

Архитектура и системы для комбинированных рекомендаций и поиска эстетичных изображений

Келим Илья, 17.Б11-мм

Научный руководитель: к.т.н., доцент Брыксин Т.А.

Консультант: Основатель “graphica.ai” Брыксин М.А.

Введение

Предметная область:

Дизайнерский контент – изображения, демонстрирующие работы или идеи дизайнеров.

Практическая ценность:

Облегчение поиска вдохновения для дизайнеров.

Постановка задачи

Цель

Целью данной работы является разработка **архитектуры** и создание системы **комбинированных рекомендаций** и **поиска** для дизайнерского контента

Задачи

- Обзор существующих решений
- Реализовать модуль поиска изображений, похожих по **доминантным цветам**
- Реализовать модуль поиска изображений, похожих по **текстовой информации** (автор, ключевые слова в описании, название стиля)
- Реализовать модуль поиска изображений, похожих по **визуальному стилю**
- Спроектировать общую архитектуру системы и **интегрировать модули** друг с другом
- Реализовать прототип системы и **провести апробацию** в проекте Graphica.ai

Пример поиска в интерфейсе

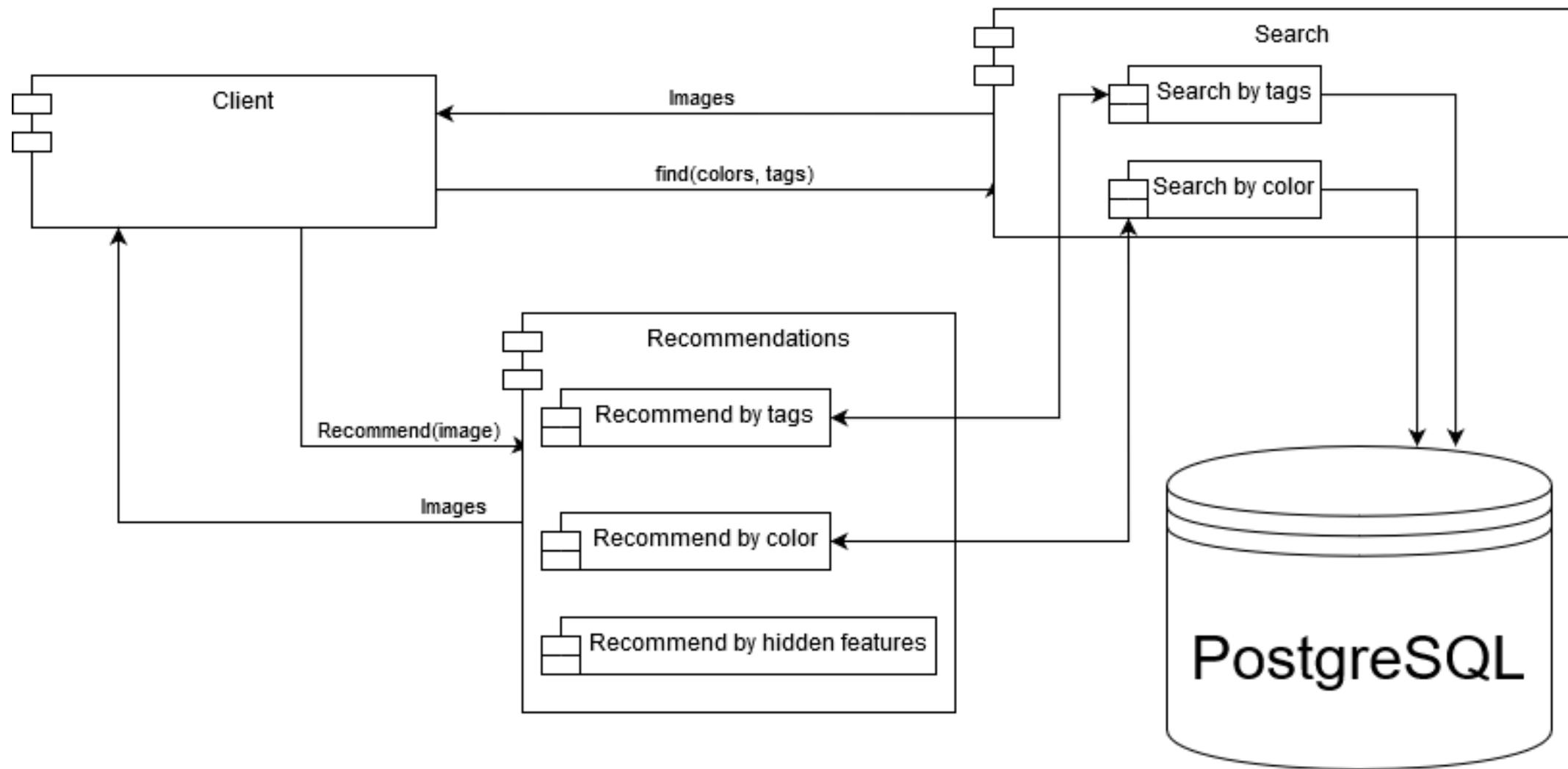
The screenshot displays the Graphica.ai search interface. On the left is a sidebar with navigation options: Collections (Visual, Patterns, Components like Checkbox, Datepicker, Dropdown, File Upload; Pages, Emails, Navigation; Research), Boards (Instagram Post #1, Domains & Nabvar), and Add Collection/Board buttons. The main search bar contains the query "#E93423 #FFCE45 Graphic Design". The search results are a grid of six items:

