

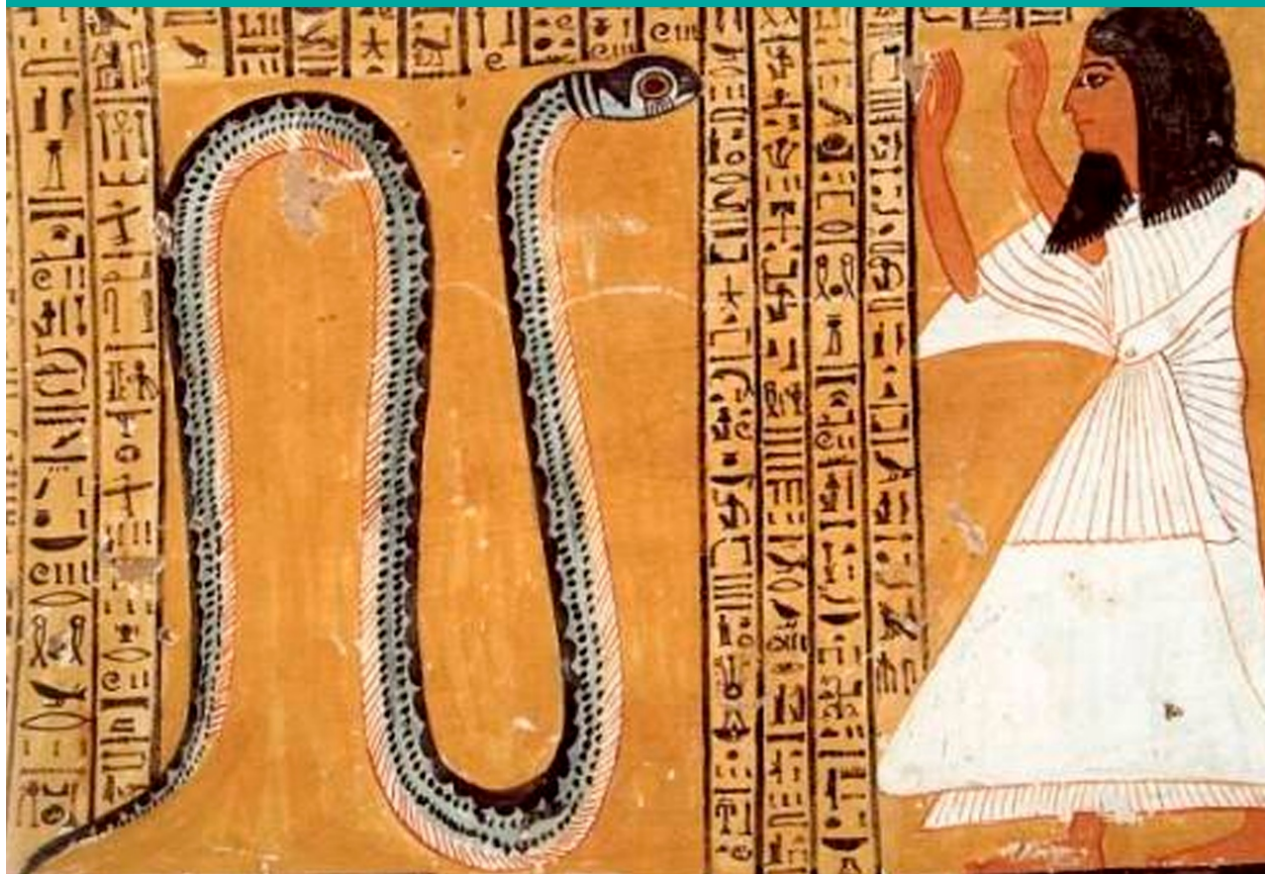


# Информатика



**Лекция №4. Тема: «Основы регулярных выражений. Протоколы, форматы обмена информацией и языки разметки документов. Основы формальных грамматик.»**

