.2024 13:28:55.909] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/56638/makarony-s-kurinyum-farshem)
13 [06.11.2024 13:28:58.259] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/59267/salat-muzhskoj-kapriz-so-svininoj)
14 [06.11.2024 13:29:00.532] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/55405/bifshteks-s-jajcom)
15 [06.11.2024 13:29:02.885] Рецепт успешно получен с URL(https://www.gastronom.ru/recipe/21868/phali-iz-svezhej-zeleni)

```

Рисунок 3.5 – Лог-файл *parser.log*

```

1 [06.11.2024 13:28:32.621] Рецепт с URL(https://www.gastronom.ru/recipe/45015/kurica-s-cvetnoj-kapustoj-i-gribami) успешно добавлен в БД.
2 [06.11.2024 13:28:34.134] Рецепт с URL(https://www.gastronom.ru/recipe/38797/salat-iz-grushi-s-granatom) успешно добавлен в БД.
3 [06.11.2024 13:28:35.873] Рецепт с URL(https://www.gastronom.ru/recipe/54792/zapechennye-mini-frikadelki-s-ajvoj-i-orehami) успешно добавлен в БД.
4 [06.11.2024 13:28:39.302] Рецепт с URL(https://www.gastronom.ru/recipe/1859/chizhejk-s-tykvoj) успешно добавлен в БД.
5 [06.11.2024 13:28:40.946] Рецепт с URL(https://www.gastronom.ru/recipe/57966/kartoshka-v-duhovke) успешно добавлен в БД.
6 [06.11.2024 13:28:43.668] Рецепт с URL(https://www.gastronom.ru/recipe/50555/salat-iz-fejhoa-s-kopchenoj-kuricej) успешно добавлен в БД.
7 [06.11.2024 13:28:46.553] Рецепт с URL(https://www.gastronom.ru/recipe/56244/pasta-picca-na-shpazhkah) успешно добавлен в БД.
8 [06.11.2024 13:28:48.553] Рецепт с URL(https://www.gastronom.ru/recipe/56466/brusketty-s-korolevskimi-krevetkami) успешно добавлен в БД.
9 [06.11.2024 13:28:50.354] Рецепт с URL(https://www.gastronom.ru/recipe/55491/bystraja-kurinaja-grudka-v-chesnochno-medovom-souse) успешно добавлен в БД.
10 [06.11.2024 13:28:52.447] Рецепт с URL(https://www.gastronom.ru/recipe/31193/tort-tri-moloka) успешно добавлен в БД.
11 [06.11.2024 13:28:54.054] Рецепт с URL(https://www.gastronom.ru/recipe/56259/midii-na-skovorode-so-smetanoj) успешно добавлен в БД.
12 [06.11.2024 13:28:55.955] Рецепт с URL(https://www.gastronom.ru/recipe/56638/makarony-s-kurinyum-farshem) успешно добавлен в БД.
13 [06.11.2024 13:28:58.308] Рецепт с URL(https://www.gastronom.ru/recipe/59267/salat-muzhskoj-kapriz-so-svininoj) успешно добавлен в БД.
14 [06.11.2024 13:29:00.581] Рецепт с URL(https://www.gastronom.ru/recipe/55405/bifshteks-s-jajcom) успешно добавлен в БД.
15 [06.11.2024 13:29:02.934] Рецепт с URL(https://www.gastronom.ru/recipe/21868/phali-iz-svezhej-zeleni) успешно добавлен в БД.

```

Рисунок 3.6 – Лог-файл *writer.log*

Пример полученной БД, записанной в файле *recipes.db* представлен на рисунке 3.7

| #  | id | name  | url                                     | image_url                                 | calories     | cook_time | portions | protein      | fat      | carbohy... | ingredients                      |
|----|----|---|---|---|--------------|-----------|----------|--------------|----------|------------|----------------------------------|
| 1  |    | Курица с цветной капустой и грибами             | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 489.99 кк... | PT55M     | 4        | 489.99 кк... | 39.30 г. | 28.94 г.   | курица – 1 шт. (весом 1,2...     |
| 2  |    | Салат из груши с гранатом                       | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> |              | PT10M     | 4        |              |          |            | 2 груши, 1 гранат, полови...     |
| 3  |    | Запеченные мини-фрикадельки с айвой и орехами   | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 687.20 кк... | PT1H      | 4        | 687.20 кк... | 21.74 г. | 54.53 г.   | 500 г, 1 шт., 1 ч.л., 3 шт.,...  |
| 4  |    | Чизкейк с тыквой                                | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 404.00 кк... | PT15M     | 12       | 404.00 кк... |          |            | 2 стакана крошек сдобно...       |
| 5  |    | Картошка в духовке                              | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 352.49 кк... | PT40M     | 4        | 352.49 кк... | 4.39 г.  | 22.39 г.   | картофель – 800 г, оливк...      |
| 6  |    | Салат из фейхоа с копченой курицей              | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> |              | PT30M     | 4        |              |          |            | 300 г спелой мягкой фей...       |
| 7  |    | Пицца на шпакках из пасты                       | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 447.63 кк... | PT35M     | 4        | 447.63 кк... | 20.51 г. | 22.00 г.   | 250 г пасты пенне, 4 сос...      |
| 8  |    | Брускетты с королевскими креветками             | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 310.75 кк... | PT30M     | 6        | 310.75 кк... | 16.14 г. | 19.06 г.   | 500 г крупных сырых кре...       |
| 9  |    | Быстрая куриная грудка в чесночно-медовом соусе | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 419.12 кк... | PT25M     | 4        | 419.12 кк... | 39.49 г. | 12.41 г.   | 4 крупных филе куриной ...       |
| 10 |    | Торт Три молока                                 | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 381.98 кк... | PT8H      | 8        | 381.98 кк... | 11.37 г. | 11.38 г.   | 4 яйца, 1 стакан с горькой...    |
| 11 |    | Мидии на сковороде со сметаной                  | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> |              | PT30M     | 2        |              |          |            | 500 г замороженных очи...        |
| 12 |    | Макароны с куриным фаршем                       | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> |              | PT40M     | 6        |              |          |            | 500 г макарон вашего лю...       |
| 13 |    | Салат Мужской каприз со свиной                  | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 468.03 кк... | PT45M     | 6        | 468.03 кк... | 21.75 г. | 36.75 г.   | 300 г, 250 г, 150 г, 3 шт., 1... |
| 14 |    | Бифитексо с яйцом                               | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> | 873.13 кк... | PT40M     | 4        | 873.13 кк... | 44.42 г. | 72.88 г.   | 800 г говяжьего фарша, ...       |
| 15 |    | Пхали из свежей зелени                          | <a href="https://w...">https://w...</a> | <a href="https://im...">https://im...</a> |              | PT30M     | 4        |              |          |            | 1 кг свежей зелени (ботв...      |

Рисунок 3.7 – БД из файла *recipes.db*



## 4 Тестирование

В таблице 4.1 представлены функциональные тесты для разработанного программного обеспечения. Все тесты пройдены успешно.