- Top Left:** A magazine spread with a red background and yellow stars. Text includes "INTELLIGENCE IN LIFESTYLE", "OTTOBRE 2009", "13", "Verde, a destra!", "L'altra Engadina", and "L'amo nel traffico". At the bottom, Chinese characters "中國綜合症" and "(Sindrome cinese)" are displayed, along with the text "Come e perché il più grande Paese comunista del mondo sta salvando il Capitalismo".
- Top Middle:** A woman in a grey cardigan on a red background. Text includes "boldybae", "Next Color ORANGE", and "– Cuz i'm still looking looking for trouble".
- Top Right:** A collage of magazine pages with red and white elements. Text includes "Start Art", "You say it!", "nora.", "A doll's House", and "Neverready".
- Bottom Middle:** A blue ribbon graphic on a red background with the text "THE PHENOM".
- Bottom Right:** A red background with the text "Bloomberg Businessweek" and "Neverready Inside the slow-motion collapse of".

Существующие решения

- Агрегаторы изображений (Google Images, Yandex Images)
 - Работают в более широкой области, не учитывают специфику
 - Не имеют рекомендательных систем
- Дизайнерские платформы (Dribbble, Behance)
 - Рекомендательные и поисковые системы не учитывают специфику предметной области
- Общие наборы данных для определения визуального стиля (aesthetic dataset)
 - Используют субъективное определение стиля
 - Не готовый продукт

Структура проекта



Поиск по цвету

1. Выделение доминантных цветов изображений
 - Задача выполняется в **отдельном потоке** для каждой новой картинки
 - **Доминантные цвета** вычисляются с помощью кластеризации (scikit-learn)
 - Цвета переводятся в цветовое пространство «**Mean Shift**» (CV2)
2. Построение индекса векторов
 - Построение индекса выполняется в отдельном потоке **раз в сутки**
 - Для каждой картинки цвета объединяются в **вектор**
 - Вектор **балансируется** для повышения точности поиска
 - Все вектора объединяются в **индекс** и записываются на диск
3. Использование индекса для поиска ближайших соседей данного вектора
 - Для заданного цвета строится вектор, для которого ищутся похожие (**Approximate Nearest Neighbor**)

Поиск по тексту

1. При загрузке изображения сохраняется информация, окружающая его
 - Автор
 - Название
 - Описание
 - Категория
 - Шрифт
 - Стиль
2. Полнотекстовой поиск
 - С помощью встроенного в PostgreSQL движка полнотекстового поиска **находятся вхождения** искомых слов или фраз
 - Запрос **кэшируется**

