

# Обработка и анализ фильмов

Импортируем библиотеки для работы с данными

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

import warnings
warnings.filterwarnings("ignore")
```

Импортируем данные и выведем некоторые строки

```
df = pd.read_csv("habr_articles.csv", index_col="Unnamed: 0")
df.head()
```

	text
0	Привет, Хабр! Сегодня разберём, как реализовать...
1	Старший iOS-разработчик red_mad_robot Аня Коче...
2	Еще до начала MWC в Барселоне было очевидно, ч...
3	Взлом Instagram*-аккаунта — популярный запрос ...
4	Привет, Хабр! Меня зовут Кирилл Прямов, я мене...

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 1131 entries, 0 to 1130
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype
---  -
 0   text    1131 non-null    object
dtypes: object(1)
memory usage: 17.7+ KB
```

Проверим наличие дубликатов и удалим их

```
df.duplicated().sum()

24

df.drop_duplicates(inplace=True)
```

## Предварительная обработка данных

Очистим данных для их последующей векторизации:

- Приведение к нижнему регистру

- Очистка от латинских символов
- Очистка от пунктуации и спец символов
- Очистка от цифр
- Очистка от лишних пробеллов
- Очистка от стоп слов
- Токенизация
- Лемматизация

Импортируем библиотеки для предобработка данных

```
import re
import string

import nltk
import pymorphy3
import wordcloud

stopwords = nltk.corpus.stopwords.words("russian") + [
    "хабр", "хабра"
]

punctuation = string.punctuation + "--"
morph = pymorphy3.MorphAnalyzer()
```

Создадим функции для удобной обработки текста

```
def remove_punctuations(text: str) -> str:
    return re.sub(f"[{punctuation}]", " ", text)
    return "".join([char for char in text if char not in punctuation])

def remove_digits(text: str) -> str:
    return "".join([char for char in text if not char.isdigit()])

def remove_latinic(text: str) -> str:
    return re.sub("[a-z]", "", text)

def remove_multiple_spaces(text: str) -> str:
    return re.sub(r"\s+", " ", text.replace("\n", " ").replace("\t", " "))

def tokenize(text: str) -> list:
    return re.split("\W+", text)

def lemmatize(tokenized_text: list) -> list:
    return [morph.parse(word)[0].normal_form for word in
    tokenized_text]

def remove_stop_words(tokenized_text: list) -> list:
    return [word for word in tokenized_text if word not in stopwords]
```

```
def join_to_string(tokenized_text: list) -> str:
    return " ".join(tokenized_text)
```

Поочерёдно применим все эти функции

```
df["preprocessed_text"] = df["text"].apply(lambda x:
remove_punctuations(x.lower()))
df.head()
```

```
text \
0 Привет, Хабр! Сегодня разберём, как реализовать...
1 Старший iOS-разработчик red_mad_robot Аня Коче...
2 Еще до начала MWC в Барселоне было очевидно, ч...
3 Взлом Instagram*-аккаунта – популярный запрос ...
4 Привет, Хабр! Меня зовут Кирилл Прямов, я мене...
```

```
preprocessed_text
0 привет хабр сегодня разберём как реализовать...
1 старший ios разработчик red mad robot аня коче...
2 еще до начала mwc в барселоне было очевидно ч...
3 взлом instagram аккаунта популярный запрос ...
4 привет хабр меня зовут кирилл прямов я мене...
```

```
df["preprocessed_text"] = df["preprocessed_text"].apply(lambda x:
remove_digits(x))
df.head()
```

```
text \
0 Привет, Хабр! Сегодня разберём, как реализовать...
1 Старший iOS-разработчик red_mad_robot Аня Коче...
2 Еще до начала MWC в Барселоне было очевидно, ч...
3 Взлом Instagram*-аккаунта – популярный запрос ...
4 Привет, Хабр! Меня зовут Кирилл Прямов, я мене...
```

```
preprocessed_text
0 привет хабр сегодня разберём как реализовать...
1 старший ios разработчик red mad robot аня коче...
2 еще до начала mwc в барселоне было очевидно ч...
3 взлом instagram аккаунта популярный запрос ...
4 привет хабр меня зовут кирилл прямов я мене...
```

```
df["preprocessed_text"] = df["preprocessed_text"].apply(lambda x:
remove_latinic(x))
df.head()
```

```
text \
0 Привет, Хабр! Сегодня разберём, как реализовать...
1 Старший iOS-разработчик red_mad_robot Аня Коче...
2 Еще до начала MWC в Барселоне было очевидно, ч...
```

```
3 Взлом Instagram*-аккаунта – популярный запрос ...
4 Привет, Хабр! Меня зовут Кирилл Прямов, я мене...
```

```
preprocessed_text
0 привет хабр сегодня разберём какреализовать...
1 старший разработчик аня кочешкова рассказы...
2 еще до начала в барселоне было очевидно что ...
3 взлом аккаунта популярный запрос в поисков...
4 привет хабр меня зовут кирилл прямов я мене...
```

```
df["preprocessed_text"] = df["preprocessed_text"].apply(lambda x:
remove_multiple_spaces(x))
df.head()
```

```
text \
0 Привет, Хабр! Сегодня разберём, какреализовать...
1 Старший iOS-разработчик red_mad_robot Аня Коче...
2 Еще до начала MWC в Барселоне было очевидно, ч...
3 Взлом Instagram*-аккаунта – популярный запрос ...
4 Привет, Хабр! Меня зовут Кирилл Прямов, я мене...
```

```
preprocessed_text
0 привет хабр сегодня разберём какреализовать те...
1 старший разработчик аня кочешкова рассказывает...
2 еще до начала в барселоне было очевидно что хо...
3 взлом аккаунта популярный запрос в поисковиках...
4 привет хабр меня зовут кирилл прямов я менедже...
```

```
df["preprocessed_text"] = df["preprocessed_text"].apply(lambda x:
tokenize(x))
df.head()
```

```
text \
0 Привет, Хабр! Сегодня разберём, какреализовать...
1 Старший iOS-разработчик red_mad_robot Аня Коче...
2 Еще до начала MWC в Барселоне было очевидно, ч...
3 Взлом Instagram*-аккаунта – популярный запрос ...
4 Привет, Хабр! Меня зовут Кирилл Прямов, я мене...
```

```
preprocessed_text
0 [привет, хабр, сегодня, разберём, какреализова...
1 [старший, разработчик, аня, кочешкова, рассказ...
2 [еще, до, начала, в, барселоне, было, очевидно...
3 [взлом, аккаунта, популярный, запрос, в, поиск...
4 [привет, хабр, меня, зовут, кирилл, прямов, я,...
```

```
%time
df["preprocessed_text"] = df["preprocessed_text"].apply(lambda x:
lemmatize(x))
df.head()
```

```
CPU times: total: 3min 44s
Wall time: 3min 47s
```

```
text \
0 Привет, Хабр! Сегодня разберём, как реализовать...
1 Старший iOS-разработчик red_mad_robot Аня Коче...
2 Еще до начала MWC в Барселоне было очевидно, ч...
3 Взлом Instagram*-аккаунта – популярный запрос ...
4 Привет, Хабр! Меня зовут Кирилл Прямов, я мене...
```

```
preprocessed_text
0 [привет, хабра, сегодня, разобрать, как реализо...
1 [старший, разработчик, аня, кочешкова, рассказ...
2 [ещё, до, начало, в, барселона, быть, очевидно...
3 [взлом, аккаунт, популярный, запрос, в, поиско...
4 [привет, хабра, я, звать, кирилл, прям, я, мен...
```

```
df["preprocessed_text"] = df["preprocessed_text"].apply(lambda x:
remove_stop_words(x))
df.head()
```

```
text \
0 Привет, Хабр! Сегодня разберём, как реализовать...
1 Старший iOS-разработчик red_mad_robot Аня Коче...
2 Еще до начала MWC в Барселоне было очевидно, ч...
3 Взлом Instagram*-аккаунта – популярный запрос ...
4 Привет, Хабр! Меня зовут Кирилл Прямов, я мене...
```

```
preprocessed_text
0 [привет, сегодня, разобрать, как реализовать, т...
1 [старший, разработчик, аня, кочешкова, рассказ...
2 [ещё, начало, барселона, очевидный, хотя, опер...
3 [взлом, аккаунт, популярный, запрос, поисковик...
4 [привет, звать, кирилл, прям, менеджер, развит...
```

```
%time
df["preprocessed_text"] = df["preprocessed_text"].apply(lambda x:
join_to_string(x))
df.head()
```

```
CPU times: total: 31.2 ms
Wall time: 36.9 ms
```

```
text \
0 Привет, Хабр! Сегодня разберём, как реализовать...
1 Старший iOS-разработчик red_mad_robot Аня Коче...
2 Еще до начала MWC в Барселоне было очевидно, ч...
3 Взлом Instagram*-аккаунта – популярный запрос ...
4 Привет, Хабр! Меня зовут Кирилл Прямов, я мене...
```

```
preprocessed_text
```

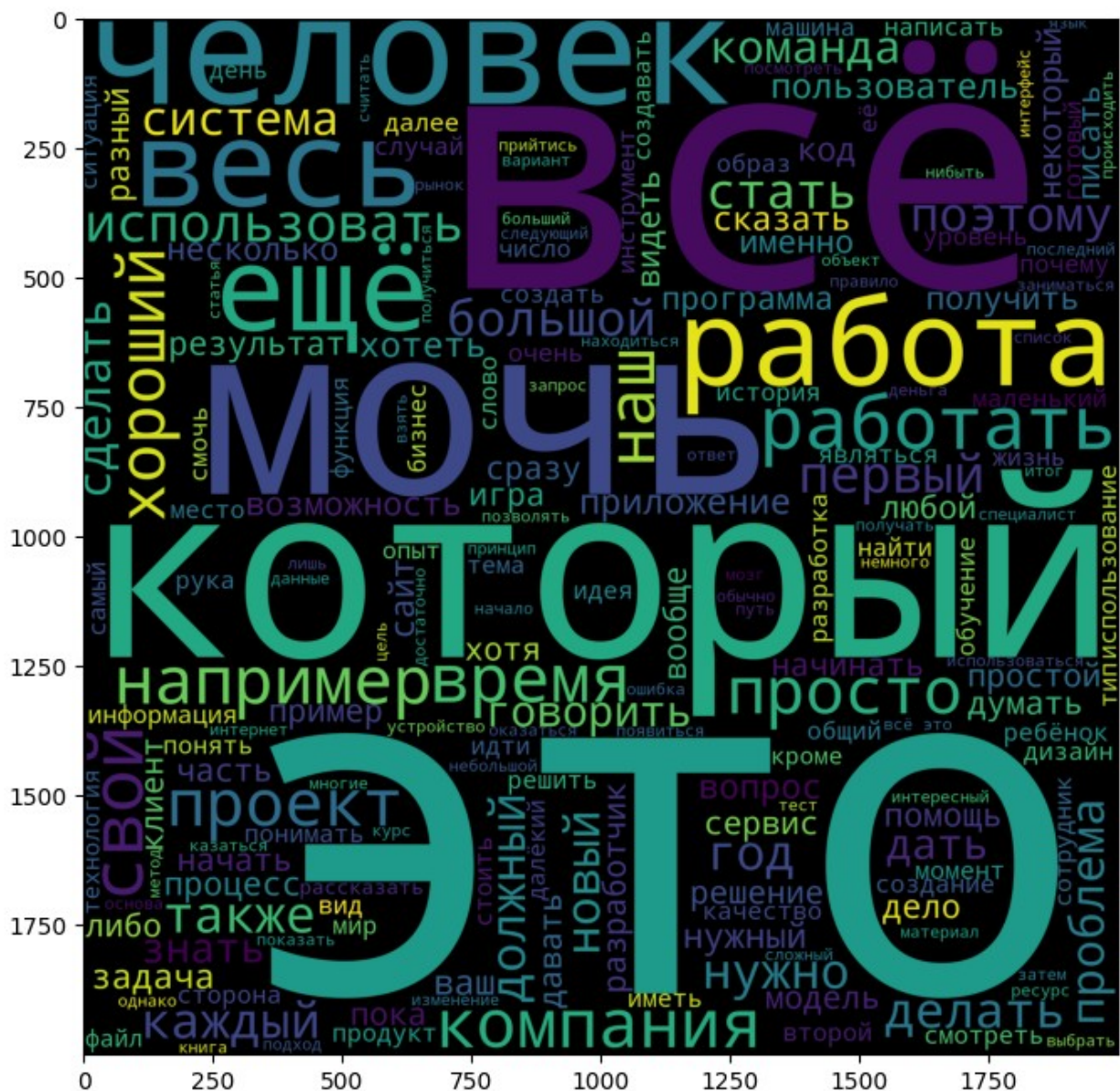
```
0 привет сегодня разобрать как реализовать тестир...
1 старший разработчик аня кочешкова рассказывать...
2 ещё начало барселона очевидный хотя оператор п...
3 взлом аккаунт популярный запрос поисковик поэт...
4 привет звать кирилл прям менеджер развитие дав...
```

Теперь, после предварительной обработки текста можно отобразить облако слов и расширить список слов.

```
%%time
from wordcloud import WordCloud

text = " ".join(word for word in df["preprocessed_text"])

plt.figure(figsize=(8, 8))
plt.imshow(WordCloud(background_color="black", width=2000,
height=2000, random_state=42).generate(text))
plt.show()
```



CPU times: total: 10.7 s  
Wall time: 10.9 s

## Векторизация

Так как модели кластеризации и классификации умеют работать только с цифрами и не умеют работать с текстом, текст необходимо векторизировать.

Важно учитывать не только частоту слов но и их важность, так как многие слова в тексте не несут смысловой нагрузки. Для векторизации нам подойдёт модель TfidfVectorizer



```
%time
from sklearn.feature_extraction.text import TfidfVectorizer

# Создаём и обучаем модель
tfidf_vectorizer = TfidfVectorizer()
tfidf_matrix = tfidf_vectorizer.fit_transform(df["preprocessed_text"])

CPU times: total: 766 ms
Wall time: 767 ms

tfidf_matrix.shape

(1107, 51326)

tfidf_vectorizer.get_feature_names_out()[100]

array(['ée', 'öü', 'еєһый', 'шəəшозэтый', 'аа', 'aaa', 'aaaa',
'aaaaa',
      'aaaaaaaaaaaaa', 'aaaaaaaaaaaaaaaaaaaaa', 'аарон', 'аба', 'абаджи',
      'абак', 'абасмирзоев', 'абб', 'аббатта', 'аббревиатура',
      'авгдеёзийклмнопрстуфхцьыэ', 'абгарян', 'абдул',
'абдильманов',
      'абель', 'аберрация', 'абзац', 'абзацный', 'абидно', 'абие',
      'абий', 'абитуриент', 'абк', 'аблация', 'аблютофобия', 'абон',
      'абонемент', 'абонент', 'абонентский', 'абориген', 'аборт',
      'абразив', 'абразивность', 'абракадабра', 'абрам', 'абрамзона',
      'абрамов', 'абрамович', 'абрахам', 'абрин', 'абрином',
'аброботка',
      'абрус', 'абсолютный', 'абсолют', 'абсолютизировать',
'абсолютно',
      'абсолютность', 'абсолютный', 'абсорбировать',
'абсорбироваться',
      'абсорбция', 'абстрагирование', 'абстрагировать',
      'абстрагироваться', 'абстрактно', 'абстрактность',
'абстрактный',
      'абстракционизм', 'абстракционизмажнейросетевой',
'абстракционист',
      'абстракция', 'абсурд', 'абсурдно', 'абсурдность', 'абсурдный',
      'абсцисса', 'абу', 'абуз', 'абузоустойчивость', 'абуталим',
      'абъять', 'абызов', 'абышко', 'абьюзить', 'ава', 'авадон',
      'авангард', 'авангардист', 'авангардный', 'авангардпласт',
'аванс',
      'авансовый', 'авансом', 'авантюра', 'авантюризм', 'авантюрист',
      'авантюрный', 'аварийка', 'аварийность', 'аварийный',
'авария'],
dtype=object)
```

# Кластеризация

Теперь, так как текст векторизован, можно кластеризовать данные



Для кластеризации будем использовать модель KMeans. Алгоритм k-means используют для группировки объектов в кластеры на основе их схожести. В основе работы k-means лежит принцип минимизации расстояния между объектами внутри одного кластера.