**Древние египтяне пытаются освоить  
регулярки на питоне. Фото в цвете.**



# Функции в регулярных выражениях в Python



REGULAR EXPRESSION 2 matches (8 steps, 6.1ms)

`r" \w+"` "gm

TEST STRING

Hello, World!

re.search()

```
In [5]: print(re.search(r'\w+', 'Hello, World!'))  
        <re.Match object; span=(0, 5), match='Hello'>
```

re.findall()

```
In [7]: print(re.findall(r'\w+', 'Hello, World!'))  
        ['Hello', 'World']
```



**re.compile()**

re.match()

re.search()

re.fullmatch()

re.findall()

re.split()

re.sub()

re.finditer()

```
% python -m timeit -s "import re" "re.match('hello', 'hello world')"  
1000000 loops, best of 3: 0.661 usec per loop
```

```
% python -m timeit -s "import re; h=re.compile('hello')" "h.match('hello world')"  
1000000 loops, best of 3: 0.285 usec per loop
```

```
% python -m timeit -s "import re" "h=re.compile('hello'); h.match('hello world')"  
1000000 loops, best of 3: 0.65 usec per loop
```

```
% python --version  
Python 3.6.5 :: Anaconda, Inc.
```



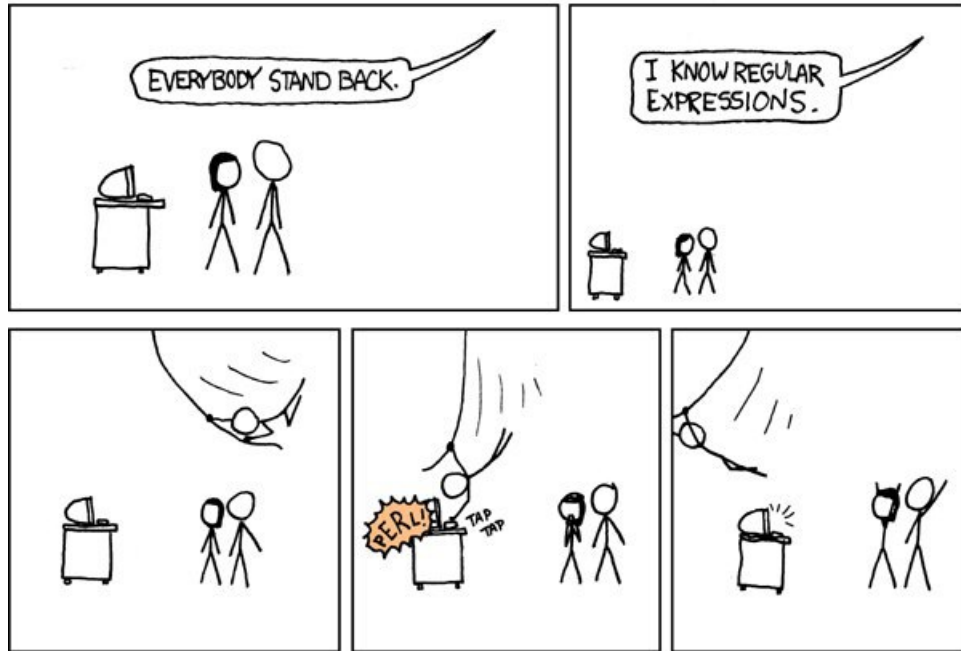
Заменить цифру на "!", если после неё стоит буква:

```
test = "1jgl1jgigi1Y4Tg4 iojg8GG92  
4jg8492ygggn9840n87qyn0q4ty94n 4a6632П ш9К 2Ёё "  
pattern = r"[0-9](?=[a-za-яё])"  
result = re.sub(pattern, "!", test, flags=re.IGNORECASE)  
print(result)
```



<https://habr.com/ru/post/349860/> - Много примеров, заданий и объяснений







## GEEKS IN A BAR

## Немного юмора про RegEx (2)







**WYSIWYG** — What you see is what you get

**WYSIWYM** — What you see is what you mean

**Markup Language\*** — система/язык для маркировки и выделения документа, которая указывает его логическую структуру, и даёт инструкции для представления и компоновки содержимого документа, особенно при электронной передаче и отображении.