Выделение скрытых признаков

1. Подготовка нейросети

- Была выбрана нейросеть NasNet (опробована RasNet)
- Последние слои были обучены на данных aesthetics dataset

2. Выделение скрытых признаков

- Обученная нейросеть выдает для изображения вектор фиксированной длины

3. Запись вектора в базу данных

4. Построение индекса векторов

Рекомендации

По цвету

1. Поиск по цвету данной картинке

По тексту

1. Поиск категорий, автора, стилей и шрифтов

Комбинированные модули

Поиск

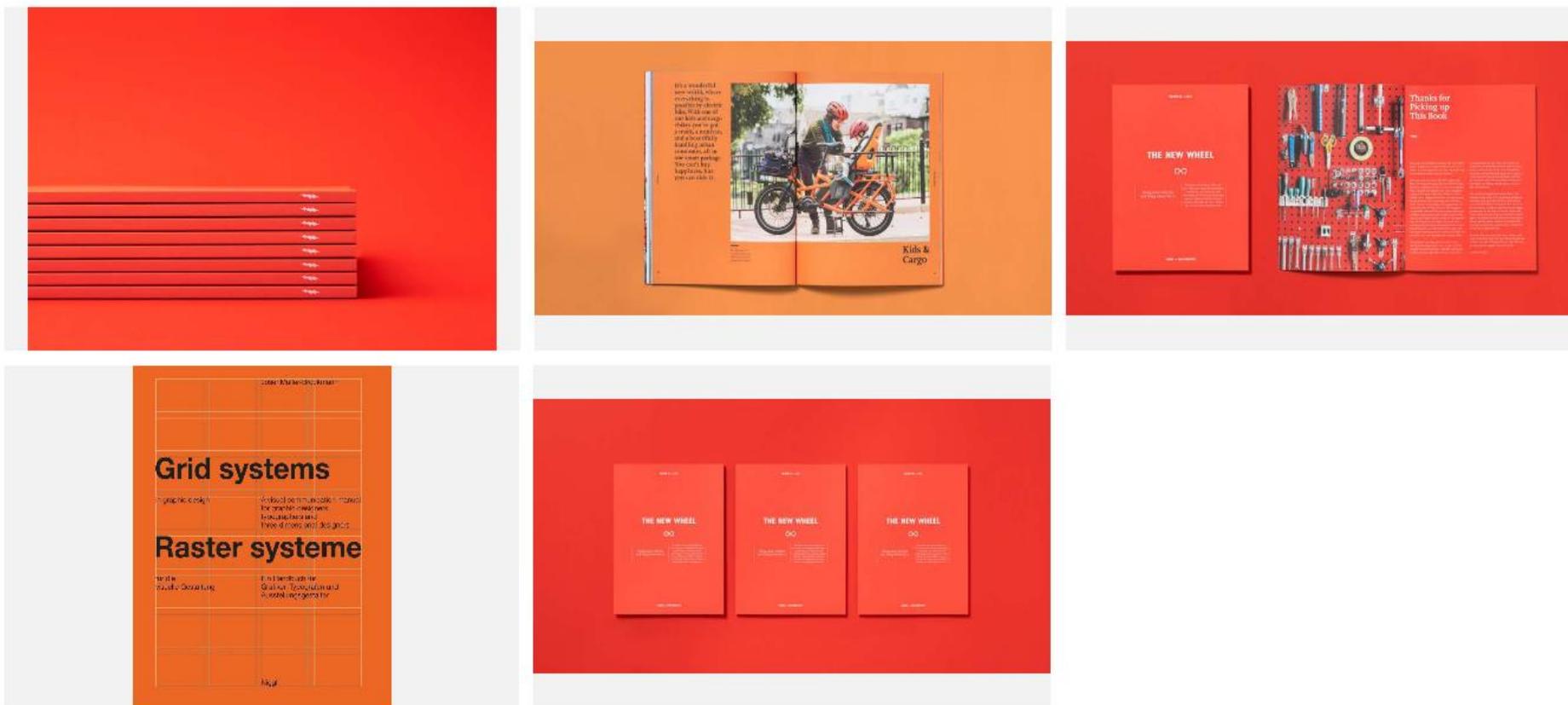
1. Разделение запроса на цвета, ключевые слова и фразы
2. Поиск по цвету
3. Поиск по тексту на результатах поиска по цвету