Основные преимущества k-means это простота и быстрота реализации, а так же эффективность при работе с большими наборами данных

Для начала уменьшим размерность матрицы. Это упростит обучение моделей путём уменьшения вычислительной сложности

```
from sklearn.decomposition import NMF
from sklearn.cluster import KMeans

nmf_model = NMF(n_components=7, random_state=0)
W = nmf_model.fit_transform(tfidf_matrix)

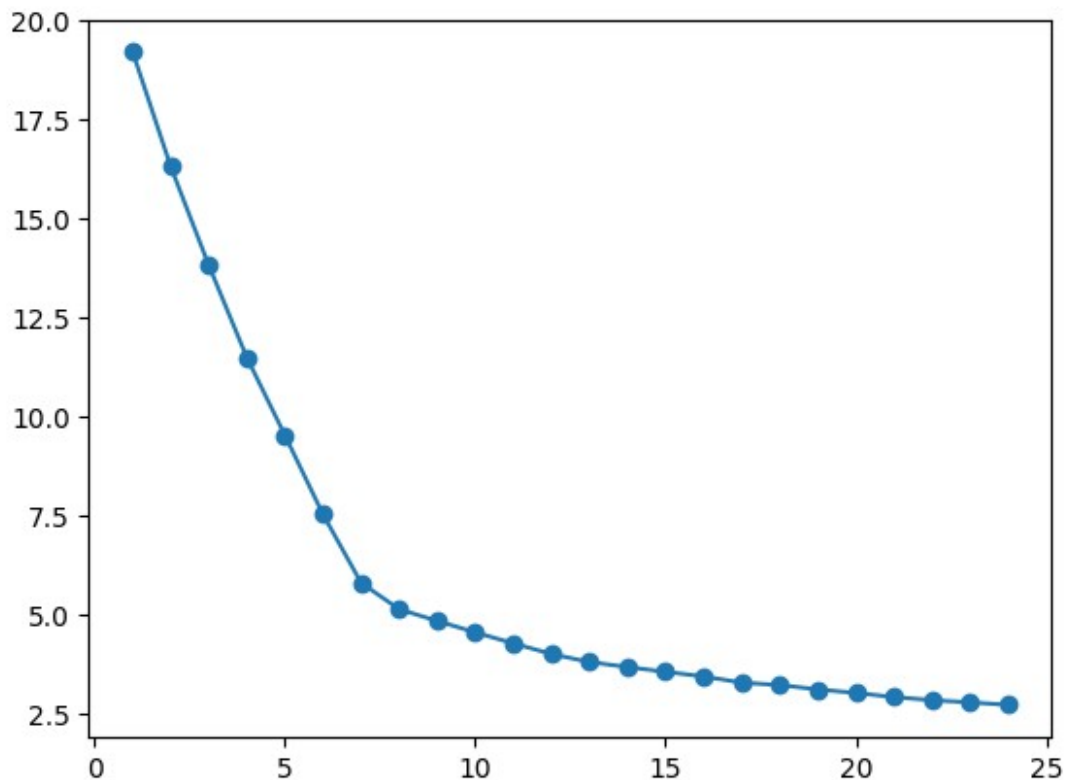
%%time

cluster_range = range(1, 25)
wcss = []

# Обучаем модели несколько раз
for i in cluster_range:
    # Создаём и обучаем модель
    kmeans = KMeans(n_clusters=i, n_init=10, init="k-means++",
                    random_state=42).fit(W)
    wcss.append(kmeans.inertia_)

CPU times: total: 10.1 s
Wall time: 1.41 s

plt.plot(cluster_range, wcss, marker="o")
plt.show()
```



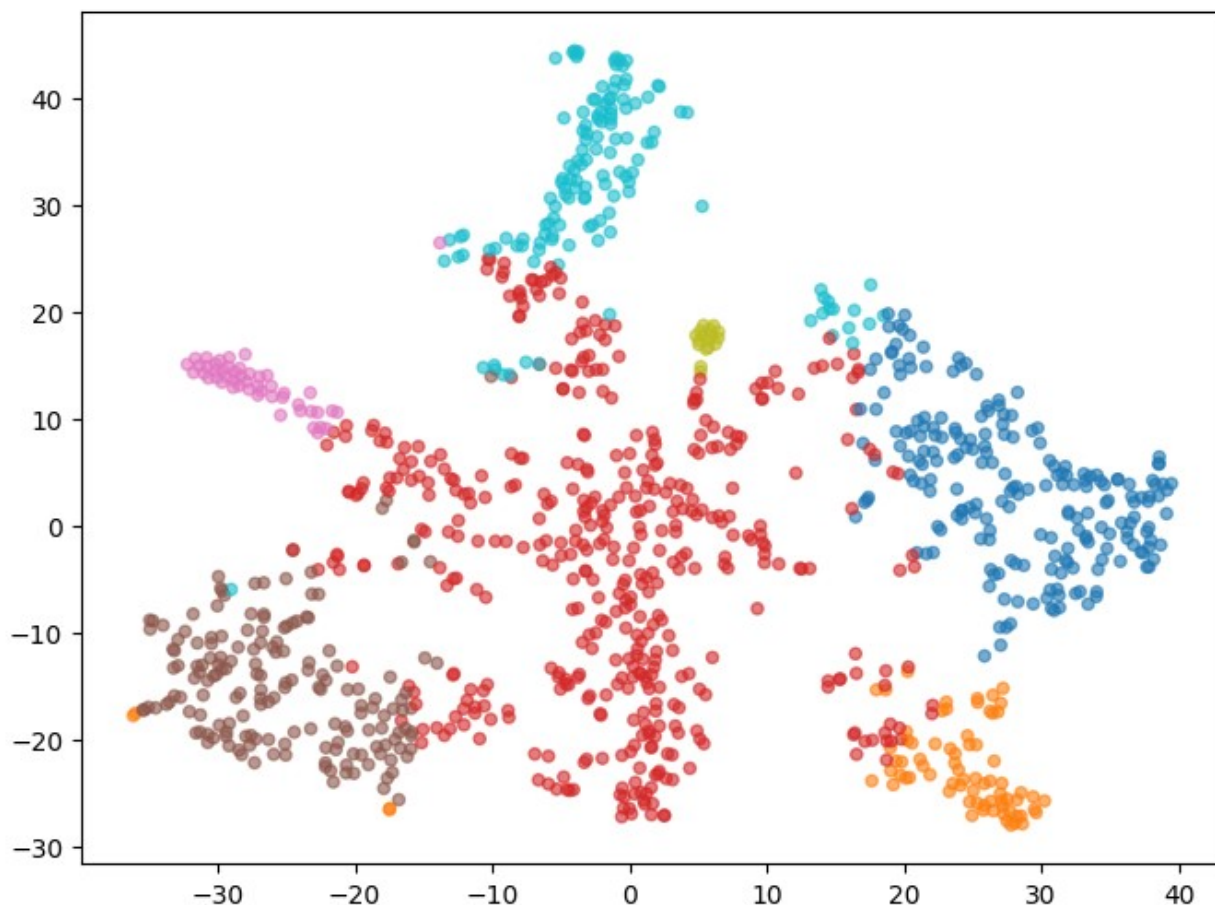
```
kmeans = KMeans(n_clusters=7, n_init=10, random_state=42)
labels = kmeans.fit_predict(W)
centers = kmeans.cluster_centers_
```

Визуализируем данные, каждый цвет прикреплен к кластеру. Так же отобразим центры кластеров красными крестиками

```
from sklearn.manifold import TSNE

x_tsne = TSNE(n_components=2, random_state=42).fit_transform(W)

plt.figure(figsize=(8, 6))
plt.scatter(x_tsne[:, 0], x_tsne[:, 1], c=labels, s=20, cmap="tab10",
            alpha=0.6)
plt.show()
```



По данному графику распределения видно, что 1 кластер расположен на большом удалении и содержит малое количество записей. Добавим кластеры в исходный набор данных

Добавим информацию о принадлежности к кластеру в набор данных

```
df["cluster"] = labels
df.sample()
```

	text \	
684	С геймификацией быта большинство из нас впервые...	
	preprocessed_text	cluster
684	геймификация быт большинство впервые столкнуться...	2

Выведем распределение кластеров

```
df["cluster"].value_counts()

cluster
2      453
0      208
```

```
3    170
6    134
1     76
4     49
5     17
Name: count, dtype: int64
```

Теперь выведем несколько записей для каждого кластера

```
pd.set_option("display.max_colwidth", 100)

for i in range(7):
    print(f"cluster: {i}: ")
    print(f"{df[df['cluster'] == i].sample(1)[['text']].values}\n")

cluster: 0:
[['Несколько дней назад команда инженеров Facebook отличилась – ее удостоили награды Most Influential POPL Paper Award. В среде специалистов по машинному обучению это весьма почетно. Награду вручили за работу Compositional Shape Analysis by Means of Bi-abduction, которая раскрывает нюансы Project Infer. Сам проект предназначен для обнаружения и ликвидации багов в коде мобильного приложения перед его деплоем.\n\nБаги в программном обеспечении для мобильных устройств обходятся очень дорого как разработчикам, так и пользователям. Что касается первых, то обнаружение проблемы в уже размещенном в каталогах приложении – ночной кошмар для любого специалиста. Конечно, софт тестируется, работу программ проверяют по определенным шаблонам. Но чаще разработчики не могут предусмотреть всего, и баги в приложении обнаруживаются уже после деплоя.\n\nНапоминаем: для всех читателей «Хабра» – скидка 10 000 рублей при записи на любой курс Skillbox по промокоду «Хабр».\n\nSkillbox рекомендует: Онлайн-курс «Аналитик данных на Python».\n\nProject Infer сканирует код мобильных приложений и позволяет найти разного рода проблемы, паттерны которых хранятся в базе (а она все время обновляется). Сам проект был представлен три года назад. Почти сразу после анонса Facebook открыл код, после чего его стали использовать в таких компаниях, как Amazon Web Services, Spotify и Uber.\n\nКак это работает?\n\nProject Infer использует специализированный набор алгоритмов для анализа работы кода. В исходном коде любого крупного приложения могут быть миллионы комбинаций, которые приводят к появлению ошибок. Традиционные процедуры анализа кода не могут помочь обнаружить все. Infer от Facebook, самообучаясь, расширяет свою базу, так проект позволяет обнаружить массу проблем в коде.\n\nВ общем смысле процесс работы Facebook Infer можно разделить на два этапа: сбор данных и их анализ. Жизненный цикл (lifecycle) также разделяется на две части: глобальную и дифференциальную.\n\nНа этапе сбора данных Infer переводит исходный код в собственный язык. Этап анализа посвящается изучению мельчайших деталей структуры кода, которые могут потенциально привести к появлению ошибки. Если Infer встречает уже знакомое сочетание
```

факторов, идентифицируемое как паттерн ошибки, анализ останавливается для конкретного метода или функции, но остальные методы и функции продолжают анализироваться. Вот пример работы Infer.

С точки зрения исполнения Infer может работать в двух модальностях – Global и Differential, как и говорилось выше. В первом случае Infer анализирует все файлы. Для проекта, который компилируется с использованием Gradle, запуск Infer производится командой

```
infer run -- gradle build
```

Дифференциальный процесс используется в инкрементных системах сборки, характерных для мобильных приложений. В этом случае нужно сначала запустить сбор данных Infer, чтобы получить все команды компиляции, а затем провести анализ только изменений. Для этого используется такой набор команд:

```
gradle clean
infer capture -- gradle build
edit some/File.java
# make some changes to some/File.java
infer run -- reactive -- gradle build
```

Отчеты Infer можно просмотреть при помощи команды

```
InferTraceBugs
infer run -- gradle build
inferTraceBugs
```

Основа Project Infer от Facebook основан на двух новых математических методах: логике разделения и Bi-abduction.

Ключевая особенность логики разделения – возможность локальных рассуждений (local reasoning). Она появилась благодаря наличию в утверждениях пространственных связей (англ. spatial connectives) между частями кучи. В этом случае нет необходимости учитывать весь объем памяти на каждом из этапов.

Основной элемент логики разделения – оператор  $*$  (и отдельно), который называется разделяющим соединением. Формула  $X \mapsto Y * Y \mapsto X$  может быть прочитана как « $X$  указывает на  $Y$ , а отдельно  $Y$  указывает на  $X$ », что очень похоже на то, как работают указатели памяти.

В контексте Infer Bi-abduction можно рассматривать как метод логического вывода, который дает платформе возможность обнаруживать свойства, касающиеся поведения независимых частей кода приложения. Bi-abduction совместно выводит антифреймы (отсутствующие части состояния) и фреймы (те части, которые не затронуты операцией). Математически проблема Bi-abduction выражается с использованием следующего синтаксиса:  $A * ? \text{Antiframe} \vdash B * ? \text{Frame}$ .

В Infer от Facebook метод дает возможность вывести спецификации pre/post из чистого кода при условии, что мы знаем спецификации для примитивов на базовом уровне кода.

Создание FI стало возможно благодаря анализу работы специалистов по машинному обучению, который проводился много лет. В ходе работы над Infer были опубликованы такие ключевые для всей сферы статьи:

- Compositional Shape Analysis by means of Bi-abduction. Как раз за эту работу была получена премия, о которой говорится выше. Работа знакомит читателя с композиционным анализом формы. Это дополнение к традиционному анализу формы (shape analysis), дающее возможность применить метод для анализа исходного кода приложений.
- A Local Shape Analysis Based on Separation Logic: эта статья описывает логику разделения в качестве механизма анализа исходного кода приложений. Авторы показывают возможность изучения отдельных ячеек в куче памяти, без изучения всей кучи целиком. Таким образом, определенные ячейки составляют связанный список и без полного анализа.
- Smallfoot: Modular Automatic Assertion Checking with Separation Logic: в этой работе описывается предшественник Facebook

Infer, который называется Smallfoot.\n AL: A new declarative language for detecting bugs with Infer: AL позволяет любому разработчику проектировать новые чекеры без полного понимания внутренней кухни Infer. AL – это декларативный язык.\n Moving Fast with Software Verification: Наконец, статья, которая раскрывает, как Facebook использует Project Infer для собственных нужд. В документе рассказывается о том, как разработчики Facebook интегрировали Infer в свой процесс разработки для обеспечения статического анализа для мобильных приложений, таких как Instagram, Facebook Messenger и приложения Facebook для Android и iOS.\n\r\nПока что Infer может использоваться лишь для мобильных приложений. Но некоторые его принципы применимы к приложениям общего назначения. Возможно, в будущем возможности Infer станут шире, и с его помощью разработчики смогут анализировать десктопные или серверные приложения.\n\nSkillbox рекомендует:\n\n\nПрактический курс «Мобильный разработчик PR0».\n\nонлайн-курс «Профессия Frontend-разработчик».\n\nПрактический годовой курс «PHP-разработчик с 0 до PR0».']]

cluster: 1:

[[ 'Делимся с вами подборкой вебинаров на тему разработки игр. Вы узнаете, как сделать простую консольную игру на PHP, 3D-арканойд на движке Unreal Engine 4, космическую аркаду и AR-приложение на Unity. Заходите под кат – будет интересно.\n\nЗнакомство с PHP: пишем простую консольную игру\r\nНа вебинаре вы познакомитесь с базовой информацией о PHP, где и как часто этот язык применяется. Узнаете об основных конструкциях, синтаксисе и операторах + расскажем, почему PHP лучше начать изучать именно на консольных приложениях.\n\n\nПо теме:\r\n – 5 заблуждений о разработке на PHP;\r\n – Выбираем язык программирования: что нужно знать о PHP;\r\n – Практический онлайн-курс «PHP-разработчик».\n\nСоздаём игру на Unity\r\nВ этом видео мы будем учиться писать простой скролл-шутер: создавать сцены, управлять камерой и космическим кораблем. Напишем пару скриптов на C# и разберёмся с ассетами: моделями, текстурами и эффектами. Но начнём мы со знакомства с Unity – одним из самых популярных игровых движков.\n\n\nПо теме:\r\n – Подборка «Это всё – Unity. История самых заметных игр на движке с 2010 по 2019»;\r\n – Сыграйте в игру от студентов Skillbox про главные навыки во время самоизоляции;\r\n – Бесплатный интенсив «Ваша первая игра на Unity за 3 вечера» 18-20 мая.\n\n\nСобираем настоящий космос\r\nВторая часть вебинара про создание игр на Unity. В предыдущем видео мы научили корабль двигаться, теперь нужно добавить ему противников и препятствия, а также набросать пару взрывов. Также мы добавим звуки в игру и организуем подсчёт очков.\n\n\nПо теме:\r\n – Подведение итогов интенсива и ревью работ участников;\r\n – Руководство «Создаём простую 3D-гонку на Unity»;\r\n – Практический онлайн-курс «Разработчик игр на Unity».\n\nСоздаём приложение дополненной реальности\r\nНа вебинаре мы будем осваивать Unity в разрезе создания AR-приложения: рассмотрим базовые инструменты, разместим объекты в сцене и познакомимся с компонентами.\n\n\n\nБиблиотека 8th wall для приложений дополненной реальности\n\r\n

Второй мастер-класс посвящён работе с библиотекой 8th wall для приложений дополненной реальности. Мы познакомимся с набором инструментов и библиотек, рассмотрим примеры их использования и займёмся удалённой отладкой приложения на мобильном устройстве.

По теме:

- Подведение итогов интенсива и ревью работ участников
- Практический онлайн-курс «Разработчик VR & AR»
- Разработка игры на Unreal Engine 4

На вебинаре мы научимся создавать персонажа, разберёмся с настройкой управления скорости объекта, спроектируем окружение и его реакцию на действия игрока. Но, как обычно, начнём со знакомства с редактором Unreal Engine 4.

Собираем проект

В следующей части мастер-класса мы продолжим развивать персонажа и создадим движущуюся платформу. Научимся спавнить новый объект, добавим условия для его уничтожения и возможность проиграть. Также будем работать со звуком и соберём проект в готовую игру.

По теме:

- Подведение итогов интенсива и ревью работ участников
- Расписание наших бесплатных интенсивов
- Практический онлайн-курс «Middle-разработчик на Unreal Engine 4»

cluster: 2:

Самые необычные игровые аксессуары в истории индустрии

Современная игровая индустрия напоминает море во время штиля: крупные издатели и платформодержатели предпочитают оставаться на плаву, эксплуатируя проверенные временем рецепты привлечения целевой аудитории, тогда как подлинная жизнь бурлит в глубинных водах инди-сегмента. Это касается не только самих игр, но и аппаратного обеспечения. Вспомните, когда последний раз Sony или Microsoft радовали нас чем-то необычным? Да что говорить: даже «инновационный» Switch, как бы его ни расхваливала пресса, является не чем иным, как дальнейшим эволюционным развитием идей, заложенных еще во времена создания первого Game Boy. Однако на заре становления игровой индустрии даже крупные компании частенько пускались во все тяжкие, создавая поистине безумные устройства в попытках отвоевать свой кусочек рынка. О наиболее интересных из них мы и расскажем в сегодняшнем материале.

Киберпанк, который мы не заслужили: ретрофутуризм для геймеров

Одним из самых частых объектов для инженерных опытов закономерно становились игровые контроллеры. С помощью необычных устройств ввода платформодержатели стремились предложить игрокам принципиально новый пользовательский опыт, чтобы затмить конкурентов и привлечь внимание целевой аудитории к своей консоли. Далеко не все из подобных проектов оказались удачными (а если говорить совсем откровенно, выстреливали лишь единицы), однако от этого они не становились менее интересными.

Atari Mindlink: как тебе такое, Илон Маск?