\*<https://www.merriam-webster.com/dictionary/markup%20language>

```
<recipe>
  <title>Peanut-butter On a Spoon</title>
  <ingredientlist>
    <ingredient> Peanut-butter </ingredient>
  </ingredientlist>
  <preparation>
    Open a jar of peanut-butter
    Stick a spoon in a jar of peanut-butter
    Scoop and pull out a big glob of peanut-butter
    Close a jar
  </preparation>
</recipe>
```

William Warren Tunnicliffe  
(1922-1996)



GenCode → GML → SGML

# SGML Document Components

1967: GenCode — William Warren Tunnicliffe

1969: Generalized Markup Language — Charles F. Goldfarb



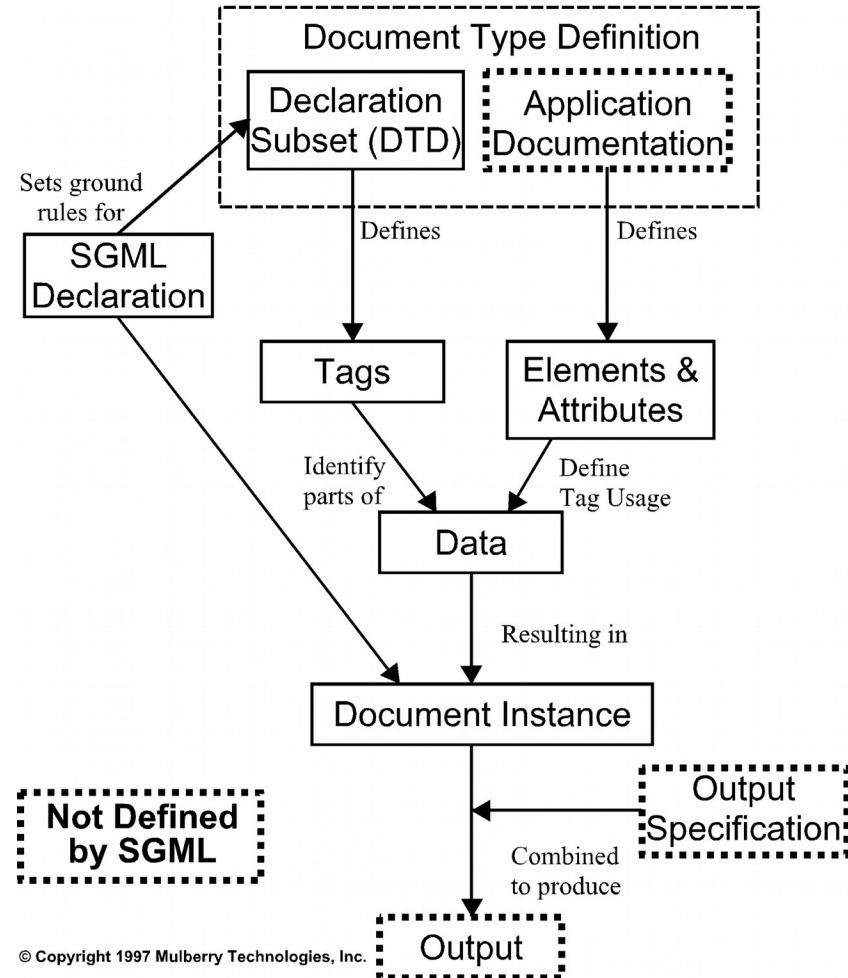
1986: Standard Generalized Markup Language

Пример синтаксиса SGML:

