

# **EXTERNAL JAVA TRAINING. TASK INFORMATION HANDLING LIGHT**

---

---

[www.training.by](http://www.training.by)

**Legal Notice:** This document contains privileged and/or confidential information and may not be disclosed, distributed or reproduced without the prior written permission of EPAM Systems.

---

April 26, 2019

## Task. Information handling light

*Создать приложение, разбирающее текст из файла и позволяющее выполнять с текстом **три** различных операции.*

### Требования

- Разобранный текст должен быть представлен в виде объекта, содержащего, например, абзацы, предложения, лексемы, слова, символы. *Лексема - часть текста, ограниченная пробельными символами.* Использовать Composite.
- Классы с информацией являются классами сущностей и не должны быть перенгружены методами логики.
- Разобранный текст необходимо восстановить в первоначальном виде. Пробелы и знаки табуляции при разборе могут заменяться одним пробелом.
- Для деления текста на составляющие следует использовать регулярные выражения. Регулярные выражения для приложения определять как литеральные константы.
- Код, выполняющий разбиение текста на составляющие части, следует оформить в виде классов-парсеров с использованием Chain of Responsibility.
- При разработке парсеров, разбирающих текст, необходимо следить за количеством создаваемых объектов-парсеров.
- Для записи логов использовать Log4J.
- Реализовывать индивидуальные задания по работе над текстом.

### Индивидуальные задания

- 1 Отсортировать абзацы по количеству предложений.
- 2 Отсортировать слова в предложении по длине.
- 3 Отсортировать предложения в абзаце по количеству слов.

### Пример текста

It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English.

It is a established fact that a reader will be of a page when looking at its layout.

Bye.

REVISION HISTORY					
Ver.	Description of Change	Author	Date	Approved	
				Name	Effective Date
<1.0>	Первая версия	Игорь Блинов Ольга Смолякова	08-05-2014		
<1.1>	Незначительные правки формулировок предложений	Ольга Смолякова	24-09-2015		
<1.2>	Изменения общей формулировки задачи. Удаление устаревших вариантов. Добавление новых вариантов	Игорь Блинов	29-09-2015		
<1.3>	Добавление нового варианта задания. Уточнение общей формулировки	Игорь Блинов	06-10-2016		
<1.4>	Удаление варианта задания. Уточнение общей формулировки и некоторых заданий	Игорь Блинов	28-11-2016		
<1.5>	Упрощение общей формулировки задания	Игорь Блинов	25-04-2019		