Прототип Atari MindLink – гарнитуры, позволяющей играть «силой мысли»

После кризиса игровой индустрии 1983 года разработчики консолей стремились всеми силами привлечь внимание потребителей к своей продукции, в том числе с помощью необычных геймерских гаджетов. Одним из таких проектов был Atari MindLink, позиционировавшийся как первый контроллер, позволяющий управлять персонажами видеоигр силой мысли. На самом деле регистрировать биоэлектрическую активность мозга, преобразуя сигналы ЭЭГ в команды



управления, футуристического вида устройство не умело: вместо этого инфракрасные сенсоры наголовного модуля лишь фиксировали движение мышц лба. К сожалению (а может быть, и к счастью), необычный девайс так и не добрался до розницы: проект MindLink пришлось свернуть на этапе тестирования, а разрабатываемые для него игры (Bionic Breakthrough, Telepathy и Mind Maze) – отменить. Все дело в том, что инженеры Atari так и не смогли добиться достаточной точности, в результате чего геймплей с таким устройством превращался в настоящую пытку. Кроме того, при длительном использовании у MindLink проявлялся серьезный побочный эффект: из-за постоянного напряжения мышцы игрока спазмировались, что вызывало сильную головную боль.

Robotic Operating Buddy – ваш самый надежный тиммейт

Robotic Operating Buddy – робот-напарник для игр\ха0Многие эксперты называют Robotic Operating Buddy спасителем игровой индустрии, и вполне заслуженно. Все дело в том, что в посткризисный период дистрибьюторы бытовой электроники стали относиться к игровым консолям с большим скепсисом, из-за чего у Nintendo возникли серьезные проблемы с продвижением NES на мировом рынке. Во многом именно R.O.B. помог японской корпорации завоевать доверие продавцов и популяризовать домашние игровые системы: инновационный робот, способный помочь в прохождении видеоигр или стать напарником в кооперативе (во всяком случае, именно это обещали рекламные развороты журналов), быстро «завирусился» в СМИ, став одним из самых желанных подарков для детей конца 80-х.

R.O.B. рядом с фирменной коробкой

Хотя на самом деле R.O.B. был способен играть лишь в специально разработанные для него проекты вроде Gyromite и Stack-Up, необычная игрушка имела оглушительный успех, во многом поспособствовав популяризации NES на американском и европейском рынках. Но слава Robotic Operating Buddy оказалась недолгой: концепт робота-напарника так и не получил дальнейшего развития, что во многом обусловлено излишней честностью его разработчиков. Дело в том, что R.O.B. работал именно так, как и было заявлено: он не получал данные напрямую с консоли по проводу, а действительно «смотрел» на экран телевизора, считывая визуальную информацию и принимая на ее основе решения о дальнейших действиях согласно прописанным алгоритмам. Все это происходило крайне медленно и накладывало ограничения на перечень поддерживаемых игр: R.O.B. прекрасно подходил для решения пазлов, однако не мог играть в такие популярные жанры, как платформеры, файтинги или гонки.

Гейл Тилден, возглавлявшая в те времена отдел маркетинга Nintendo of America, вспоминала: «Смотреть, как R.O.B. играет, – это примерно то же самое, что наблюдать за растущей травой: все происходило очень и очень медленно. А нам требовалось каким-то образом сделать процесс захватывающим. Для меня как маркетолога это было одним из самых серьезных вызовов за всю карьеру». Чтобы добавить геймплею динамики, для R.O.B. создали необычный двухклавишный контроллер, а его поведение усложнили: вместо того, чтобы нажимать на кнопки манипуляторами, робот стал устанавливать на них вращающиеся на огромной скорости волчки, предварительно раскручивая их на специальном стенде. Как именно легендарный инженер Nintendo Гумпей Ёкой, которому и принадлежит данное изобретение, дошел до столь причудливой схемы

управления, история умалчивает, однако это выглядело достаточно эффектно и сыграло важную роль в продвижении проекта. Но время шло, NES становилась все популярнее, а после релиза Super Mario Bros. консоль и вовсе стали расхватывать, словно горячие пирожки, и в какой-то момент R.O.B. оказался попросту никому не нужен, поскольку развить проект в нечто большее технологии того времени попросту не позволяли. R.O.B. и его миниатюрная amiibo-копия. Впрочем, в Nintendo не забыли о своем спасителе. В разное время R.O.B. появлялся в качестве игрового персонажа в различных частях Super Mario Kart, Super Smash Bros. и в ряде других проектов. Кроме того, в честь Robotic Operating Buddy Nintendo выпустила специальную серию фигурок Amiibo. Power Glove – Перчатка Бесконечности от Nintendo. Перчатка-геймпад Power Glove для Nintendo Entertainment System. Помимо R.O.B., Nintendo пробовала свои силы и в создании мышь-контроллеров. Результатом изысканий японской корпорации стала выпущенная в 1989 году перчатка Power Glove, которая позволяла управлять происходящим на экране при помощи движения руки. Метод отслеживания движений, используемый в Power Glove, был довольно остроумен. В саму перчатку была встроена пара ультразвуковых передатчиков, передающих звук на частоте 40 кГц. УЗ-излучение фиксировалось специальным модулем, который требовалось закрепить на телевизоре: миниатюрное устройство определяло на основе полученных данных положение руки в пространстве и генерировало управляющие сигналы. К продвижению Power Glove в Nintendo подошли весьма рационально. Для перчатки были выпущены два стартовых эксклюзива – 3D-головолонка Super Glove Ball и бит-ем-ап Bad Street Brawler, которые можно было без проблем пройти на обычном геймпаде, однако при использовании инновационного контроллера в играх появлялись дополнительные движения. Аналогичные идеи лежали в основе Glove Pilot, Manipulator Glove Adventure и Tech Town, однако эти игры так и не добрались до релиза, поскольку Power Glove провалилась в продажах: аппарат был сложен в использовании, да и цена \$80 не особо способствовала росту его популярности. Лишь спустя десятилетия Хакермен смог раскрыть истинный потенциал Power Glove. Зато Power Glove оставила отпечаток в массовой культуре. Так, например, отсылку на необычный контроллер можно встретить в фильме ужасов «Кошмар на улице Вязов 6», где похожим девайсом пользуется Фредди Крюгер, а в короткометражной пародии на боевики 80-х «Кунг-фьюри» гениальный Хакермен использует уже настоящий Power Glove для взлома пространственно-временного континуума. Konami LaserScope: мы играли до хрипоты «Да стреляй ты уже, зараза!» – примерно так выглядели игры с Konami LaserScope. Задолго до того, как бросить все силы на выпуск автоматов пачинко, Konami активно экспериментировала не только в сфере разработки видеоигр, но и в создании необычных контроллеров. Magnum opus инженеров японской компании стала напольная гарнитура LaserScope, увидевшая свет в 1991 году и выпущенная специально для NES. Необычный девайс разрабатывался специально для Laser Invasion, однако, поскольку его оптический модуль функционировал точно так же, как и NES Zapper, контроллер можно было использовать с любой игрой, поддерживающей световой пистолет. В Laser Invasion игроку предстояло взять на себя роль пилота боевого

вертолета. С помощью гарнитуры LaserScope геймер мог наводить орудия на врагов, а чтобы сделать выстрел, в микрофон было необходимо скомандовать «Огонь!». Laser Invasion – первый и единственный эксклюзив для Konami LaserScope. Но если система наведения работала с достаточно высокой точностью, то с распознаванием голоса у LaserScope были серьезные проблемы: чтобы сделать выстрел, в миниатюрный микрофон приходилось буквально орать, так что длительные игровые сессии могли с легкостью привести к потере голоса из-за перенапряжения голосовых связок. Как следствие, гарнитура не снискала популярности среди геймеров и провалилась в продажах.

**U-Force: ловкость рук и никакого толку**  
U-Force – система управления жестами. В 1989 году компания Broderbund представила инновационный игровой контроллер для Nintendo Entertainment System, позволяющий управлять происходящим на экране телевизора с помощью жестов. Устройство представляло собой складную панель с инфракрасными сенсорами, которые определяли положение рук игрока, преобразуя их движения в управляющие сигналы. Столь революционный девайс был обречен на успех. И U-Force действительно попал в топ... худших контроллеров за всю историю индустрии по версии IGN. Из-за низкой точности любая игра с U-Force превращалась в кромешный ад, поскольку буквально каждый жест приходилось повторять по несколько раз. Для динамичных игр вроде файтингов и бит-ем-апов контроллер не подходил в принципе: выполнить эффектное комбо на таком своеобразном геймпаде было не под силу даже киберспортсменам.

**Sega Activator – лучшее лекарство от гиподинамии**  
Моушн-контроллер Sega Activator в сборе. Другой неудачный эксперимент в области создания моушн-контроллеров предприняла в 1993 году Sega. Необычная система управления представляла собой напольный восьмиугольник, по периметру которого были расположены инфракрасные сенсоры. Согласно задумке, игрок вставал в центр восьмиугольника и имитировал различные движения (например, удары руками и ногами во время игры в файтинг), а контроллер должен был преобразовывать их в управляющие сигналы. На старте продаж девайс поддерживал такие игры, как Best of the Best: Championship Karate, Eternal Champions, Greatest Heavyweights, Streets of Rage 2 и 3, Mortal Kombat и Street Fighter 2. Также Sega планировала выпустить два эксклюзива, Air Drums и Bounty Hunter, призванные в полной мере раскрыть потенциал моушн-контроллера. В теории звучало красиво, однако на деле Sega Activator оказался абсолютно непредсказуем. Устройство часто ошибалось, путая удары руками и ногами, и реагировало даже на тени от конечностей игрока, из-за чего любое движение приводило к неожиданным результатам, а уж о том, чтобы крутить хитроумные комбинации, и вовсе не могло быть речи. И хотя данный контроллер сумел занять достойное место в истории индустрии, из-за оглушительного провала разработка так и не получила дальнейшего развития.

**Roll & Rocker: поиграл, очнулся – гипс**  
Контроллер-балансир Roll & Rocker. В отличие от систем управления жестами, контроллер-балансир Roll & Rocker, выпущенный для NES в 1989 году, преследовал несколько иные цели. Данный девайс был призван радикально решить нехватку кнопок на стандартном геймпаде, а заодно заставить заядлых геймеров хоть немного двигаться во время игровых сессий. Устройство

представляло собой напольный балансир, встав на который вы могли управлять движениями персонажа, наклоняя платформу в ту или иную сторону. Сама по себе идея была не так уж и плоха, но, как обычно, подкачала реализация. Хотя контроллер был куда точнее, чем упомянутый выше Sega Activator, он все равно обладал значительно меньшей отзывчивостью, нежели крестовина традиционного геймпада. К тому же Roll & Rocker получился весьма травмоопасным: из-за крутизны наклонов игрок имел все шансы вписаться головой в телевизор или же вывихнуть лодыжку. А поскольку родители не горели желанием рисковать здоровьем своих детей, необычный контроллер потерпел фиаско.

Reflex Control Paradox – enfant terrible трекбола и геймпада

Геймпад со встроенным трекболом Paradox от компании Reflex Control

При всем своем удобстве, геймпады проигрывают компьютерным мышкам в скорости и точности, что особенно хорошо заметно при игре в шутеры и RTS. Во времена PlayStation 2 решить данную проблему попыталась компания Reflex Control, выпустив для игровой консоли гибридный геймпад Paradox, в котором вместо правого рычажка красовался массивный трекбол. Как нетрудно догадаться, изобретение Reflex Control не имело коммерческого успеха. Хотя Paradox отлично показывал себя в стратегиях, шутерах и тактических играх, контроллер оказался бесполезен в проектах, требовавших использования сразу двух аналоговых стиков. В дальнейшем, когда стандартом консольных стрелялок окончательно стала система «умного» наведения, а прочие целевые жанры почти исчезли из игротек платформодержателей, необходимость в подобных девайсах и вовсе отпала.

Chameleon X1 – не мышонок, не лягушка

Chameleon X1 – помесь геймпада и компьютерной мышки

В 2011-м подвиг Reflex Control попыталась повторить Shogun Bros., задумав раз и навсегда стереть грань между консольным и ПК-геймингом с помощью гибридной мыши-геймпада Chameleon X1. Характеристики «грызуна» были вполне стандартными: сенсор чувствительностью 1600 DPI с возможностью выбора разрешающей способности и 5 функциональных кнопок (ЛКМ, ПКМ, колесико мыши + 2 боковые программируемые клавиши) мало кого могли удивить даже 10 лет назад. Но стоило перевернуть Chameleon X1 кверху брюхом – и заурядная игровая мышь превращалась в миниатюрный 14-кнопочный геймпад, у которого была даже пара аналоговых стиков! Разумеется, адекватно работать такой гибрид не мог по определению. Из-за наличия органов управления на днище и специфического рельефа основания мышка имела отвратительный баланс и очень плохо скользила по коврику, что негативно сказывалось на скорости и точности прицеливания в шутерах. Геймпад также не отличался эргономикой: микроскопические размеры, отсутствие курков и бамперов на их законных местах и крайне неудобные стики, к тому же не отличающиеся хорошей чувствительностью, превращали игру в любой слэшер или файтинг в настоящую пытку. Как следствие, «революционный» девайс провалился в продажах, ведь мало кому из геймеров улыбалось потратить \$55 на неудобную мышь и неполноценный контроллер.

eDimensional Access Controller: геймплей на кончиках пальцев

eDimensional Access Controller можно было кастомизировать как вашей душе угодно

В 2008 году инженер и консольный моддер Бен Хек представил eDimensional Access Controller – уникальный в своем роде

контроллер для людей с ограниченными возможностями. Необычный геймпад состоял из основания с пятью типовыми разъемами и нескольких модулей управления с кнопками, крестовиной и аналоговыми стиками, которые пользователь мог располагать так, как ему удобно. Контроллер поддерживал ПК и игровые консоли PlayStation 2 и PlayStation 3, для подключения к которым в комплекте поставлялись специальные переходники. К сожалению, на деле eDimensional Access Controller оказался вовсе не таким уж удобным и универсальным, как мечтал его создатель. Комфортно пользоваться геймпадом могли лишь обладатели крупных ладоней и достаточно длинных пальцев, поскольку гнезда для съемных модулей были расположены слишком далеко друг от друга. По этой же причине контроллер не подходил больным ревматоидным артритом или людям с посттравматическими контрактурами, которые не могли полностью раскрыть ладонь из-за ограничений подвижности в суставах. Проблему можно было бы решить разделением геймпада на две или три независимые части: так владелец девайса получил бы возможность самостоятельно регулировать расстояние между модулями управления. Но, поскольку eDimensional Access Controller провалился в продажах, проект так и не получил дальнейшего развития. Геймпад в нагрузку: эксклюзивные контроллеры для видеоигр «Современные проблемы требуют современных решений», а уникальный геймплей – уникальных контроллеров. Стремясь предложить игрокам новый опыт, разработчики зачастую загоняли сами себя в угол, поскольку возможностей обычных геймпадов в какой-то момент оказывалось уже недостаточно для реализации всех задумок. Тогда на помощь программистам приходили инженеры, создавая нестандартные девайсы для конкретных проектов. \ха0Скейт для Tony Hawk: Ride Издание Tony Hawk: Ride для игровой консоли Xbox 360 со скейтом-контроллером в комплекте В 2009 году Activision решилась на необычный эксперимент, выпустив игру Tony Hawk: Ride – симулятор скейтбордиста, в котором геймеру предлагалось управлять своим персонажем с помощью контроллера-скейта, оснащенного датчиками движения. По замыслу издателя, такой контроллер должен был обеспечить большее погружение в игровой процесс, а заодно вернуть франшизе былую славу. К сожалению, как и прочие устройства ввода на базе инфракрасных сенсоров, скейты не отличались точностью, так что ни о какой иммерсивности не могло быть и речи. Проект закономерно провалился: хотя Tony Hawk: Ride вышла одновременно на всех актуальных на тот момент консолях, ее стартовые продажи составили лишь немногим более 114 тысяч копий. «Инновационный геймплей и отзывчивое управление» не оценили ни геймеры, ни критики: средний балл игры на Metacritic так и не перешагнул отметку 44 пункта. Впрочем, одно неоспоримое достоинство у необычного контроллера все же было: обозреватель GameSpot отметил, что доску практически нереально сломать, и даже накинул за это проекту 1 дополнительный балл. Как журналисту удалось это выяснить, остается загадкой. Можно лишь предполагать, что, находясь под глубочайшим впечатлением от самой игры и контроллера, рецензент решил проверить скейт на прочность. Несмотря на слабые продажи, спустя всего год Activision вновь наступила на те же грабли, выпустив Tony Hawk: Shred, также заточенную под «иммерсивное» управление с помощью контроллера-доски. Сиквел побил

антирекорды оригинала: в первую неделю продажи составили около 3000 копий — старт был в 40 раз хуже по сравнению с выпуском предшественника. Wi-BOWL для серии игр Wii Sports. Контроллер Wi-BOWL для Wii Sports. Технически Wi-BOWL, созданный компанией CTA Digital для игровой консоли Nintendo Wii, являлся лишь дополнением Wii Remote: контроллер игровой консоли устанавливался внутрь корпуса, габариты которого в точности повторяли обычный шар для боулинга, если, конечно, не обращать внимания на кнопки управления. Аксессуар предназначался для поклонников серии игр Wii Sports, в рамках которых геймеры могли поиграть в виртуальный боулинг. Wi-BOWL обеспечивал максимальное погружение при игре в Wii Sports. На поверхности Wi-BOWL присутствовали три отверстия, благодаря чему вы могли взять его в руки точно так же, как и настоящий шар для боулинга, и использовать во время игры привычные движения. В свою очередь, ложе контроллера было спроектировано таким образом, чтобы в него можно было поместить Wii Remote вместе с подключенным к нему аддоном Wii MotionPlus, повышающим точность отслеживания движений. Wi-BOWL можно было держать в руках точно также, как и обычный шар для боулинга. Хотя сам по себе Wi-BOWL не привносил в геймплей ровным счетом ничего нового, по задумке своих создателей он должен был обеспечить больший уровень погружения во время игры. И они действительно добились своего: Wi-BOWL настолько хорошо имитировал реальный шар, что во время игры было немудрено разбить телевизор, запустив контроллером в экран, или же вывихнуть запястье страховочным ремешком. Маракасы для Samba de Amigo. Контроллер-маракас для Samba de Amigo. Контроллеры в виде музыкальных инструментов давно не являются чем-то из ряда вон выходящим (для той же Guitar Hero выпускалась масса разнообразных гитар и даже целая барабанная установка), однако необычные экземпляры встречались и среди них. Яркий пример — устройство ввода для Samba de Amigo, портированной на Sega Dreamcast в 2000 году. В комплекте с музыкальной ритм-игрой поставлялся уникальный контроллер в виде пары маракасов, в каждый из которых был встроен датчик движения. Геймплей сводился к тому, что игроку нужно было трясти одним или сразу двумя маракасами в нужной позиции (высокой, средней или низкой), следуя подсказкам на экране. Чем лучше вы играли в Samba de Amigo, тем больше забавных персонажей приходило потанцевать под музыку. Наученная горьким опытом с Activator, в этот раз Sega подошла к созданию контроллера со всей ответственностью — маракасы работали практически безупречно. Что же касается самой игры, то, помимо необычной системы управления, она выделялась на фоне конкурентов благодаря оригинальному способу визуализации прогресса через интерактивное окружение: чем лучше игрок справлялся с ролью музыканта Амиго, тем больше вокруг его протагониста собиралось радостно танцующих персонажей. Wii Remote и Nunchuk с надетыми на них внешними корпусами в виде маракасов. В 2008 году Samba de Amigo была переиздана и для консоли Nintendo Wii (кстати, ее портированием занималась Gearbox Interactive, известная современным геймерам по франшизе Borderlands), однако никаких дополнительных контроллеров в комплекте с игрой не поставлялось, поскольку их успешно заменяли Wii Remote и Nunchuk. Однако на релиз оперативно отреагировали сторонние