`<quote type="example">`

typically something like `<italics>this</italics>`

`</quote>`

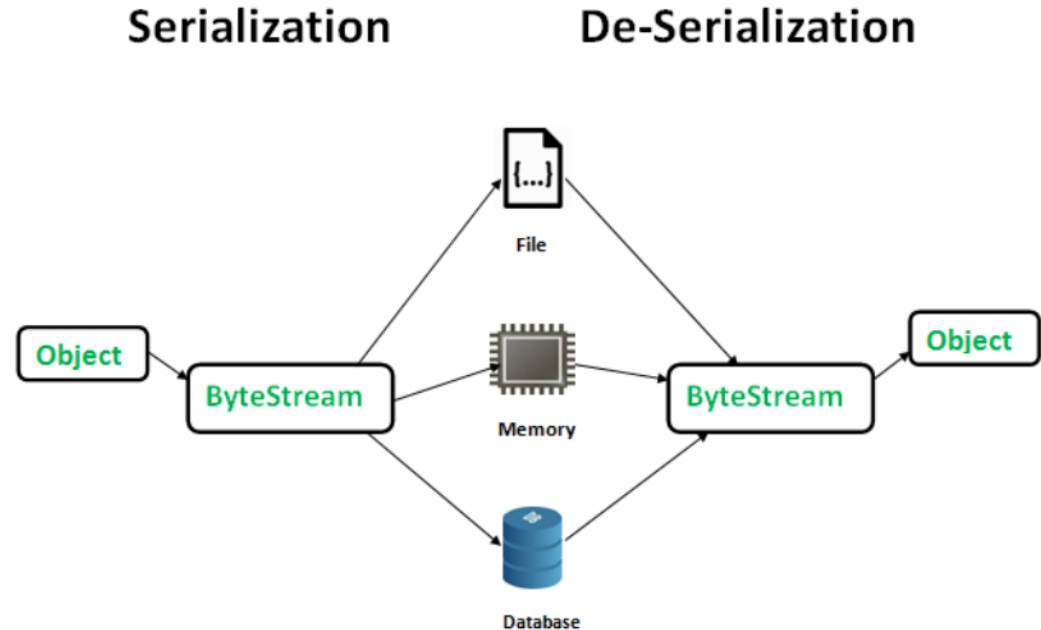




**Сереализация** — процесс перевода структур данных или состояния объекта в формат, который может быть сохранён или передан и реконструирован позже. Обратный процесс — **десериализация**.

Сереализация применяется для:

- передачи данных по сетям;
- сохранения данных (базы данных);
- удалённых вызовов процедур;
- распространения объектов;
- обнаружения изменений в данных, изменяющихся со временем.

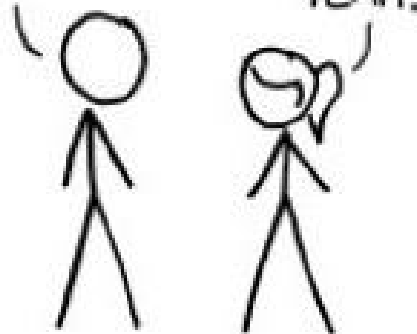




## HOW STANDARDS PROLIFERATE: (SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)

SITUATION:  
THERE ARE  
14 COMPETING  
STANDARDS.

14?! RIDICULOUS!  
WE NEED TO DEVELOP  
ONE UNIVERSAL STANDARD  
THAT COVERS EVERYONE'S  
USE CASES.



SOON:

SITUATION:  
THERE ARE  
15 COMPETING  
STANDARDS.