Рекомендации

1. Рекомендации по цвету
2. Рекомендации по тексту на результатах рекомендаций по цвету

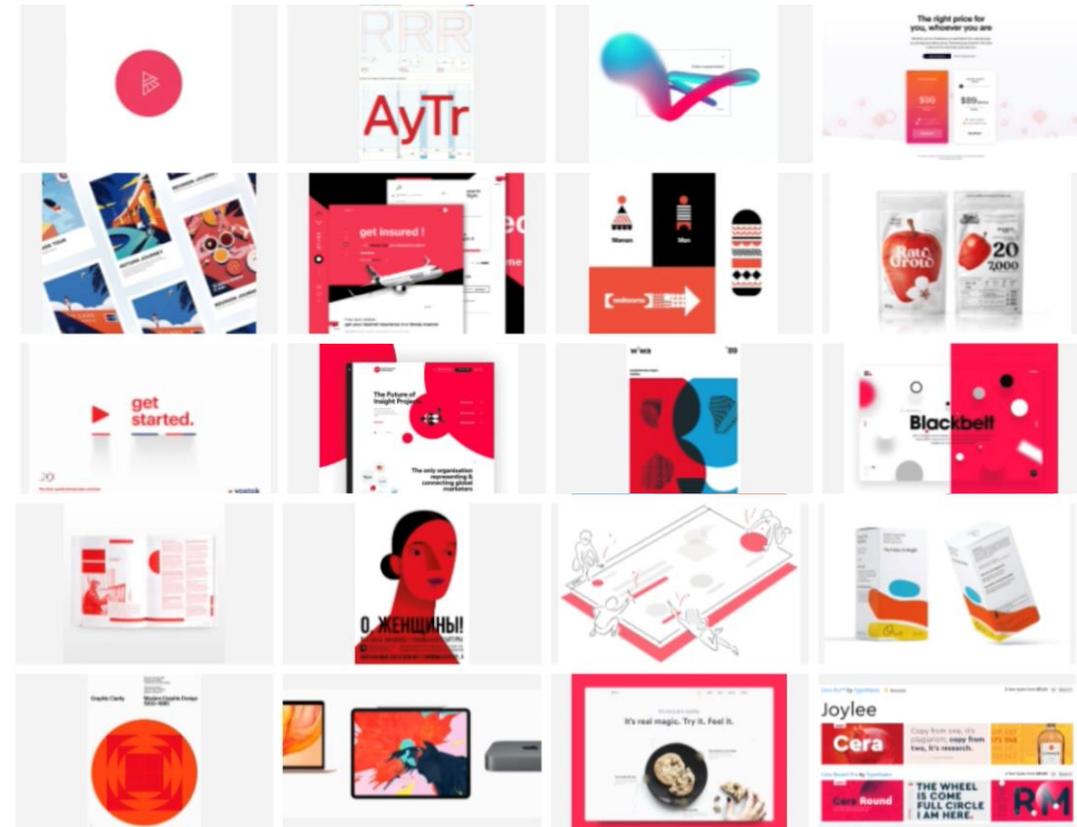
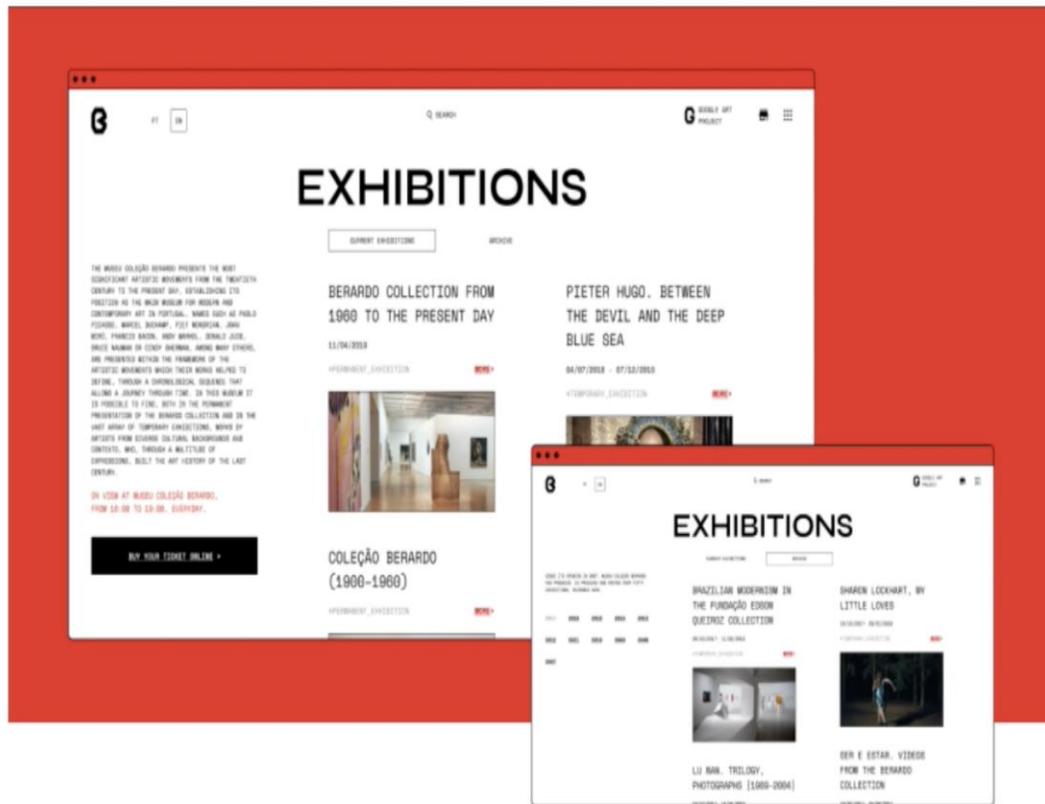
Демонстрация