производители аксессуаров для игровых консолей, выпустив для моушн-контроллеров стилизованные пластиковые корпуса. И не прогадали: среди поклонников игры такие изделия пользовались огромной популярностью. DK Bongos для Donkey Konga Контроллер-бонго для игры Donkey Konga Еще один необычный контроллер был выпущен в 2003 году для игры Donkey Konga – музыкального спин-оффа культового платформера, повествующего о приключениях гориллы Донки. Девайс представлял собой миниатюрную копию кубинских бонго (сдвоенных барабанов) и имел встроенный микрофон. По геймплею Donkey Konga представлял собой типичную ритм-игру с той лишь разницей, что игроку нужно было не только бить в барабаны в такт музыке, но и вовремя хлопать в ладоши. Проект оказался достаточно успешным и даже получил два сиквела. Пульт управления RailDriver для Microsoft Train Simulator Контроллер RailDriver для железнодорожных симуляторов С учетом того, что Train Simulator 2022 со всеми DLC обойдется вам почти в 100 тысяч рублей, покупка специализированного контроллера за пару сотен баксов для столь премиального хобби уже не выглядит чем-то из ряда вон выходящим. И таковой действительно существует: внушительных размеров пульт RailDriver от P.I. Engineering, первая версия которого увидела свет еще в 2003 году, был разработан специально для железнодорожных симуляторов, чтобы каждый любитель поездов мог почувствовать себя машинистом, не вставая с любимого кресла. Оригинальная модель была разработана для Microsoft Train Simulator, изданной в 2001 году. Поскольку контроллер не имел нативной поддержки со стороны игры, программистам P.I. Engineering пришлось также создать собственный эмулятор для интерпретации поступающих с пульта команд. Однако RailDriver оказался настолько удачным, что вскоре разработчики трейнсимов стали самостоятельно внедрять поддержку необычного девайса в свои проекты, благодаря чему сегодня RailDriver прекрасно работает с Train Simulator, Train Sim World, Trainz RailRoad Simulator, Open Rails и многими другими играми. Модульный контроллер для Steel Battalion Пульт управления мехом для Steel Battalion Из всех меха-симуляторов самым амбициозным и по сей день остается Steel Battalion – уникальный в своем роде эксклюзив, выпущенный для игровой консоли Xbox в 2002 году японской компанией Sarcasm. Данная игра позволяла любому желающему почувствовать себя пилотом боевого шагохода, обеспечивая невероятный уровень погружения за счет уникального контроллера Mega-Jockey-9000. Внушительных размеров пульт имел 44 элемента управления, в том числе два джойстика для выбора направления движения и наводки орудий, блок из трех педалей (газ, тормоз, форсаж), переключатель скоростей и множество кнопок для активации тех или иных подсистем боевого робота. Сама игра получилась по-настоящему хардкорной. Так, например, чтобы запустить шагоход, игроку требовалось активировать в нужном порядке все необходимые подсистемы, а в дальнейшем отслеживать их состояние во время боя, считывая множество разнообразных параметров с индикаторов виртуального кокпита. Приборы кокпита позволяли отслеживать состояние меха. При неосторожном управлении боевой робот мог с легкостью перевернуться на бок и стать абсолютно беззащитным, поэтому геймеру требовалось постоянно оттачивать навыки управления. С подсистемами было тоже не все гладко: двигатель мог перегреться из-за слишком высокой нагрузки,



а бортовой компьютер – «зависнуть» в самый неподходящий момент. В Steel Battalion даже была неотключаемая permadeath: если шагohод получал критические повреждения, игроку нужно было успеть катапультироваться до того, как боевая машина взорвется, иначе виртуальный персонаж погибал, а сохранения удалялись и кампанию приходилось начинать заново, уже с новым пилотом. В общем, все как в жизни, если бы, конечно, боевые роботы реально существовали. Steel Battalion – тот редкий случай, когда экспериментальный и в высшей степени хардкорный продукт нашел свою аудиторию и принес прибыль создателям: несмотря на невероятную сложность и внушительную цену \$200 за комплект диск + контроллер, на момент релиза Steel Battalion заняла пятую строчку в списке самых продаваемых игр Японии.

\xa0Finger Dance Mat для Dance Dance RevolutionFinger Dance Mat – танцевальный коврик для пальцевДля чего люди покупают танцевальные симуляторы? Очевидно, для того чтобы оторваться на полную катушку в компании друзей или же размяться под заводной трек после многочасового сидения в офисном кресле. Однако разработчики Finger Dance Mat посчитали, что окажут геймерам неоценимую услугу, создав танцевальный контроллер для пальцев.Миниатюрный пульт позволял играть в Dance Dance Revolution, Just Dance, Dance Central и другие хиты всего двумя пальцами, а для большего погружения геймер мог превратить собственную руку в миниатюрную танцовщицу, воспользовавшись идущим в комплекте стикером. Но, вопреки ожиданиям создателей гаджета, неблагодарная публика не оценила идею, так что выпуск контроллеров пришлось свернуть.

Sega Fishing Rod для Sega Marine Fishing и других рыболовных симуляторовКонтроллер-удочка Sega Fishing RodНа волне бума популярности симуляторов рыбалки в начале нулевых Sega выпустила для Dreamcast уникальный геймпад-спиннинг и не прогадала: устройство пользовалось огромным спросом и помогло привлечь дополнительное внимание к игровой консоли со стороны вовлеченного и преданного сообщества. Sega Fishing Rod работала точно так же, как и самая обычная удочка: чтобы закинуть крючок или подсечь добычу, игроку требовалось делать характерные движения, а для того, чтобы вытащить рыбу из воды,\xa0– вращать ручку для сматывания виртуальной лески. Контроллер отличался достаточно хорошей точностью и поддерживал практически все рыболовные симуляторы, выпущенные для Dreamcast, включая Sega Marine Fishing, Bass Fishing, GetBass и многие другие игры.Геймпад-катана для Onimusha 3: Demon SiegeКатана-контроллер The Soul для PS2-версии игры Onimusha 3: Demon SiegeK выходу экшена Onimusha 3: Demon Siege Capcom подготовила для фанатов особое издание игры, в которое входил моушн-контроллер для консоли PlayStation 2, выполненный в виде катаны «The Soul». Полный комплект включал в себя сам бутафорский меч со съемным пластиковым клинком (это позволяло играть даже в тесной комнате), ножны и декоративную подставку.К сожалению, нормально пользоваться таким геймпадом было невозможно: стремясь максимально удешевить устройство, Capcom закупила для изготовления катан чуть ли не самые низкокачественные датчики движения, какие только были на рынке, так что контроллер не мог похвастаться достаточной точностью для комфортной игры в энергичный

слэшер.\ха0Геймпад-клавиатура для Phantasy Star OnlineГеймпад-клавиатура для Phantasy Star Online позволяла игрокам общаться в чате с сопартийцамиPhantasy Star Online, выпущенная в 2000 году студией Sonic Team для Sega Dreamcast, а впоследствии переизданная на Xbox, Nintendo GameCube и Windows, стала первой в мире консольной MMORPG. В процессе ее создания разработчики столкнулись со множеством вызовов, наиболее острым из которых стала проблема коммуникации между игроками. Поскольку в эпоху dial-up с поминутной тарификацией ни о какой голосовой связи не могло быть и речи, Sonic Team реализовали в игре многофункциональный текстовый чат, позволяющий геймерам общаться между собой с помощью полнотекстового ввода, иконок-эмотиконов и конструктора фраз, в котором игроки могли составлять предложения из заранее сформированного списка наиболее часто употребляемых слов.Чтобы владельцы консолей не чувствовали себя ущемленными и могли общаться наравне с ПК-игроками, специально для Phantasy Star Online был разработан геймпад со встроенной клавиатурой. Хотя из-за внушительных размеров пользоваться таким девайсом было не так удобно, как обычным контроллером, со своей основной задачей данное устройство справлялось вполне сносно: геймеры могли коммуницировать между собой без каких-либо ограничений.\ха0Бутыль с джином для Barbie: Magic Genie AdventureКомплект Barbie: Magic Genie Adventure для ПК с контроллером-бутылкойОригинальная Barbie: Magic Genie Adventure была разработана компанией Vicarious Visions и издана Mattel Interactive осенью 2000 года эксклюзивно для портативной консоли Game Boy Advance. Несмотря на название, с привычными адвенчурами игра имела мало общего, по сути представляя собой набор довольно примитивных пазлов и мини-игр. Хотя игра была прохладно встречена критиками и закономерно провалилась в продажах, спустя несколько лет Mattel подарила своему детищу вторую жизнь, подготовив своеобразный ремастер для персональных компьютеров.Суть геймплея осталась практически неизменной, однако новоявленный ПК-эксклюзив обзавелся полноценной 3D-графикой и уникальным контроллером в виде волшебной бутылки, которую игрок должен был открывать и тереть, чтобы призывать себе на помощь Барби-джинна или же совершать разнообразные магические действия (при этом «драгоценные камни» на бутылки начинали светиться).Несмотря на креативную систему управления, новая инкарнация Barbie: Magic Genie Adventure также с треском провалилась, что лишний раз доказало простую истину: даже самый экстравагантный девайс не сможет продать посредственную игру.Со всеми удобствами: полезные девайсы для геймингаВпрочем, помимо сомнительных устройств, годящихся разве что на роль музейных экспонатов, рынок игровой периферии может предложить и немало действительно полезных девайсов, способных стать отличным подспорьем для каждого любителя виртуальных развлечений.Сплит-клавиатуры и кейпадыСплит-клавиатура из двух половинИзначально сплит-клавиатуры разрабатывались в качестве меры профилактики синдрома запястного канала. Так называется заболевание, вызванное сдавливанием срединного нерва между костями запястья, поперечной связкой кисти и сухожилиями мышц руки. Все дело в том, что при использовании традиционных клавиатур наши кисти постоянно отведены кнаружи, что и

приводит к сдавлению срединного нерва, провоцируя онемения, боли и нарушение чувствительности пальцев и ладони. Неправильное положение кистей чревато развитием синдрома запястного канала. Решить данную проблему помогает разделение клавиатуры на два буквенно-цифровых блока, которые можно расположить относительно друг друга таким образом, чтобы кисть и предплечье всегда находились на одной оси. Сплит-клавиатура обеспечивает естественное положение кистей рук. Логичным развитием данной концепции стало появление кейпадов – эргономичных устройств ввода для одной руки. В самом простом случае кейпад представляет собой мини-клавиатуру с набором самых часто используемых в компьютерных играх клавиш и некоторым количеством программируемых кнопок. Наиболее продвинутые модели оснащаются регулируемым по высоте и наклону упором для ладони и дополнительными органами управления вроде мультимедийных клавиш, аналоговых стиков или трекболов. Кейпад от Razer с аналоговым стиком и колесом прокрутки. Обеспечивая естественное положение кисти во время игры, кейпады помогают разгрузить запястье и способны стать отличным подспорьем для заядлых геймеров, проводящих в виртуальных мирах по несколько часов кряду.

**Вертикальные мыши** Вертикальная мышь Logitech MX Vertical. Вертикальные мыши преследуют ту же цель, что и сплит-клавиатуры: такие устройства ввода обеспечивают наиболее физиологичное положение руки во время работы или игры, помогая минимизировать пронацию кисти относительно центральной оси предплечья и предотвратить ущемление срединного нерва. Одним из первых таких устройств стала мышь Logitech MX Vertical, напоминающая по своей форме акулий плавник с углом наклона 57 градусов. Похожий дизайн имеют и подавляющее большинство аналогов, представленных на современном рынке, однако в каталогах интернет-магазинов попадаются и более экзотические девайсы. Вертикальная мышь от 3M в форме джойстика. Хотя вертикальные мыши помогают заметно разгрузить запястье, рассматривать их в качестве полноценной замены традиционных игровых «грызунов» все же нельзя. Такой девайс отлично подойдет для пошаговой стратегии или тактики вроде X-Com, а вот для онлайн-шутеров или MOBA, в которых на первый план выходят скорость, точность и безупречный микроконтроль, подобное устройство будет малоприспособно, поскольку для перемещения курсора вам придется двигать уже не кистью, а всей рукой. Тем не менее для многочасовых партий в Civilization или Stellaris вертикальная мышь – то, что доктор прописал, как в прямом, так и в переносном смысле.

**Хабы для игровых ноутбуков** WD\_BLACK D50 – многофункциональный хаб и внешний SSD в одном корпусе. В связи со сложившейся ситуацией на рынке комплектующих, приобретение игрового ноутбука становится куда более выгодным вложением средств, нежели покупка топовой видеокарты. Но зачастую, стремясь уместить мощное железо в компактный корпус, производители ноутбуков жертвуют удобством конечного пользователя. В качестве примера здесь можно привести крайне скудные возможности апгрейда: в большинстве случаев самое большее, что вы можете сделать, – увеличить RAM и заменить (не добавить, а именно заменить, поскольку дополнительных M.2-слотов во многих моделях просто нет) штатный SSD на более емкий и быстрый. Другая проблема состоит в катастрофической

нехватке портов для подключения всей необходимой периферии. Так, лэптоп с достаточно быстрой видеоподсистемой может получить единственный выход для дополнительного монитора, хотя легко справился бы даже с двумя, а вместо четырех USB-портов 3-го поколения вы можете обнаружить лишь один, тогда как остальные будут поддерживать устаревший стандарт USB 2.0, характеризующийся гораздо меньшей пропускной способностью. К счастью, современные хабы с поддержкой интерфейса Thunderbolt 3 способны нивелировать все перечисленные недостатки. Посмотрим, на что способны такие устройства, на примере WD\_BLACK D50 Game Dock. WD\_BLACK D50 позволяет подключить всю необходимую периферию WD\_BLACK D50 – это многофункциональная игровая станция, объединившая в себе 5-портовый USB-хаб (2 × USB Type-C + 3 × USB Type-A с поддержкой стандарта USB 3.2 Gen 2), внешний NVMe SSD емкостью 1 или 2 ТБ, гигабитную сетевую карту, разъемы DisplayPort 1.4 и Thunderbolt 3 для подключения дополнительных мониторов, а также комбинированный аудиовход/выход для игровой гарнитуры. WD\_BLACK D50 предлагает набор портов на все случаи жизни. Появление такого комбайна «все в одном» стало возможным благодаря особенностям Thunderbolt третьего поколения: объединяя в себе PCI Express Gen 3 и DisplayPort 1.4, данный интерфейс обеспечивает возможность передачи любых типов данных по одному кабелю. В сочетании с рекордной пропускной способностью 5 ГБ/с это позволяет создавать по-настоящему универсальные решения наподобие WD\_BLACK D50, существенно расширяющие функционал вашего игрового ноутбука. Так, например, используя игровой док, вы сможете подключить дополнительные мониторы для мультибоксинга, если речь идет о MMO, или же для лучшего обзора и погружения в игровой процесс, если говорить о космических симуляторах, шутерах или гонках. Наличие встроенного SSD позволит вам забыть о нехватке места для новинок, благо скорости накопителя, достигающей 3 ГБ/с в операциях чтения, вполне хватит для комфортной игры даже в самые технологичные AAA-тайтлы. В свою очередь, широкий выбор портов поможет наиболее оптимально организовать подключение периферии: с WD\_BLACK D50 вам больше никогда не понадобится перетягивать десяток кабелей каждый раз, когда вы захотите отвлечься от работы и поиграть. WD\_BLACK D50 может похвастаться стильной светодиодной подсветкой. Кстати, у данного хаба есть еще одна интересная особенность, которая поможет вам дополнительно сэкономить место на рабочем столе, избавившись от лишних проводов, \xa0– функция сквозной зарядки через все тот же порт Thunderbolt 3, благодаря чему игровой док сможет заменить собой внешний блок питания. Ну а в качестве вишенки на торте стоит упомянуть полностью настраиваемую светодиодную подсветку. ']]

cluster: 3:

[[ 'Мартовский дайджест радует разнообразием курсов, вебинаров, конференций и митапов IT HR \xa0 сообщества. Пока они еще в безопасном онлайн-формате, но тяга людей к личному общению не слабеет и, скорее всего, ближе к концу месяца мы увидим больше очных мероприятий. Под спойлером я традиционно сохранила образовательные курсы и интересные события общей эйчар-тематики, которые встретились мне при подготовке

этого дайджеста. Следите за обновлениями поста и за комментариями, я монитрю новые анонсы и добавляю их по мере появления.