Parquet

XHTML

eXtensible HyperText Markup  
Language





<div style="

```
font-size: 88%;  
line-height: 1.2em;  
position: absolute;  
z-index: 2;  
left: Ошибка выражения:  
неожидаемый оператор <%;
```

top: Ошибка выражения:  
неожидаемый оператор <%">  
Басильевский остров



<div style="

```
font-size: 88%;  
line-height: 1.2em;  
position: absolute;  
z-index: 2;  
left: Ошибка выражения:  
неожидаемый оператор <%;
```

top: Ошибка выражения:  
неожидаемый оператор <%">

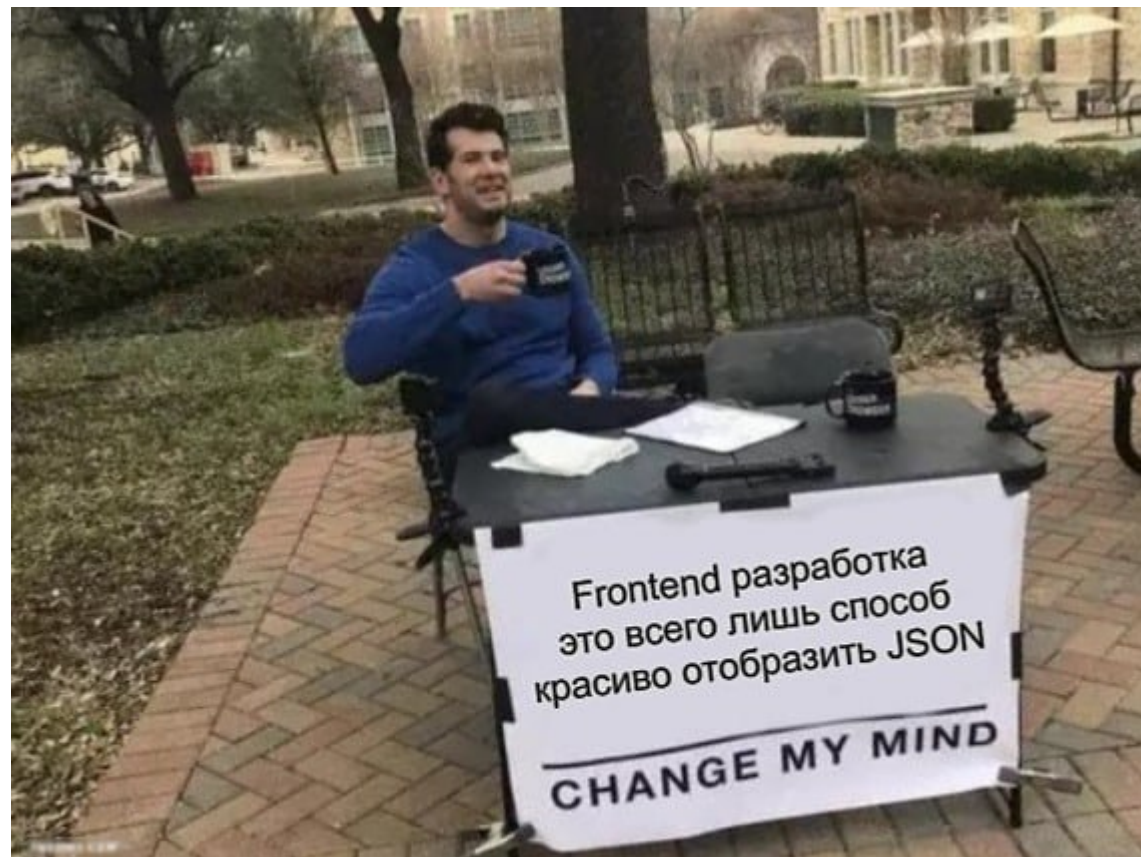
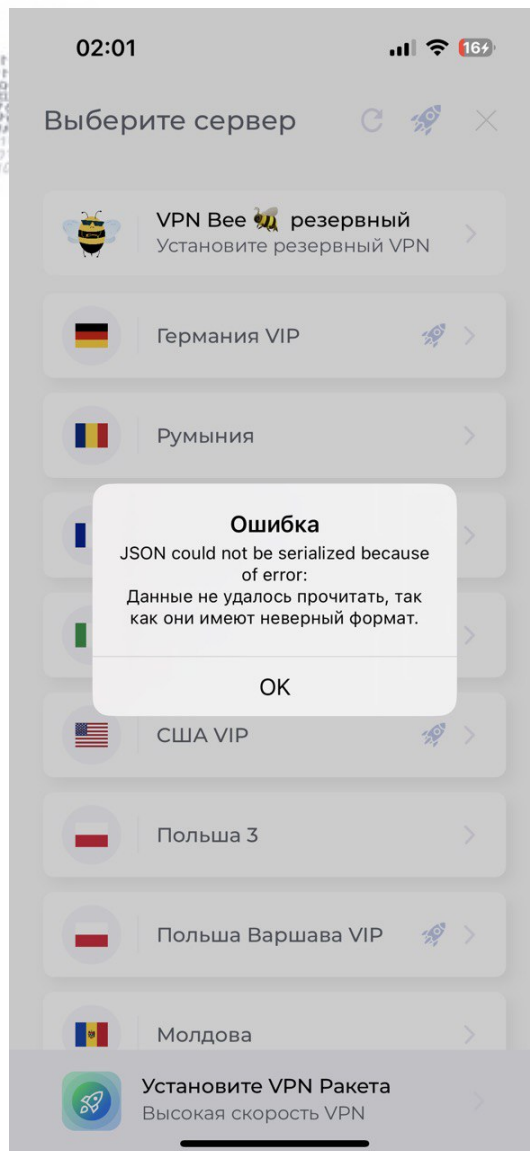


