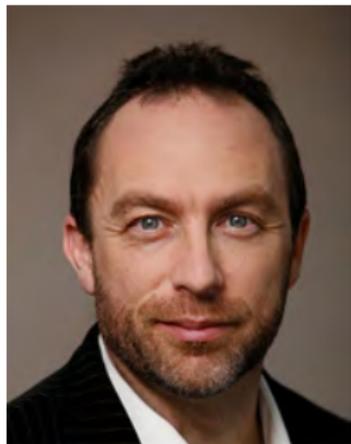


Интернет, окно возможностей

Я не хочу десять лет спустя сказать: “Да я еще тогда знал, что интернет станет огромной сетью”, и ничего не сделать.

- Jimmy Wales



Интернет сегодня

- Поиск
 - Google
 - Yahoo
 - Baidu
 - Live.com
- Интернет-магазины
 - Amazon
 - Ebay
 - Taobao
- Социальные сети
 - Facetube
 - Youbook
 - Twitter

Интернет завтра

- Семантический веб
- Социальные медиа
- Облако

Интернет завтра

- Семантический веб
Интернет будет знать о нас больше
- Социальные медиа
- Облако

Интернет завтра

- Семантический веб
Интернет будет знать о нас больше
- Социальные медиа
в частности потому что мы сами ему об этом расскажем
- Облако

Интернет завтра

- Семантический веб
Интернет будет знать о нас больше
- Социальные медиа
в частности потому что мы сами ему об этом расскажем
- Облако
*а еще все файлы, фотографии, документы и прочее ему отдадим.
Сами.*

Основные навыки для разработки веб-проектов

- Программирование на языке быстрого прототипирования
- Умение работать с базами данных
- Умение быстро обрабатывать большие количества данных

Основные навыки для разработки веб-проектов

- Программирование на языке быстрого прототипирования
 - Linux: Release early, release often*
 - Google: Because beta is better than nothing*
 - PHP
 - Ruby
 - Python
- Умение работать с базами данных
- Умение быстро обрабатывать большие количества данных





- 140 000 000 пользователей;
- Примерно 300 000 000 твитов в обычный день;
- Интересный факт: землетрясения распространяются медленнее, чем твиты о них.

twitter



Интерпретатор и средства разработки

- GNU/Linux, OS X: большинство дистрибутивов устанавливает Python по умолчанию.
- Windows:
<http://www.python.org/ftp/python/2.7.2/python-2.7.2.msi>.
- Windows x86_64 <http://www.python.org/ftp/python/2.7.2/python-2.7.2.amd64.msi>

Очень быстрое введение в Python

Вывод текста в консоль на Си

```
#include <stdio.h>

int main(void)
{
    printf("Hello World!\n");
    return 0;
}
```

Вывод текста в консоль на Python

```
print "Hello World!"
```

Очень быстрое введение в Python

Вывод текста в консоль на Си

```
#include <stdio.h>

int main(void)
{
    printf("Hello World!\n");
    return 0;
}
```

Вывод текста в консоль на Python

```
print "Hello World!"
```

Типы данных, динамическая типизация, приведение типов

- String: строковые переменные.
- Integer: может быть любым целым числом.
- Float: любое вещественное число.
- Boolean: логический тип, может принимать значение True или False.
- List: список.
- Dict: ассоциативный массив.

Типы данных, динамическая типизация, приведение типов

Типы в Си

```
int a = 2;  
char b[] = "Строка";  
float c = 3.141592654;
```

Типы данных, динамическая типизация, приведение типов

Типы в Си

```
int a = 2;  
char b[] = "Строка";  
float c = 3.141592654;
```

Типы в Python

```
a = 2  
a = "Строка"  
c = 3.141592654  
d = True  
e = [1, 2, 3, "str", True]  
f = {1: "str", 2: True, "str": 3.1}
```

Операции со списками

```
test_list = [1, 2, 3, True]
list2 = ["str1", 3.14]
empty_list = []
```

```
a = test_list[2]
```

```
b = test_list[:2]
```

```
c = test_list[1:3]
```

```
d = test_list + list2
```

Операции со списками

```
test_list = [1, 2, 3, True]
list2 = ["str1", 3.14]

test_list.append(False)

test_list.pop()

a = test_list[:2] + list2 + test_list[2:]

len(a)
```

Ветвление, циклы

```
a = 1

if a == 1:
    print "1"
elif a == 2:
    print "2"
else:
    print "hm"
```

Ветвление, циклы

```
a = [10, 1, 2, 3, 4]
```

```
for i in a:  
    print i
```

```
i = 0  
while i < len(a):  
    print a[i]  
    i = i + 1
```

Функции

```
def my_function(a, b):  
    print a+b  
    return a+b
```

```
a = my_function(1, 2)
```

- 1 Файл `twitter_rv_19000000.net`
Текстовый файл, содержит построчно номер пользователя и номер фоловера (3.5 Гб)
- 2 Файл `numeric2screen`
Текстовый файл, содержит построчно соответствие номера пользователя и его ника (793 Мб)
- 3 Файл `celebrities_profiles.txt`
Текстовый файл, содержит построчно информацию из профилей “звёзд” (3 Мб)

Задание №1 (С подвохом)

Прочитать файл и построить индекс.

Индекс - список всех пользователей, которые есть в базе
twitter_rv_19000000.net

Работа с файлами

```
f = open("filename", "r")

index = []

data = f.readlines()
for l in data:
    print l
    values = l.split('\t')
    index.append(int(values[0]))

f.close()
```

- зачем объявлен пустой список `index = []` ?
- что здесь происходит: `data = f.readlines()` ?
- а здесь: `for l in data:` ?

Работа с файлами

```
f = open("filename", "r")

index = []

for l in f:
    print f
    data = l.split('\t')
    index.append(int(values[0]))

f.close()
```

- что происходит в строке "for l in f:" ?
- что делает функция split ?
- что делает функция int() ?