**Профессия IT-рекрутер (онлайн-курс)**  
Когда: начало 1 марта  
Условия участия: 93 600 рублей  
Организатор: Skillbox  
На этом курсе вас научат разбираться в технологиях и особенностях подбора в IT-сфере, правильно составлять вакансии и проводить собеседования, оценивать hard и soft skills, а также быстро находить ценных кандидатов. По окончании курса вы построите личный бренд рекрутера. Программа состоит из двух взаимосвязанных блоков: рекрутинг и технологии. Всего в ней 136 уроков, разбитых на 20 тематических модулей. Преподаватели курса: Даниил Пилипенко, директор кадрового центра SymbioWay; Дмитрий Котов, эксперт по управлению и подбору персонала в IT; Денис Александров, Leading Software Developer, Arcadia.

→ Подробности и регистрация **IT HR Online Meetup (онлайн-митап)**  
Когда: 2 марта, 17:00  
Условия участия: бесплатно  
Организатор: Spice IT Recruitment  
Новый онлайн-митап Moscow IT HR Community и Spice IT Recruitment будет о культуре в IT-командах госсектора, внутреннюю кухню и работу с ценностями. И об обучении в IT: образовательные платформы, внутренние программы обучения, инструменты повышения вовлеченности.

Доклады и спикеры: «Вы сформулировали ценности и рассказать о них команде, что дальше?» – Ксения Левыкина, Mos.ru; «Где, как и чему учить айтишников» – Мария Кулешова, Альфабанк.

→ Подробности и регистрация **Геймификация – лучший друг HR: зачем сотрудникам играть (вебинар)**  
Когда: 4 марта, 18:00  
Условия участия: бесплатно  
Организатор: Hurma System  
Геймификация в HR-сфере помогает разбавить рутинную рабочую атмосферу интерактивами, тем самым повысить вовлеченность и эффективность сотрудников.

Темы вебинара: Что такое геймификация: очередной маркетинговый прием или тренд; Основные приемы и игровые механики для обучения, развития и вовлечения сотрудников; Геймификация с психологической точки зрения: почему все любят играть; Примеры крутых HR-кейсов по геймификации.

Ведущая – Ксения Шаповалова, креативный директор employer-брендингового агентства makeLove с пятилетним опытом работы, преподает на курсах «Диджитализация HR» в школах MACS и «makeLove school».

→ Подробности и регистрация **Выгорание в IT сфере (онлайн-митап)**  
Когда: 12 марта, 17:00  
Условия участия: бесплатно  
Организатор: Lineate  
Выгореть можно в любой сфере и на любой должности, и IT здесь – не исключение. На митапе речь пойдет о том, как распознать синдром выгорания.

Доклады и спикеры: «Гори, но не сгорай: синдром эмоционального выгорания у эйчаров в IT» – Алёна Петрова, HRD Lineate; «Выгорание: от бесплатного печенья до депрессии» – Константин Волков, Dev Unit Lead Skyeng; «Выгорание глазами инженера: “кто обидел, кому отомстить”» – Тимур Хайруллин, генеральный директор ETHR.

→ Подробности и регистрация **IT-Recruiter. Интенсив (онлайн-курс)**  
Когда: начало 13 марта  
Условия участия: 35 990 рублей  
Организатор: Level UP  
Курс разработан для тех, кто хочет начать заниматься подбором IT-специалистов. Программа состоит из технического и рекрутингового модулей.

В программе: Что и как устроено в IT-сфере в целом. Где и для чего используются те или иные языки программирования, технологии, фреймворки. Чем занимаются те, кого вам предстоит искать. Как правильно

снимать заявку и работать с нанимающим менеджером. Как самостоятельно составлять вакансию со сложными техническими терминами. Тонкости и лайфхаки поиска IT-специалистов разного уровня. Приёмы поиска, о которых еще не написали в книгах. Как правильно позиционировать компанию на высококонкурентном рынке IT. Основы HR брендинга и HR аналитики. → Подробности и регистрация Современное решение в помощь HR (онлайн-конференция) Когда: 17 марта, 11:00 Условия участия: бесплатно Организатор: MOZONлайн-конференция и выставка digital-решений для эйчаров. Обзор трендов упаковки эйчар-контента в диджитал: инструменты, лайфхаки, распространённые ошибки. \xa0Программа конференции: «HR в эпоху цифровых технологий. Тренды современных IT-решений для HR сегодня и завтра» – Наталия Даниленко, независимый эксперт; Обсуждение «Готовое решение от провайдера или внутренняя разработка. Какие есть плюсы и минусы?» – Инна Карелина (НСПК) и Татьяна Конева (HOFF); «Топовые HR-решения для внедрения в компании: разбор полётов» – Оксана Погодаева, HR&ED-tech; Выставка решений и продуктов для эйчаров Параллельные треки: «Инновационные решения для обучения и развития», «Тонкости упаковки контента в интерактивные продукты», «Assessment 2021». → Подробности и регистрация АйтиРекрут (онлайн-курс) Когда: начало 19 марта Условия участия: 35 000 рублей Организатор: Черногорцева Ирина Это курс на понимание IT-сферы, он охватывает самые востребованные профессии в IT: разработку, аналитику, тестирование, дизайн, управление проектами. Программа состоит из шести видеоуроков, после каждого предусмотрена практика из 20 заданий на реальных резюме и вакансиях. Автор и преподаватель курса – ведущий IT-рекрутер с десятилетним стажем Черногорцева Ирина. → Подробности и регистрация IT HR Online Meetup (онлайн-митап) Когда: 30 марта, 17:00 Условия участия: бесплатно Организатор: Spice IT Recruitment Новая встреча Moscow IT HR Community будет посвящена эйчар-трансформации компании во время бурного роста и работе с возражениями кандидатов-айтишников. Доклады и спикеры: «Работа с возражениями кандидатов в IT» – Элина Хамитова, OZON. «Трансформация компании в период бурного роста количества сотрудников» – Виталий Калинин, Dodo Brands. → Подробности и регистрация Рекрутинг в IT (видеокурс) Когда: начало в любое время Условия участия: бесплатно Организатор: GeekBrains Бесплатный видеокурс Гикбрейнс, который поможет вам укрепить знания в IT-рекрутинге. Все видеоуроки доступны после регистрации. Программа курса: Урок 1. Вводная часть. Кто такой IT-рекрутер. Основные обязанности и компетенции. Урок 2. Копирайтинг в IT-рекрутменте. Как писать, чтобы нанимать лучших на рынке. Урок 3. Каналы поиска ит-специалистов. Программы автоматизации для hr/рекрутеров. Урок 4. Boolean-запросы. Как упростить поиск: лайфхаки и расширения. Урок 5. Коммуникации с кандидатами / заказчиками. Как проверить технические скиллы, если вы не технический рекрутер. → Подробности и регистрация Введение в IT для HR (онлайн-курс) Когда: начало в любое время Условия участия: 10 990 рублей Организатор: Udemu Курс об особенностях сферы разработки программного обеспечения для специалистов по подбору персонала в IT. Речь в нем пойдет о сфере IT в целом и о профилях IT-специалистов. В программе курса: Классификация IT-

проектов и процессов их разработки; Описание ИТ-профессий и навыков; Практическое занятие по анализу резюме ИТ-специалистов на достоверность и адекватность опыта работы. → Подробности и регистрация Секреты ИТ-подбора (онлайн-курс) Когда: начало в любое время Условия участия: 11 000 рублей за все модули Организатор: GetIT Видеокурс, который поможет вам научиться разбираться в ИТ-позициях, находить редких кандидатов и мотивировать айтишников работать в вашей компании. Программа курса состоит из пяти модулей, которые можно купить сразу и пройти полное обучение или оплачивать и проходить по отдельности. Модуль 1. Погружение в ИТ-рынок. Модуль 2. Поиск ИТ-специалистов. Модуль 3. Сорсинг ИТ-специалистов. Модуль 4. Работа с кандидатом. Модуль 5. Работа с заказчиком. Автор и ведущая курса – Светлана Петровичева, основательница IT Recruiter School и GetIT. → Подробности и регистрация Полезные эйчар-мероприятия Любопытный курс о психотерапии для эйчаров стартует 1 марта. Компания «Люди и коммуникации» 3-4 марта в Москве организует форум HR-контент 2021. SPALAN проведет мастер-класс об эйчар-маркетинге и бренде работодателя 10 марта. QUORUM запустит онлайн-марафон по эйчар-аналитике и метрикам 10 марта. У Hurma System запланированы вебинары о системе слепого найма (11 марта) и о международном рекрутинге (18 марта). В самом разгаре онлайн-марафон HR-трансформация, в рамках которого 10-12 и 25-26 марта пройдут однодневные тематические мини-конференции. TalentTech предлагает видеокурс о цифровой трансформации эйчар-процессов с доступом в любое время. Об организаторах мероприятий: Черногорцева Ирина – ИТ-рекрутер, обучающий профессии с нуля. GeekBrains – образовательный портал, который помогает начать карьеру в digital с нуля. GetIT – ИТ рекрутинговое агентство. Hurma System – HRIS/HRM-система. Level UP – ИТ-образование в Санкт-Петербурге и удалённо. Lineate – российско-американская компания в сфере разработки программного обеспечения. MOZ – офлайн- и онлайн-решения, развивающие сотрудников. Skillbox – онлайн-университет современных digital-профессий. Spice IT Recruitment – ИТ-специализированное кадровое агентство. Udemu – международная платформа для онлайн-обучения. Если вы организатор мероприятий для эйчаров и рекрутеров в ИТ и хотите, чтобы они появлялись в наших дайджестах, пишите нам на ящик [entries@habr.team](mailto:entries@habr.team).']]

cluster: 4:

[[ 'Калифорнийские регуляторы предъявили компании Raiser-CA, дочерней организации Uber, штраф в 7,3 миллиона долларов за непредоставление отчётности о доступности транспортных средств для инвалидов, безопасности и служебной информации в 2014 году. У компании есть возможность в течение тридцати дней подать апелляцию или оплатить штраф. Uber обжалует решение. \n\n\n\r\nRaiser-CA предоставляет услуги UberX в Калифорнии. Комиссия по регулированию общественных услуг (California Public Utilities Commission) не получила от компании отчёты о количестве заказов такси, сделанных людьми с ограниченными возможностями, и о наличии автомобилей, способных выполнить такие заказы. Вторым нарушением компании стало непредоставление информации о



количестве заказанных поездок, количестве отказов и сумме денег, выплаченных клиентами. Третье нарушение – отсутствие информации о дорожно-транспортных происшествиях с участием водителей UberX. \n\r\nКомпании дали тридцать дней на оплату или обжалование штрафа в 7,3 миллиона долларов. В ином случае деятельность компании на территории Калифорнии будет приостановлена. \n\r\nПредставитель Uber заявил, что штраф «глубоко разочаровывает». Компания обжалует решение, потому что Uber уже предоставил значительные объемы данных в Комиссию по регулированию общественных услуг – это информация, которую компания предоставляет в других штатах. По мнению Uber, запросы «выходят за рамки Комиссии и не улучшат общественную безопасность».\n\r\nВ последние месяцы у Uber были проблемы с обслуживанием некоторых клиентов: на компанию в сентябре 2014 года подали в суд Сан-Франциско, утверждая, что водители UberX нарушают права инвалидов, отказываясь возить слепых людей с животными. \n\r\n\r\nUber по разным причинам преследуют по всему миру, часто из-за вмешательства инициативных сотрудников официальных такси. В мае UberX запретили в Южной Корее: новый закон запрещает оказывать транспортные услуги без лицензии. В том же месяце миланский суд вынес решение о запрете деятельности UberX на территории Италии. Услуги UberPOP во Франции были прекращены после того, как водители официальных такси начали переворачивать и сжигать машины Uber. \n\r\n\r\nКомиссия по защите конкуренции Болгарии в июле 2015 года оштрафовала две партнерских компании Uber на 100 тысяч евро: услуги Uber аналогичны услугам официальных такси, но компании не платят налогов, не получают лицензии и не отчитываются регуляторам о доходах, что даёт им конкурентные преимущества. \n\r\n\r\nФедеральная антимонопольная служба России не считает, что сервисы Uber, Gett и «Яндекс.Такси» нарушают законодательство. Замглавы ведомства Андрея Цариковского заявил что сочувствует водителям такси, но интернет-сервисы, подобные Uber, очень удобные.']]

cluster: 5:

[[ 'Spot танцуют, стойко сносят пинки, буксируют грузовики и катают Адама Сэвиджа на рикше. У этих роботов отличный пиар, но что вы на самом деле знаете об их возможностях? Давайте разберемся, как устроены Spot, как ими управлять, как под них разрабатывать. И главное, зачем серьезным компаниям покупать четвероногих роботов по 75 000 долларов за штуку. Случайный пиар Никто в Boston Dynamics не ожидал, что их роботы станут частью массовой культуры. По крайней мере, так говорит Марк Райберт (Marc Raibert) – основатель, председатель совета директоров и бывший генеральный директор компании. Однажды некто скачал несколько ранних роликов с сайта Boston Dynamics и залил на YouTube. Они привлекли много внимания, и закрутилось. Мы получили положительную реакцию от наших клиентов и спонсоров. После мы решили: «Хорошо, нужно завести свой канал на YouTube». Сейчас на Boston Dynamics подписано больше полутора миллионов человек. Правда, известность сыграла с компанией злую шутку. Из-за всех этих роликов роботов часто переоценивают. Разработка Spot Дизайн добавляет путаницы. Spot напоминает собаку, поэтому люди думают, что он будет вести себя

так же. Например, будет поворачиваться, чтобы что-то рассмотреть. Да, в конструкции Spot есть элементы бионического дизайна, но перед инженерами не было задачи сделать робота похожего на собаку и ведущего себя, как животное. Можно сказать, что его внешность – результат сходящейся эволюции. Для многих робот – просто компьютер, который раздает команды механизмам и получает обратную связь при помощи сенсоров. В реальности, с одной стороны механизмами управляет компьютер, а с другой – действуют гравитация, трение, закон сохранения импульса и вся остальная бессердечная физика. Чтобы робот мог нормально работать за пределами лаборатории, алгоритмы, механика и окружающая среда должны гармонично взаимодействовать. В Boston Dynamics добиваются этого, постепенно совершенствуя роботов. Инженеры компании собирают прототип, прогоняют его через серии практических тестов, и на основе полученных данных проектируют новую версию. Они начинают с конструкций на базе заводских модулей, и поколение за поколением продвигаются ко все более сложному кастомному дизайну, рассчитанному на 3D-печать. Слева – классическая конструкция: винты, кабели и разъемы. Справа – усовершенствованная. Она больше похожа на анатомический рисунок. Часть бедра с гидравликой, каркас, приводы и вся остальная машинерия будто образуют единое целое. Так, за 15 лет brutальный 110 кг BigDog, превратился в Spot – робота, которого многие хотели бы увидеть в будке около своего дома.

**Анатомия Spot**  
**Механика**  
Длина Spot – 110 см, а в холке робот достигает 84 см. Формально у Spot есть перед – там расположены светодиодные индикаторы состояния. Сзади находится набор разъемов. Однако, робот одинаково быстро движется и вперед, и назад. Максимальная скорость достигает 5,76 км/ч – его можно обогнать трусцой. В лапы встроено 12 электрических приводов – по два в бедренных суставах, и по одному в каждом колене. На конце каждой конечности расположены протекторы, отдаленно напоминающие автомобильные покрышки. Такая конструкция позволяет Spot двигаться приставным шагом, разворачиваться на месте, приседать, ложиться на землю, уверенно взбираться по склонам с уклоном до 30 градусов, подниматься по лестницам и вставать после падений. Spot весит 32,5 кг и снабжен ручками для переноски. Спецы из iFixit еще не дотянулись до Spot, а Boston Dynamics не спешит делиться коммерческими секретами. Поэтому мы мало знаем о том, как именно этот робот поддерживает равновесие. В своих ранних машинах Boston Dynamics использовала акселерометры, твердотельные гироскопы и гиростабилизаторы. Распределение веса контролировалось при помощи тензодатчиков и датчиков крутящего момента в суставах. Все вычисления, необходимые для работы Spot, выполняются «на борту», но, опять же, неясно, какая электроника за них отвечает. Мне попадались слухи о двух процессорах Intel Core I7, но подтвердить их не удалось. Корпус Spot имеет рейтинг IP54. Защита от воды хуже, чем у iPhone, но зато не разобьется при падении. Рассмотреть начинку робота мешает двухслойный защитный корпус. Снаружи – покрытие из поликарбоната, внутренний слой сделан из пенопласта. Такая конструкция смягчает удары о препятствия. Spot не боится дождя, строительной пыли и может работать при температуре от -20°C до 45°C. Батарея. К каждому Spot прилагается пара сменных литий-

ионных аккумуляторов емкостью 600 Вт-ч каждый и 400 Вт зарядное устройство. На зарядку аккумулятора требуется 2 часа. Батареи вставляются в брюхо робота и обеспечивает 90 минут активной работы или около 4 часов сидя в режиме ожидания. Пара батарей и зарядное устройство (выход: 35-58,6 В постоянного тока, 7,2 А Макс.) В документации к Spot упоминается, что при некоторых движениях приводы не тратят, а генерируют ток. В результате, напряжение на шине робота может превышать напряжение батареи, но вряд ли это заметно продлевает время работы. Сенсоры и восприятие пространства Spot видит все вокруг, но недалеко. Инженеры снабдили его пятью стереокамерами, обеспечивающими 360-градусный обзор. Сенсоры выдают черно-белую картинку и позволяют построить карту глубины, но дальность обзора составляет всего 4 метра. Чтобы ориентироваться на местности и избегать препятствий, Spot использует не только камеры, но и одометрию – данные о движении приводов. Они формируют машинный аналог человеческой проприоцепции – ощущения положения тела в пространстве. При помощи камер и одометрии Spot поддерживает и обновляет модель, которая описывает позу робота и положение относительно других объектов. Связь Все приложения связываются с компьютером Spot через сетевое соединение по протоколу TLS 1.2 или 1.3. Хотя у робота есть разъем RJ-45, для управления обычно используют встроенный модуль Wi-Fi 802.11. Spot может подключаться к существующим Wi-Fi сетям или сам выступать в качестве точки доступа. При прямом подключении к роботу можно рассчитывать на максимальную дальность управления в 50 метров. Чтобы увеличить это расстояние, Boston Dynamics рекомендует подключать к Spot LTE-модем. Управление Spot В YouTube-роликах Spot кажутся довольно самостоятельными, но еще в прошлом году это было не так. Когда их начали рассылать заказчикам, собаки Boston Dynamics не могли ориентироваться в пространстве без помощи оператора. Контроллер Обычно для управления Spot используют контроллер с незамысловатым названием Spot Explorer. Это 7-дюймовый планшет, который выглядит как помесь Nintendo Switch с геймпадом Xbox. Boston Dynamics в этом никогда не признается, но его можно купить на Aliexpress под названием JXD S192K: разрешение дисплея – 1920x1200, процессор Rockchip 1,8 ГГц, 4 ГБ оперативной, 64 ГБ встроенной памяти и аккумулятор на 10 000 мАч. В магазинах этот планшет поступает с Android 5.1 на борту, но версия Boston Dynamics обновлена до Android 8.1 и заполнена кастомным софтом для управления роботом. Чтобы заставить Spot двигаться, достаточно пары стиков по бокам от экрана планшета. Левый отвечает за движение вперед-назад и вправо-влево, а правый управляет вращением робота вокруг своей оси. Если прикоснуться к экрану, робот переместится к выбранной точке, совсем как в Google Street View. Врожденный интеллект Чтобы освоиться с управлением, нужны считанные минуты. Дело в том, что большую часть работы Spot берет на себя. Робот самостоятельно избегает столкновений, переступает через препятствия и адаптируется к неровностям местности. Это работает удивительно хорошо, но не идеально. Spot все еще теряет равновесие на скользких поверхностях и спотыкается о кабели и веревки. Он может наступать на банки с краской, ведра и другие предметы ниже 30 см, плохо распознает зеркальные и очень яркие