Медиафайлы на Викискладе





## Примеры из жизни (2)







МУНИЦИПАЛЬНОЕ ПРЕДПРИЯТИЕ ГОРОДА РЯЗАНИ

Кустовой Вычислительный Центр

Телефоны:  
33 - 75 - 78  
33 - 78 - 03

Новости

О предприятии

Законодательство

Тарифы

Телефоны экстренных служб

Вопросы

Личный кабинет

Главное меню

Unexpected token < in JSON at position 0

Когда случайно открыл xml-файл с помощью Visual Studio

Viber: JSON

WhatsApp: JSON

Telegram: TL (Type Language, собственный ML телеграма)/JSON


VK: JSON

Twitter: JSON





## Markdown — облегчённый язык разметки



example.Rmd

←

→

📄

🔍

ABC

🔗

Knit HTML

⚙️

📁

📤

📥

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

# Header 1

This is an R Markdown document. Markdown is a simple formatting syntax for authoring webpages.

Use an asterisk mark to provide emphasis, such as *italics* or **bold**.

Create lists with a dash:

- Item 1

- Item 2

- Item 3

```

Use back ticks to create a block of code

```

Embed LaTeX or MathML equations,

$$\frac{1}{n} \sum_{i=1}^n x_i$$

Or even footnotes, citations, and a bibliography. [^1]

[^1]: Markdown is great.

1:1 Header 1 R Markdown

~/Desktop/example.html

example.html

Open in Browser

Find

Header 1

This is an R Markdown document. Markdown is a simple formatting syntax for authoring web pages.

Use an asterisk mark to provide emphasis, such as *italics* or **bold**.

Create lists with a dash:

- Item 1
- Item 2
- Item 3

Use back ticks to create a block of code

Embed LaTeX or MathML equations,  $\frac{1}{n} \sum_{i=1}^n x_i$

Or even footnotes, citations, and a bibliography. <sup>1</sup>

1. Markdown is great.↩



<https://typora.io/> - !!! Рекомендуемый многими Markdown-редактор !!!

<https://dillinger.io/> - !!! Markdown online !!!

<https://github.com/typst/typst> - !!! Markdown!!!

<https://www.pvsm.ru/java/70568/> - !!! Сравнение JSON и YAML !!!

<https://habr.com/post/248147/> - !!! Сравнение XML, YAML и JSON !!!

<https://habr.com/company/wrike/blog/279797/> - !!! Parquet !!!