Таблица 4.1 – Описание тестовых случаев

| № | Входные данные                                 | Ожидаемый результат   | Результат теста |
|---|--|---|-----------------|
| 1 | Нажат <i>Enter</i>                             | Успешная загрузка $n$ страниц, сохранение в <i>recipes.db</i> | Пройден         |
| 2 | Отключенное интернет-соединение                | Вывод сообщений об ошибках при попытке загрузки страниц       | Пройден         |
| 3 | Отключение интернет-соединения во время работы | Вывод сообщений об ошибках при попытке загрузки страниц       | Пройден         |

## 5 Исследование

### 5.1 Технические характеристики

Характеристики используемого оборудования:

- операционная система — Windows 11 Home [2];
- память — 16 Гб;
- процессор — 12th Gen Intel(R) Core(TM) i7 – 12700H @ 2.30 ГГц [3].

### 5.2 Описание исследования

Во время выполнения, в программе фиксируются моменты времени начала и конца обработки рецепта очередным этапом конвейера. Простои обработчиков можно увидеть на рисунке 5.1.

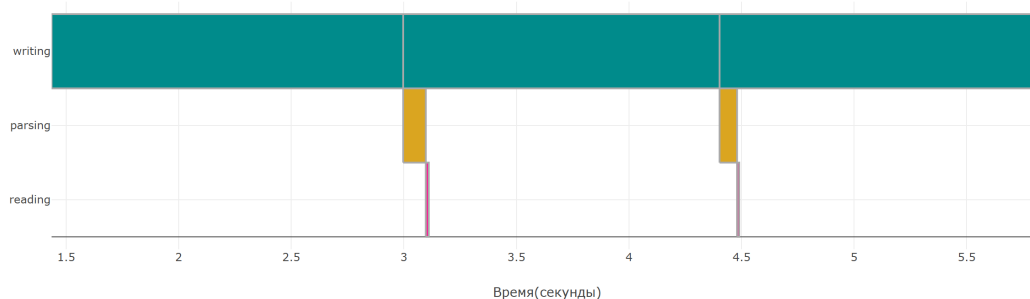


Рисунок 5.1 – График зависимости времени выполнения от количества страниц

Суммарное время простоя и работы каждого этапа представлено на таблице 5.1.

Таблица 5.1 – Описание тестовых случаев

| Название этапа | Среднее время работы(с) | Среднее время простоя(с) |
|----------------|-------------------------|--------------------------|
| reading        | 1.6131                  | 0.0000                   |
| parsing        | 0.0873                  | 1.6392                   |
| writing        | 0.0106                  | 1.7123                   |

В результате исследования было получено, что при выполнении параллельных вычислений в конвейере возможны простои отдельных его частей. Это означает, что в определенные моменты времени часть конвейера может простаивать, ожидая поступления данных от предыдущих этапов обработки.

## ЗАКЛЮЧЕНИЕ

Цель работы достигнута: разработано ПО, выполняющего скачивание страниц и парсинг рецептов с сайта `gastronom.ru` с помощью параллельных вычислений по конвейерному принципу.

В ходе выполнения данной лабораторной работы были решены следующие задачи:

- описаны входные и выходные данные;
- разработано ПО, выполняющего парсинг рецептов в конвейере;
- выполнить анализ характеристик разработанного ПО на данных о интервалах времени стадий обработки рецепта.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *David A. Patterson J. L. H.* Computer Organization and Design: The Hardware/Software Interface. — 5th Edition — Morgan Kaufmann, 2013. — 800 p.
2. Windows technical documentation for developers and IT pros [Электронный ресурс]. — Режим доступа: <https://microsoft.com/en-us/windows/> (дата обращения: 05.10.2024).
3. Intel® Core™ i7-12700H Processor [Электронный ресурс]. — Режим доступа: <https://ark.intel.com/content/www/us/en/ark/products/132228/intel-core-i7-12700h-processor-24m-cache-up-to-4-70-ghz.html> (дата обращения: 05.10.2024).