препятствия. В инструкции отдельно подчеркивается, что Spot может врезаться в стеклянные двери. А еще, если придет робоапокалипсис, вы можете попробовать спастись в высокой траве. AutowalkSpot может запомнить дорогу и самостоятельно пройти по сохраненному маршруту. Эта функция делает из игрушки с радиоуправлением инструмент для автоматизации повторяющихся задач. Сначала Spot нужно вывести на прогулку — включить запись и провести по маршруту вручную. В режиме Autowalk робот пройдет тот же путь самостоятельно, обходя появившиеся препятствия и повторяя все действия, которые совершал раньше. Например, он будет фотографировать в одних и тех же точках маршрута. Чуть позже вы поймете, зачем это нужно. Чтобы Spot понимал где находится и куда движется, ему нужны визуальные ориентиры — путевые точки. Обычно робот отмечает их с интервалом в 2 метра. Это места с углами, мебелью, оборудованием и другими подобными особенностями. Поэтому Spot хорошо ориентируется в насыщенной среде, но может потеряться в трех соснах в длинном белом коридоре. В таких местах Boston Dynamics рекомендует развешивать фидуциары — картинки, похожие на QR-коды. Они служат искусственными ориентирами. Как минимум один такой тег обязательно нужен, чтобы отметить начало пути. Записи маршрутов хранятся в Spot Explorer. Их может повторить любой робот, к которому подключен контроллер. Продвинутая навигация Весной 2020 года Boston Dynamics выпустила обновление Spot 2.0. Теперь под руководством оператора Spot может построить подробную карту местности. Затем нужно выбрать на карте те места, которые должен посетить робот, и действия, которые он должен совершить. Spot самостоятельно построит кратчайший маршрут, пройдет по нему и выполнит свою задачу. Полезная нагрузка Чтобы заставить Spot сделать что-то полезное, его нужно хорошенько навьючить. На спине робота расположена пара алюминиевых рельсов, отдаленно напоминающих автомобильный багажник. К ним можно прикрепить до 14 кг. Это не мертвый груз, а функциональные модули, которые подключаются к роботу через пару портов DB25. Разъемы обеспечивают питание (до 150 Вт на порт) и передачу информации. Spot понимает, какой модуль к нему подключен, учитывает вес и габариты груза при ходьбе. Цель Boston Dynamics — превратить Spot в «мобильную платформу», так что компания уже разработала набор аксессуаров. Они продаются по негуманным ценам на официальном сайте компании. Spot CAM и CAM+ Основной недостаток спот — близорукость, и Boston Dynamics предлагает несколько способов его исправить. Прежде всего, это Spot Cam — панорамная камера из пяти сенсоров Sony IMX290. Их дополняет кольцо из ярких светодиодов, пара динамиков и микрофоны для двусторонней связи. Spot CAM+ — то же самое, плюс управляемая pan-tilt-zoom-камера со скромным разрешением 2 МП и нескромным 30-кратным оптическим зумом. LIDAR и Faro Trek LIDAR позволяет Spot уверенно ориентироваться на больших открытых пространствах и картировать местность. Фактически Boston Dynamics перепродает лидар Velodyne VLP-16, который сканирует объекты в радиусе 100 метров с точностью до 3 см. Кроме того, компания FARO уже выпустила для Spot собственное решение для трехмерного сканирования на базе лазерного сканера FARO Focus S. Spot CORE и Core AI Чтобы обрабатывать данные с лидара в реальном времени, Spot нужна дополнительная вычислительная мощность. CORE — обычный компьютер под

управлением Ubuntu Desktop 18.04 LTS в компактном исполнении. Внутри находится процессор Intel Core i5 восьмого поколения, 16 ГБ оперативной памяти DDR4 и SSD диск на 512 ГБ. В отличие от компьютеров, управляющих Spot, эта система открыта для разработчиков. Для машинного обучения предусмотрен отдельный модуль – Core AI. Это CPU Intel Xeon E3-1515M V5, 32 ГБ оперативной памяти, 480 ГБ SSD и видеокарта Nvidia Quadro P5000 в корпусе с защитой IP67 и возможностью работы от -20 до 45°C. Spot ARM Самый интересный официальный аксессуар для Spot – роборука. Spot Arm весит 8 кг и может поднять до 4 кг. Впервые ее показали еще несколько лет назад, но обещают выпустить в продажу лишь к началу 2021 года. Ведь мало сконструировать манипулятор, нужно научить Spot им пользоваться. У Spot Arm, как и у всего робота, должен быть «врожденный интеллект». Манипулятор должен захватывать предметы автоматически, повинаясь простым командам. Кроме того, Spot Arm должен слаженно работать с остальным телом Spot. Это хорошо видно, когда Spot открывает дверь. Робот определяет местоположение ручки, выбирает угол захвата, позиционирует манипулятор, решает, толкать или тянуть, а затем тянет за ручку и маневрирует всем телом. Научить робота действовать также в различных ситуациях – непростая задача. Сейчас инженеры по всей Америке заставляют Spot собирать разные предметы с пола своих квартир. Boston Dynamics отправила сотрудников на удаленку вместе с 71 робособакой. Но, несмотря на это, в Boston Dynamics подчеркивают: Spot не подходит, чтобы убираться на кухне или бегать во дворе дома. Не для дома Spot – промышленный робот, а не домашний любимец. Да, он тише, чем прародитель – дизельный BigDog, но все еще здорово шумит, особенно в маленькой квартире. «Spot громкий», – рассказал в разговоре с The Verge Сэм Сейферт (Sam Seifert), один из инженеров Boston Dynamics, – «Это не так заметно в промышленных условиях или больших помещениях, но в маленьких замкнутых пространствах, топот Spot резонирует». На спине этого песика закреплен лазерный сканер FARO Главная причина держаться от Spot на расстоянии – он не соответствует трем законам робототехники, ни буквально ни метафорически. Робот не может предсказывать траектории движущихся объектов, и у него нет специальных навыков для контактов с людьми. Кроме того, его конструкция довольно опасна. Шарниры могут с легкостью прищемить пальцы. Инструкция не скупится на кровавые подробности. На такой случай предусмотрен аварийный выключатель в приложении и на шасси робота, и все же, Spot не стоит использовать в тесном контакте с людьми, тем более дома. Применение Spot Как и большинство современных роботов, Spot лучше всего подходит для Three Ds: dirty, dull, and dangerous – грязной, скучной и опасной работы. Но даже инженеры Boston Dynamics сначала не знали, где именно робот окажется наиболее полезен. Осенью 2019 года Boston Dynamics отправила 150 Spot различным компаниям и исследовательским центрам для тестирования. Роботов буквально отдавали со словами: «придумайте им наилучшее применение». И применения нашлись. Строительство и инженерия Spot оказались полезны на стройплощадках. Там одновременно работает много людей разных профессий, и их усилия сложно координировать. Каким бы придирчивым ни

было руководство, рано или поздно кто-то отклоняется от чертежей. Тогда приходится сносить уже готовую стену, а, в худшем случае, реконструировать целое здание. В 21 веке вместо бумажных планов в строительстве используют BIM (Building Information Modeling). Это сложная штука, но сейчас достаточно сказать, что BIM выглядит, как подробная трехмерная модель здания со всеми коммуникациями. Во время строительства специально обученный сотрудник обходит строительную площадку с фотокамерой в руках, делает сотни снимков в заранее выбранных точках, сгружает их на компьютер и сравнивает с ранними снимками и трехмерным планом. Через неделю – повторяет все заново. Ну вы поняли... К тому же, сделанные человеком фотографии отличаются по ракурсу, и это затрудняет сверку. Spot справляется с этой задачей лучше людей. Ему не надоедает ходить по одному и тому же маршруту 24 на 7, и он может делать больше снимков с большей точностью. Так, чтобы отслеживать ход строительства на площадке Pomerleau, пара роботов Spot еженедельно делала почти 5 000 фотографий, обходя территорию более 45 000 квадратных метров. Но дело не только в фотографиях. С лидаром или лазерным сканером на спине Spot может собирать данные для создания новых и обновления старых информационных моделей. Этим собаки Boston Dynamics занимаются в Мичигане. Там расположены заводы Ford, которые модернизировали и перестраивали на протяжении десятков лет. Теперь они даже приблизительно не похожи на оригинальные инженерные планы. Чтобы полностью переоборудовать производство, Ford нужно заново перенести его «на бумагу». По словам Марка Годериса (Mark Goderis), менеджера Ford, люди сканировали бы завод до двух недель, а их работа стоила бы почти 300 000 долларов. Spot позволил справиться в два раза быстрее. Энергетический сектор Другая область, в которой роботы показали себя очень хорошо – мониторинг оборудования в опасных условиях, например, в море. Так, AkerBP отвезла Spot на судно FPSO в Норвегию. Там робособака будет снимать показания с оборудования и искать утечки нефти. Меню управления Autowalk Еще один пример – морские ветряные электростанции. Ветряки генерируют переменный ток. Чтобы снизить потери при передаче, его преобразуют в постоянный на подстанции, расположенной прямо в море, в сотнях километрах от берега. И все было бы отлично, но подстанции требуют регулярных осмотров и обслуживания. Суровые условия и сложное оборудование делают такую работу очень опасной. Смертность среди морских рабочих в США в семь раз выше, чем на среднем рабочем месте. К тому же, во время осмотров станции приходится отключать. Это означает колоссальные потери – примерно 20 мегаватт-часов электроэнергии на каждую турбину за день простоя. В этом контексте 75 000 долларов за работа – не такая уж высокая цена. Тем более, Spot может приближаться к работающему оборудованию. С установленным на спину тепловизором он может обнаруживать проблемы, которые пропустят техники, например, перегрев трансформаторов. Развлечения Четвероногих роботов, как и квадрокоптеры, можно использовать в различных шоу. Как вы можете убедиться, Spot неплохо танцует. Осталось только пошить красочный костюм – и в цирк. Cirque du Soleil всерьез рассматривал Spot как артиста, но после отмены представлений у них вряд ли остались деньги на

роботов. Медицина Зато с началом пандемии роботами Spot стали интересоваться медики. Их сразу же стали применять для дистанционных консультаций, но в Массачусетский технологический институт (MIT) и Гарвардская медицинская школа пошли дальше и решили приспособить Spot и для удаленной диагностики. Исследователи установили на Spot четыре новых камеры – инфракрасную и три монохромных, которые фильтруют различные длины световых волн: 670, 810 и 880 нанометров. Инфракрасная камера позволяет измерять температуру и частоту дыхания пациента, а монохромные фиксируют изменения цвета кожи и, таким образом, определяют насыщенность крови кислородом. Близкое знакомство Теперь, когда вы представляете, зачем нужен Spot, самое время познакомиться с роботом поближе. Для этого даже необязательно выкладывать 75 000 долларов. Spot на дистанционном управлении Formant, стартап из Сан-Франциско, создает системы дистанционного управления роботами и предлагает поуправлять Spot дистанционно из любой точки мира. Для того чтобы прогуляться по парку Золотые Ворота или по местному пляжу даже не нужен контроллер, только компьютер, надежное подключение к интернету и немного везения. Компания столкнулась с наплывом желающих, но все еще принимает заявки на тестирование. Разработка и открытый SDK Boston Dynamics выпустила для Spot комплект разработчика. Он доступен на Github и снабжен подробным руководством. SDK позволяет создавать приложения для управления роботом, получать информацию с датчиков для анализа и подключать к Spot новые модули полезной нагрузки. Чтобы сделать разработку проще и доступнее, в качестве основного языка для API Boston Dynamics выбрала Python. Компания предлагает разработчикам арендовать роботов, но существуют и симуляции. Виртуальные модели Spot доступны в Webots и, с недавних пор, в Gazebo. Они не идеальны, зато бесплатны и позволяют вволю поэкспериментировать. Spot в симуляторе Webots Крупные компании уже запустили разработку приложений для Spot. Американская HoloBuilder выпустила приложение для навигации по строительным площадкам, а Rocos представила облачную платформу для управления целым парком роботов. Что дальше Сейчас в мире работает больше двух сотен Spot. За 2020 год Boston Dynamics рассчитывала собрать еще тысячу роботов. Пандемия нарушила график, но производство продолжается. Так, на четырех механических лапах в нашу реальность крадется настоящее будущее. Вполне возможно, что навыки разработки под Spot вскоре окажутся чертовски востребованы на практике. P.S. Если вам понравилась эта статья и хочется еще, ищите в моем блоге по адресу [santry.ru](http://santry.ru) и в телеграм-канале [GeeksNote](https://t.me/geeksnote). P.P.S. @AiratGl написал хорошее дополнение к этой статье: Как работает Spot от Boston Dynamics? Возможный ответ. Там много подробностей об алгоритмах поддержания равновесия на примере роботов Robotic systems lab.']]

cluster: 6:

[[ 'Привет, хабр!\n\n0 опубликованы результаты четвертого ежегодного рейтинга экспертной группы Тэглайн – «Топ-100 ведущих веб-студий Рунета 2009». Формальным показателем построения рейтинга, как и в прошлые годы, послужил оборот компаний на рынке разработки и поддержки



Интернет-проектов.\n\r\nПервая десятка рейтинга выглядит следующим образом:\n\r\n1. Студия Артемия Лебедева\r\n2. Actis Wunderman\r\n3. ADV/web-engineering\r\n4. Defa Interaktiv\r\n5. Deluxe Interactive Moscow\r\n6. Креативное агентство Red Keds\r\n7. Promo Interactive\r\n8. РБК-Софт (Армада)\r\n9. Бюро Пирогова\r\n10. Qsoft\n\nМетодология (довольно сильно поменялась с прошлого года) учитывала голосование представителей студий и внешних экспертов, а также поправочные коэффициенты по данным, поданным самими студиями. Поиме прочего, опубликована оценка объема рынка за 2008 год.\n\r\nБалл рассчитывается по хитрой формуле: сначала считаются голоса, набранные всеми студиями (голосуют сами студии и внешние эксперты). Потом берутся первые 250 компаний (проголосовавших), и их голоса (голоса “выборщиков”) учитываются с модификатором 1, а голоса остальных – с модификатором 1/10. Помимо этого, для компенсации накруток, взаимные голосования (друг за друга) учитываются с модификатором 1/2. Полученный балл умножается на 7 разных коэффициентов (каждый в диапазоне от 1 до 1,1), данные для которых подали сами студии – и так получается финальный рейтинговый балл.\n\r\nПомимо прочего, организаторы рейтинга открыли информацию, кто и как голосовал за студии первой двадцатки, а также опубликовали значения общего поправочного коэффициента. существенно увеличив прозрачность проекта.\n\r\nЕще из интересного, опубликована статистика использования студиями коробочных CMS:\n\r\n1. 1С-Битрикс – 65,3% (+15,3%)\r\n2. UMI.CMS – 22,5% (+16,5%)\r\n3. NetCat – 18,1% (+2,1%)\r\n4. Drupal – 9,3% (+7,3%)\r\n5. Joomla! – 5,4% (+1,4%)\r\n6. MODx – 5,4% (+3,2%)\r\n7. S.Builder – 4,3% (+2,3%)\r\n8. AB0.CMS – 3,8% (-1,2%)\r\n9. WordPress – 3,2% (new)\n\nРаспределение лидеров рынка хостинга по популярности среди веб-студий:\n\r\n1. .masterhost.ru – 34,1%\r\n2. Хостинг-Центр РБК – 11,2%\r\n3. RU-Center – 10,2%\r\n4. SpaceWeb – 7,1%\r\n5. AGAVA – 6,8%\r\n6. Infobox – 5,4%\r\n7. 1Gb.ru – 3,5%\r\n8. Majordomo – 3,5% \r\n9. TimeWeb – 3,2%\r\n10. ISPserver.com – 3,2%\r\n11. Зенон Н.С.П. – 2,4%\r\n12. mtw.ru – 2,4%\r\n13. McHost.ru – 2,1%\r\n14. «Караван» – 2,1%']]

## Тематическое моделирование

Импортируем библиотеки для тематического моделирования

```
import gensim
import pyLDAvis.gensim_models as gensimvis
import pyLDAvis
```