<https://jsonformatter.org/> - !!! JSON online парсер!!!

<https://wtools.io/> - !!! Удобный конвертор между форматами !!!

<https://onlinejsontools.com/> - !!! Ещё один конвертор между форматами !!!



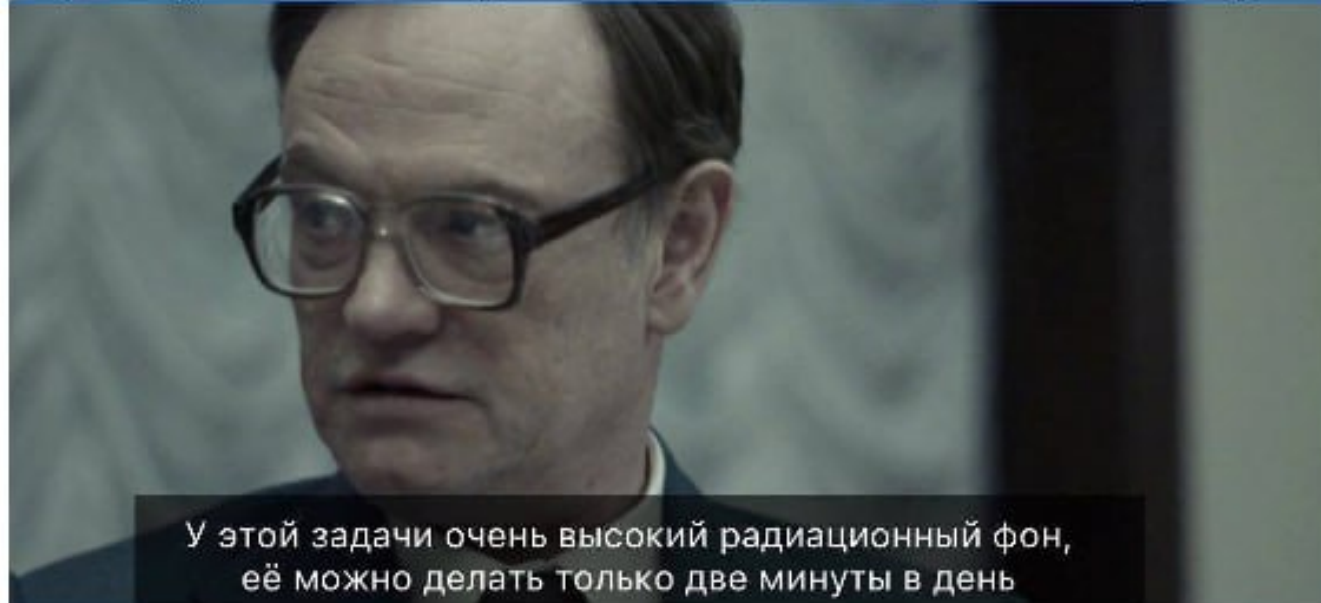
<https://al.cs.msu.ru/files/formal.languages.translation.theory.pdf>

Учебное пособие МГУ «Формальные грамматики и языки. Элементы теории трансляции»

<http://mech.math.msu.su/~vzb/2course/lect10.html>

Лекция МГУ «Формальные грамматики»

Те, кому попался вариант XML->JSON в 4 лабораторной:



У этой задачи очень высокий радиационный фон,  
её можно делать только две минуты в день

Процесс выполнения лабораторной работы №4:







**ПСЫЖ**

**АКТИВАТОР ЖИЗНЕННОЙ ЭНЕРГИИ**

МИНЕРАЛЬНАЯ ПРИРОДНАЯ ГАЗИРОВАННАЯ  
ВОДА С КАВКАЗА

ЗАРЯДКА ЭНЕРГИЕЙ АКТИВИЗИРУЕТ ЖИЗНЕННЫЙ ТОНУС. СПОСОБСТВУЕТ ВЫРАБОТКЕ СЕРОТОНИНА