Комбинированный поиск (“book”, “FF0000”)



Демонстрация

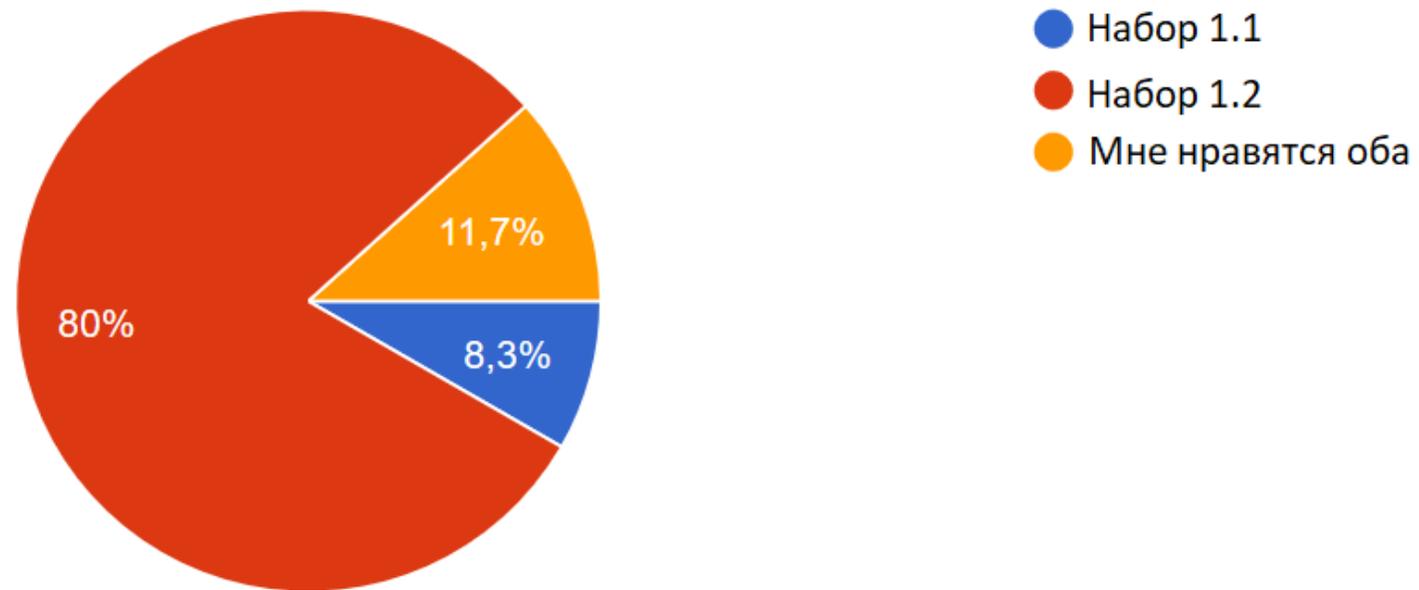
Рекомендации по цвету



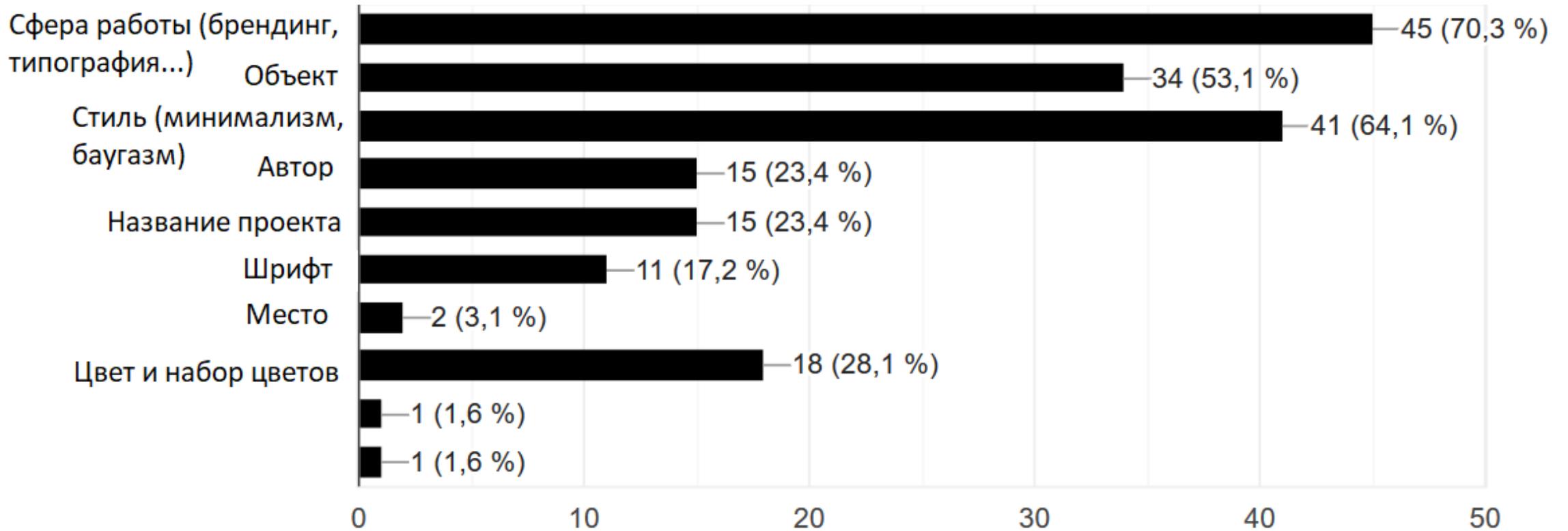
Метрики

Какой набор лучше?

60 ответов



Метрики



Результаты

- Проведен обзор существующих решений
- Реализован и апробирован модуль рекомендаций по доминантным цветам
- Реализованы и апробированы модули поиска по цвету и по тексту
- Реализован и апробирован модуль комбинированных рекомендаций по тексту и цвету
- Реализован и апробирован модуль комбинированного поиска по тексту и цвету
- Реализован модуль рекомендаций по скрытым признакам
- Результаты интегрированы в проект `graphica.ai`