Для подбора оптимального количества тем построим график зависимости когерентности от количества тем

```
%time
df["tokenized_preprocessed_text"] =
df["preprocessed_text"].apply(lambda x: tokenize(x))
```

```

gensim_dictionary =
gensim.corpora.Dictionary(df["tokenized_preprocessed_text"])
gensim_dictionary.filter_extremes(no_above=0.1, no_below=5)
gensim_dictionary.compactify()

corpus = [gensim_dictionary.doc2bow(text) for text in
df["tokenized_preprocessed_text"]]
topics_range = range(3, 25, 1)
coherence_values = []

for num_topics in topics_range:
    lda_model = gensim.models.LdaMulticore(
        corpus=corpus,
        num_topics=num_topics,
        id2word=gensim_dictionary,
        passes=10,
        random_state=42)

    coherence_model = gensim.models.CoherenceModel(
        model=lda_model,
        texts=df["tokenized_preprocessed_text"],
        dictionary=gensim_dictionary,
        coherence="c_v")

    coherence_values.append(coherence_model.get_coherence())

CPU times: total: 8min 55s
Wall time: 19min 35s

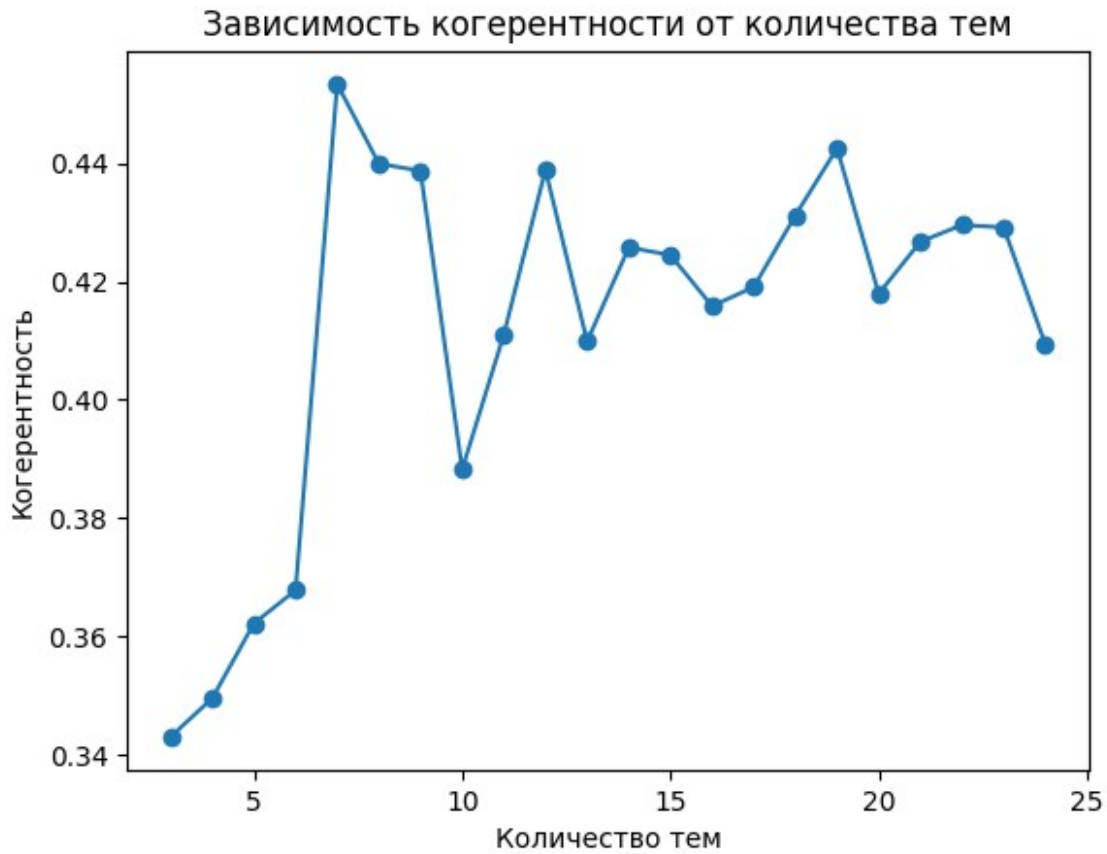
```

Выведем график на основе подсчитанных значений когерентности

```

plt.plot(topics_range, coherence_values, marker="o")
plt.xlabel("Количество тем")
plt.ylabel("Когерентность")
plt.title("Зависимость когерентности от количества тем")
plt.show()

```



На данном графике видны пики, однако при большом количестве темы получаются слишком "узкими"

Воспользуемся NMF (Non-negative Matrix Factorization) для определения скрытых тем

```
for i, topic in enumerate(nmf_model.components_):  
    print(f"Topic {i}: {",  
".join([tfidf_vectorizer.get_feature_names_out()[i] for i in  
topic.argsort()[::-1][:10]])}")
```

Topic 0: это, всё, который, человек, мочь, очень, свой, делать, время, весь

Topic 1: приложение, код, это, который, файл, использовать, функция, разработчик, всё, данные

Topic 2: курс, вебинар, разработчик, обучение, программирование, программист, язык, онлайн, студент, веб

Topic 3: такси, сервис, яндекс, компания, водитель, агрегатор, заказ, год, услуга, поездка

Topic 4: сайт, дизайн, студия, лебедев, дизайнер, артемий, веб, шрифт, страница, логотип

Topic 5: доклад, команда, компания, конференция, сотрудник, работа, проект, продукт, спикер, рассказать

Топик 6: робот, промобот, человек, робототехника, компания, технология, ия, искусственный, система, интеллект

Визуализируем данную информацию

```
lda = gensim.models.LdaMulticore(corpus, num_topics=5,  
id2word=gensim_dictionary, passes=10, random_state=42)  
vis_data = gensimvis.prepare(lda, corpus, gensim_dictionary)  
pyLDAvis.display(vis_data)
```

<IPython.core.display.HTML object>

Выведем наиболее часто встречающиеся слова для каждой темы, что бы понять, какие фильмы находятся в каждой теме

```
for i in range(7):  
    gensim_dictionary =  
gensim.corpora.Dictionary(df.loc[df["cluster"].isin([i]),  
"tokenized_preprocessed_text"])  
    doc_count = len(df.loc[df["cluster"].isin([i]),  
"tokenized_preprocessed_text"])  
  
    if doc_count >= 100:  
        gensim_dictionary.filter_extremes(no_above=0.1, no_below=5)  
    elif doc_count >= 50:  
        gensim_dictionary.filter_extremes(no_above=0.3, no_below=3)  
    else:  
        gensim_dictionary.filter_extremes(no_above=0.5, no_below=2)  
  
    gensim_dictionary.compactify()  
  
    corpus = [gensim_dictionary.doc2bow(text) for text in  
df.loc[df["cluster"].isin([i]), "tokenized_preprocessed_text"]]  
    lda = gensim.models.LdaMulticore(corpus, num_topics=5,  
id2word=gensim_dictionary, passes=10, random_state=42)  
  
    print(f"\nТемы кластера {i}:")  
  
    for j in lda.print_topics():  
        print(j)
```

Темы кластера 0:

```
(0, '0.011*"миграция" + 0.005*"нативный" + 0.005*"фронтенд" +  
0.004*"сетка" + 0.004*"доступность" + 0.004*"эксперт" +  
0.004*"партнёр" + 0.004*"студент" + 0.004*"твой" + 0.003*"дизайнер"')  
(1, '0.122*"лицензия" + 0.062*"рейтинг" + 0.012*"узел" +  
0.010*"оглавление" + 0.009*"пайплайна" + 0.008*"распределённый" +  
0.007*"граф" + 0.007*"тег" + 0.006*"коммит" + 0.006*"движок"')
```

(2, '0.022\*"анимация" + 0.014\*"герой" + 0.012\*"игровой" + 0.009\*"враг" + 0.007\*"счётчик" + 0.007\*"игрок" + 0.006\*"кадр" + 0.005\*"кластер" + 0.005\*"движок" + 0.004\*"синхронизация"')  
(3, '0.012\*"диск" + 0.008\*"терминал" + 0.007\*"контроллер" + 0.006\*"хост" + 0.006\*"эмулятор" + 0.005\*"буфер" + 0.004\*"атрибут" + 0.004\*"служба" + 0.004\*"последовательный" + 0.004\*"курсор"')  
(4, '0.009\*"метрика" + 0.008\*"тестировщик" + 0.007\*"уведомление" + 0.006\*"кластер" + 0.005\*"трассировка" + 0.005\*"инцидент" + 0.005\*"профиль" + 0.005\*"алерт" + 0.004\*"нагрузочный" + 0.004\*"заказчик"')

#### Темы кластера 1:

(0, '0.005\*"пользователь" + 0.005\*"ит" + 0.004\*"история" + 0.004\*"доступ" + 0.003\*"компьютер" + 0.003\*"игра" + 0.003\*"вуз" + 0.003\*"сеть" + 0.003\*"блок" + 0.003\*"доступный"')  
(1, '0.035\*"тестирование" + 0.009\*"тестировщик" + 0.005\*"баг" + 0.005\*"раздел" + 0.005\*"требование" + 0.004\*"кейс" + 0.004\*"рассматриваться" + 0.004\*"выпускник" + 0.004\*"модуль" + 0.003\*"автоматизация"')  
(2, '0.011\*"урок" + 0.008\*"обучать" + 0.008\*"модуль" + 0.008\*"стоимость" + 0.008\*"текст" + 0.006\*"редактор" + 0.006\*"руб" + 0.006\*"маркетинг" + 0.005\*"александр" + 0.005\*"контент"')  
(3, '0.005\*"фреймворк" + 0.004\*"вуз" + 0.004\*"ошибка" + 0.003\*"резюме" + 0.003\*"карьерный" + 0.003\*"сообщество" + 0.003\*"интервью" + 0.003\*"частый" + 0.003\*"директор" + 0.003\*"идея"')  
(4, '0.017\*"практикум" + 0.006\*"олимпиада" + 0.005\*"тренажёр" + 0.005\*"проектный" + 0.005\*"спринт" + 0.005\*"выпускник" + 0.005\*"ошибка" + 0.004\*"участник" + 0.004\*"наставник" + 0.004\*"тестирование"')

#### Темы кластера 2:

(0, '0.007\*"маска" + 0.005\*"нейрон" + 0.003\*"вирус" + 0.003\*"кора" + 0.003\*"сенсорный" + 0.003\*"аккумулятор" + 0.002\*"ток" + 0.002\*"пароль" + 0.002\*"инфекция" + 0.002\*"коронавирус"')  
(1, '0.015\*"денис" + 0.006\*"рой" + 0.005\*"кресло" + 0.005\*"рейтинг" + 0.004\*"боец" + 0.004\*"дэн" + 0.003\*"макс" + 0.003\*"кожа" + 0.003\*"призрак" + 0.003\*"марсианин"')  
(2, '0.007\*"сергей" + 0.006\*"зуб" + 0.003\*"евгений" + 0.003\*"дмитрий" + 0.003\*"визитка" + 0.003\*"рюкзак" + 0.003\*"домен" + 0.003\*"процессор" + 0.003\*"трейлер" + 0.003\*"гаджет"')  
(3, '0.006\*"макс" + 0.004\*"иван" + 0.003\*"семён" + 0.003\*"бот" + 0.003\*"чип" + 0.003\*"мама" + 0.002\*"персонаж" + 0.002\*"экзамен" + 0.002\*"тень" + 0.002\*"мечта"')  
(4, '0.005\*"ия" + 0.004\*"подпись" + 0.002\*"прототип" + 0.002\*"распознавание" + 0.002\*"полёт" + 0.002\*"джефф" + 0.002\*"нейросеть" + 0.002\*"маркетинг" + 0.002\*"пилот" + 0.002\*"предприятие"')

#### Темы кластера 3:

(0, '0.006\*"автотест" + 0.005\*"опросить" + 0.005\*"ручной" +

0.004\*"февраль" + 0.004\*"фрилансер" + 0.003\*"кадровый" + 0.003\*"пк" + 0.003\*"бэклога" + 0.003\*"подбор" + 0.003\*"митап"')  
 (1, '0.009\*"респондент" + 0.006\*"отель" + 0.006\*"модератор" + 0.005\*"авить" + 0.004\*"альфа" + 0.004\*"инвестор" + 0.004\*"бронирование" + 0.004\*"стажировка" + 0.004\*"август" + 0.004\*"сто"')  
 (2, '0.009\*"доллар" + 0.008\*"руб" + 0.006\*"стенд" + 0.005\*"октябрь" + 0.005\*"акция" + 0.005\*"миллиард" + 0.003\*"спонсор" + 0.003\*"товар" + 0.003\*"производство" + 0.003\*"р"')  
 (3, '0.005\*"компьютерный" + 0.004\*"соответствие" + 0.004\*"объект" + 0.004\*"осуществлять" + 0.004\*"защита" + 0.004\*"микросервис" + 0.003\*"поведение" + 0.003\*"п" + 0.003\*"стресс" + 0.003\*"эксплуатация"')  
 (4, '0.010\*"производственный" + 0.005\*"диджитал" + 0.005\*"эйчар" + 0.005\*"маркетинговый" + 0.005\*"прибыль" + 0.004\*"поездка" + 0.004\*"юридический" + 0.004\*"предприятие" + 0.004\*"консультант" + 0.003\*"водитель"')

#### Темы кластера 4:

(0, '0.012\*"стоимость" + 0.012\*"машина" + 0.011\*"цена" + 0.009\*"алгоритм" + 0.009\*"спрос" + 0.008\*"фас" + 0.008\*"район" + 0.006\*"минута" + 0.005\*"всё" + 0.005\*"большой"')  
 (1, '0.019\*"доставка" + 0.009\*"курьер" + 0.008\*"клиент" + 0.006\*"всё" + 0.006\*"товар" + 0.006\*"уборка" + 0.006\*"еда" + 0.006\*"дом" + 0.005\*"возможность" + 0.005\*"продукт"')  
 (2, '0.008\*"российский" + 0.008\*"пункт" + 0.007\*"доступ" + 0.007\*"интернет" + 0.006\*"список" + 0.006\*"законодательство" + 0.006\*"включая" + 0.006\*"связь" + 0.006\*"деятельность" + 0.005\*"ресурс"')  
 (3, '0.022\*"доставка" + 0.011\*"наш" + 0.010\*"система" + 0.009\*"клиент" + 0.007\*"магазин" + 0.007\*"данные" + 0.006\*"получить" + 0.006\*"товар" + 0.006\*"провайдер" + 0.005\*"дать"')  
 (4, '0.016\*"рубль" + 0.008\*"машина" + 0.007\*"доставка" + 0.007\*"наш" + 0.006\*"маршрут" + 0.006\*"клиент" + 0.006\*"тысяча" + 0.005\*"несколько" + 0.005\*"около" + 0.005\*"петербург"')

#### Темы кластера 5:

(0, '0.014\*"сказать" + 0.009\*"наш" + 0.008\*"ваш" + 0.008\*"знать" + 0.008\*"делать" + 0.007\*"велосипед" + 0.007\*"модуль" + 0.006\*"сценарий" + 0.006\*"голова" + 0.006\*"хотеть"')  
 (1, '0.007\*"российский" + 0.007\*"пока" + 0.006\*"хороший" + 0.006\*"функция" + 0.006\*"использование" + 0.005\*"сервисный" + 0.005\*"однако" + 0.005\*"млн" + 0.005\*"наш" + 0.004\*"выглядеть"')  
 (2, '0.010\*"выставка" + 0.008\*"позволять" + 0.006\*"управление" + 0.006\*"проект" + 0.006\*"большой" + 0.006\*"представитель" + 0.006\*"различный" + 0.006\*"установка" + 0.006\*"мозг" + 0.006\*"подобный"')  
 (3, '0.015\*"взаимодействие" + 0.013\*"исследование" + 0.013\*"машина" + 0.010\*"этап" + 0.009\*"пользователь" + 0.007\*"общение" +

```
0.007*"рабочий" + 0.006*"получать" + 0.006*"задача" + 0.006*"оценка"')
(4, '0.005*"управление" + 0.005*"позволять" + 0.005*"инженер" +
0.004*"история" + 0.004*"количество" + 0.004*"новость" +
0.004*"камера" + 0.004*"собирать" + 0.004*"устройство" +
0.004*"фотография"')
```

Темы кластера 6:

```
(0, '0.010*"категория" + 0.007*"доллар" + 0.007*"тыс" + 0.007*"домен"
+ 0.007*"провайдер" + 0.007*"хостинг" + 0.006*"договор" +
0.005*"сумма" + 0.004*"скидка" + 0.004*"запрос"')
(1, '0.007*"рассказ" + 0.007*"государство" + 0.006*"описывать" +
0.005*"тест" + 0.005*"компонент" + 0.005*"продуктовый" +
0.005*"доллар" + 0.005*"доступность" + 0.004*"геллер" +
0.004*"адаптивный"')
(2, '0.017*"никтолай" + 0.016*"фильм" + 0.011*"контекст" +
0.010*"отзыв" + 0.008*"нарисовать" + 0.007*"диапазон" + 0.007*"ру" +
0.007*"н" + 0.007*"генерировать" + 0.006*"руб"')
(3, '0.008*"позиция" + 0.006*"засечка" + 0.006*"промышленный" +
0.005*"компонент" + 0.004*"президент" + 0.004*"скрипт" +
0.004*"карьеря" + 0.004*"отображение" + 0.004*"красивый" +
0.004*"красота"')
(4, '0.018*"идентификация" + 0.012*"фирменный" + 0.010*"маркетинг" +
0.008*"фирма" + 0.007*"брендинг" + 0.005*"клавиатура" +
0.005*"твиттер" + 0.004*"лого" + 0.004*"собственно" +
0.004*"потребитель"')
```

Проанализировав полученный результат, по полученным словам можно выделить следующие темы:

- "IT-разработка"
- "IT-образование"
- "Технологии будущего"
- "IT в бизнесе"
- "Инфраструктура городских сервисов"
- "Технологии и человеческий фактор"
- "Брендинг и маркетинг"

Добавим в набор данных информацию о принадлежности к теме

```
df["cluster name"] = df["cluster"].map({
    0: "IT-разработка",
    1: "IT-образование",
    2: "Искусственный интеллект",
    3: "IT в бизнесе",
    4: "Инфраструктура городских сервисов",
    5: "Технологии и человеческий фактор",
    6: "Брендинг и маркетинг",
})
```

```
df["cluster name"].value_counts()

cluster name
Искусственный интеллект      453
IT-разработка                208
IT в бизнесе                  170
Брендинг и маркетинг         134
IT-образование                76
Инфраструктура городских сервисов  49
Технологии и человеческий фактор  17
Name: count, dtype: int64

df.to_csv("habr articles (cluster).csv", index=False)

pd.set_option("display.max_colwidth", 50)

df = pd.read_csv("habr articles (cluster).csv")
df.sample(1)
```

	text \
8	Когда организации переходят в облако, их систе...

	preprocessed_text	cluster \
8	организация переходить облако система начинать...	0

	tokenized_preprocessed_text	cluster name
8	['организация', 'переходить', 'облако', 'систе...]	IT-разработка

## Классификация

Для классификации нам нужно разделить набор данных на обучающую и тестовую выборку. Воспользуемся методом `train_test_split`

```
from sklearn.model_selection import train_test_split

# Оставим 66% для обучающей выборки и 33% для тестовой выборки
X_train, X_test, y_train, y_test = train_test_split(W, df["cluster"],
test_size=0.33, random_state=42)

print(X_train.shape)
print(X_test.shape)

(741, 7)
(366, 7)
```

Для кластеризации возьмём 3 модели:

KNeighbors это алгоритм классификации, который предполагает, что расположенные близко друг к другу объекты принадлежат к одному классу. Данный алгоритм популярен и подходит для нашей задачи, возьмём его для сравнения результатов



Random Forest это множество решающих деревьев, это позволяет снизить переобучения и повысить точность в сравнении с одним деревом.

DecisionTreeClassifier - это класс, способный выполнять многоклассовую классификацию на наборе данных.

Импортируем модели

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
```

Далее обучим каждую модель

```
%%time
rfc = RandomForestClassifier(random_state=42)
rfc.fit(X_train, y_train);
rfc_predict = rfc.predict(X_test)

CPU times: total: 141 ms
Wall time: 139 ms

%%time
dtc = DecisionTreeClassifier(random_state=42)
dtc.fit(X_train, y_train);
dtc_predict = dtc.predict(X_test)

CPU times: total: 0 ns
Wall time: 2.99 ms

%%time
knn = KNeighborsClassifier(weights='distance')
knn.fit(X_train, y_train);
knn_predict = knn.predict(X_test)

CPU times: total: 15.6 ms
Wall time: 17.9 ms
```

Модели обучены, теперь сравним их метрики

Для оценки качества моделей необходимо понимать метрики оценки моделей.

Precision и Recall можно описать следующим образом:

- Precision - "Сколько выбранных элементов являются релевантными"
- Recall - "Сколько релевантных элементов выбранно"

Precision демонстрирует способность алгоритма отличать один класс от других классов, а Recall демонстрирует способность алгоритма обнаруживать данный класс.

Precision и recall не зависят, в отличие от accuracy, от соотношения классов и потому применимы в условиях несбалансированных выборок.

F1-score — это гармоническое среднее между Precision и Recall, которое используется для оценки баланса между точностью и полнотой модели. Он особенно полезен в задачах с несбалансированными классами.

Поскольку классы несбалансированными, для оценки качества обучения, как основную метрику, будем использовать F1-score.

```
from sklearn.metrics import classification_report
print(classification_report(y_test, rfc_predict))
```

	precision	recall	f1-score	support
0	0.96	0.93	0.95	84
1	0.88	0.88	0.88	26
2	0.95	0.98	0.96	142
3	0.93	0.93	0.93	58
4	1.00	1.00	1.00	14
5	1.00	1.00	1.00	4
6	0.94	0.89	0.92	38
accuracy			0.95	366
macro avg	0.95	0.95	0.95	366
weighted avg	0.95	0.95	0.95	366

```
print(classification_report(y_test, dtc_predict))
```

	precision	recall	f1-score	support
0	0.93	0.96	0.95	84
1	0.88	0.88	0.88	26
2	0.97	0.95	0.96	142
3	0.95	0.95	0.95	58
4	1.00	1.00	1.00	14
5	1.00	1.00	1.00	4
6	0.92	0.92	0.92	38
accuracy			0.95	366
macro avg	0.95	0.95	0.95	366
weighted avg	0.95	0.95	0.95	366

```
print(classification_report(y_test, knn_predict))
```

	precision	recall	f1-score	support
0	0.98	0.99	0.98	84
1	0.91	0.81	0.86	26
2	0.93	0.96	0.94	142
3	0.90	0.97	0.93	58
4	1.00	0.93	0.96	14

	5	1.00	1.00	1.00	4
	6	1.00	0.87	0.93	38
accuracy				0.95	366
macro avg		0.96	0.93	0.94	366
weighted avg		0.95	0.95	0.94	366

Показатели очень хорошие, лучший F1-score у KNeighborsClassifier. Сохраним эту модель для последующего предсказания

```
from joblib import dump

dump(rfc, "RandomForestClassifier.joblib")
dump(nmf_model, "nmf_model.joblib")
dump(tfidf_vectorizer, "tfidf_vectorizer.joblib")

['tfidf_vectorizer.joblib']
```

## Результат проделанной работы

### Была проведена предварительная обработка данных

- Приведение к нижнему регистру
- Очистка от латинских символов
- Очистка от пунктуации и спец символов
- Очистка от цифр
- Очистка от лишних пробеллов
- Очистка от стоп слов
- Токенизация
- Лемматизация

**Отбор признаков** Для кластеризации использована модель KMeans, так как она проста и быстра в реализации, а так же эффективна при работе с большими наборами данных

Для кластеризации взяты модели KNeighborsClassifier, DecisionTreeClassifier и RandomForestClassifier.

- KNeighbors это алгоритм классификации, который предполагает, что расположенные близко друг к другу объекты принадлежат к одному классу. Данный алгоритм популярен и подходит для нашей задачи, возьмём его для сравнения результатов
- Random Forest это множество решающих деревьев, это позволяет снизить переобучения и повысить точность в сравнении с одним деревом.
- DecisionTreeClassifier - это класс, способный выполнять многоклассовую классификацию на наборе данных.

Для оценки качества обучения использована метрика F1-score.

Каждый кластер был поименован:

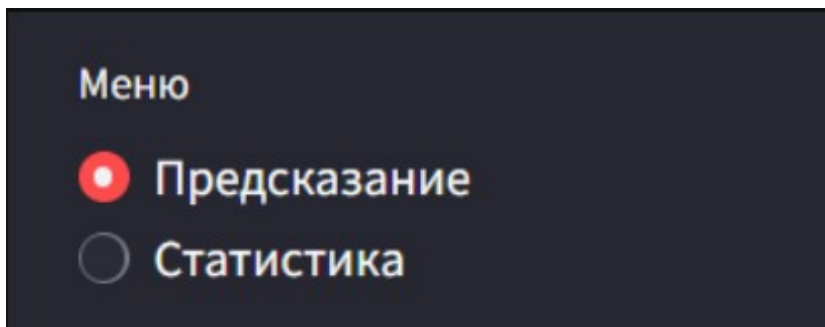
- "IT-разработка"
- "IT-образование"
- "Технологии будущего"
- "IT в бизнесе"
- "Инфраструктура городских сервисов"
- "Технологии и человеческий фактор"
- "Брендинг и маркетинг"

Выведена информация о количестве статей для каждого кластера

## Web приложение

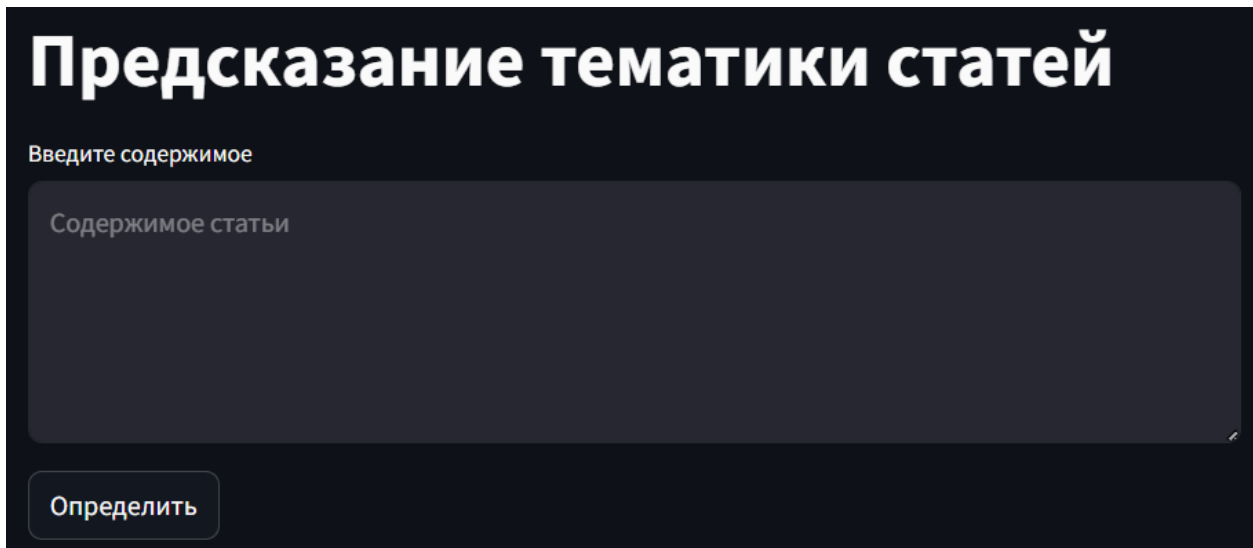
Для предсказания тематики статей создано web приложение на основе библиотек fastapi и streamlit

Приложение имеет меню навигации в левой части экрана



### Предсказание

На основной вкладке "Предсказание" имеется форма ввода текста с кнопкой определения

A dark-themed form titled "Предсказание тематики статей" (Prediction of article topics). Below the title is a label "Введите содержимое" (Enter content) and a large text input area with a placeholder "Содержимое статьи" (Article content). At the bottom left of the form is a button labeled "Определить" (Determine).

По нажатию кнопки отправляется http запрос к API, в результате которого пользователь получает распределение вероятностей тем

# Предсказание тематики статей

Введите содержимое

Нидерландская компания Yandex N.V. по итогам реструктуризации не сможет создавать и развивать сервисы, аналогичные бизнесам «Яндекса» в течение пяти лет по всему миру. Об этом говорится в документах по реструктуризации бизнеса. Отмечается, что исключением будут выделяемые зарубежные стартапы Avride, Toloka, Nebius и TripleTen, которые можно будет развивать везде, кроме России и Беларуси. «При этом независимо от рынка, им запрещено развиваться так, чтобы создавать конкуренцию сервисам

Определить

Инфраструктура городских сервисов - 0.890 %

Искусственный интеллект - 0.080 %

Технологии и человеческий фактор - 0.020 %

IT-образование - 0.010 %

IT-разработка - 0.000 %

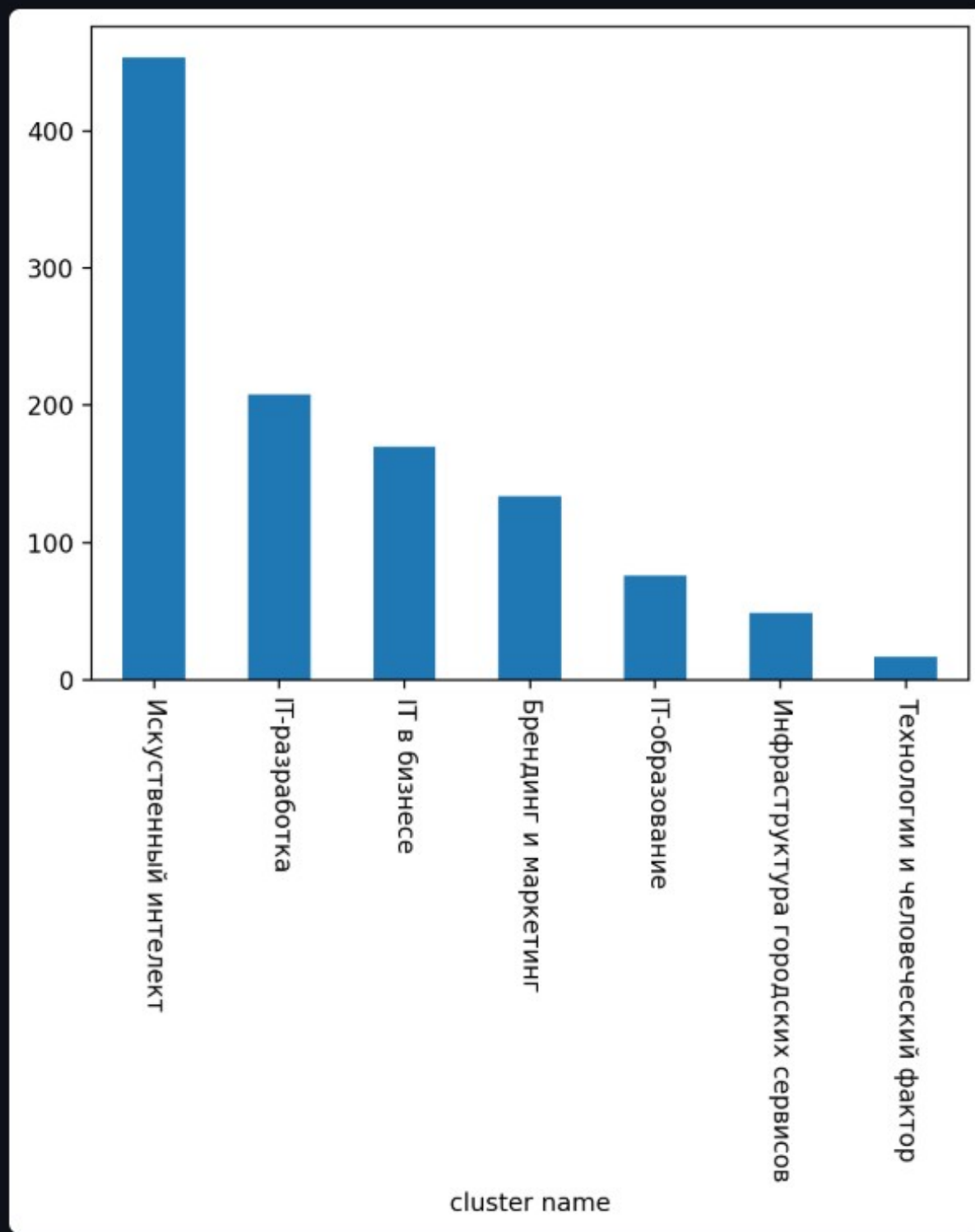
IT в бизнесе - 0.000 %

Брендинг и маркетинг - 0.000 %

## Статистика

На вкладке "Статистика" имеется информация о наборе данных, сформированном в результате проделанной работы

Распределение статей:



Код API:

```

import re
import string

import fastapi
import joblib
import nltk
import pymorphy3

app = fastapi.FastAPI()

def preprocess_text(text: str) -> str:
    punctuation = set(string.punctuation) - { "(", ")" }
    stopwords = nltk.corpus.stopwords.words("russian") + [
        "хабp", "хабpa"
    ]

    morph = pymorphy3.MorphAnalyzer()

    text = "".join([char for char in text if char not in
punctuation])
    text = "".join([char for char in text if not char.isdigit()])
    text = re.sub(r"[a-z]", "", text)
    text = re.sub(r"\s+", " ", text.strip())
    tokenized_text = re.split(r"\W+", text)
    tokenized_text = [morph.parse(word)[0].normal_form for word in
tokenized_text]
    tokenized_text = [word for word in tokenized_text if word not in
stopwords]
    return " ".join(tokenized_text)

def predict_cluster(text):
    with open('RandomForestClassifier.joblib', 'rb') as file:
        model = joblib.load(file)

    with open('tfidf_vectorizer.joblib', 'rb') as file:
        vectorizer = joblib.load(file)

    with open('nmf_model.joblib', 'rb') as file:
        nmf_model = joblib.load(file)

    tfidf_matrix = vectorizer.transform([preprocess_text(text)])
    W = nmf_model.transform(tfidf_matrix)

    prediction = model.predict(W)
    probabilities = model.predict_proba(W)[0]

    mapping = {
        0: "IT-разработка",
        1: "IT-образование",
        2: "Искусственный интеллект",
        3: "IT в бизнесе",
    }

```

```

        4: "Инфраструктура городских сервисов",
        5: "Технологии и человеческий фактор",
        6: "Брендинг и маркетинг",
    }

    probabilities_dict = { mapping[i]: float(probabilities[i]) for i
in range(len(probabilities)) }

    return {
        "probabilities" : probabilities_dict
    }

@app.post("/predict")
def predict_class(text: str):
    return predict_cluster(text)

```

Код View

```

import streamlit as st
import requests

import pandas as pd
import matplotlib.pyplot as plt

def request_predict(text: str):
    return requests.post("http://127.0.0.1:8002/predict",
params={"text": text})

def main():
    page = st.sidebar.radio("Меню", ["Предсказание", "Статистика"])

    if page == "Предсказание":
        st.title("Предсказание тематики статей")
        input_text = st.text_area("Введите содержимое", height=150,
placeholder="Содержимое статьи")

        if st.button("Определить"):
            if input_text.strip():
                try:
                    response = request_predict(input_text)
                    response.raise_for_status()
                    result = response.json()

                    probabilities = result["probabilities"]
                    sorted_probabilities =
sorted(probabilities.items(), key=lambda x: x[1], reverse=True)

                    try:
                        for comment_nature, probability in
sorted_probabilities:

```



```

                                st.write(f"{comment_nature} -
{probability:.3f} %")

                                except Exception as e:
                                    st.error(f"Ошибка обработки вероятностей:
{e}")

                                except requests.exceptions.RequestException as e:
                                    st.error(f"Ошибка соединения с API: {e}")
                                else:
                                    st.warning("Пожалуйста, введите текст")

                                else:
                                    df = pd.read_csv("habr articles (cluster).csv",
keep_default_na=False)
                                    total_len = len(df)

                                    st.markdown(f"Общая длина набора данных: {total_len}")
                                    st.markdown(f"")

                                    st.markdown(f"Распределение статей:")
                                    counts = df["cluster name"].value_counts()

                                    fig, ax = plt.subplots()
                                    counts.plot(kind="bar", ax=ax)
                                    ax.set_xticklabels(counts.index, rotation=-90)
                                    st.pyplot(fig)

                                if __name__ == "__main__":
                                    main()

```