LMU/LLS Theses and Dissertations

2023

LIKE YOU Search

Gipson Bachman
Loyola Marymount University, gipson.bachman@gmail.com

Recommended Citation

Like You Search

by

Gipson Bachman

A capstone project presented to the

Faculty of the Departments of
Computer Science and System Engineering
Loyola Marymount University

In partial fulfillment of the
Requirements for the Degree
Master of Science in Computer Science

April 28, 2022

Like YOU

Find social media influencers that look like you.

LIKE YOU search

Gipson Bachman

Course
CMSI 694 - Capstone

Professor
Dr. Elham Ghashghai

Advisors
Seaver, Computer Science: Dr. B.J. Johnson

Date
April 28, 2022

Table of Contents

Executive Summary

Background
- 49% of consumers rely on trusted social media influencers for recommendations when making purchase decisions.
- Majority of consumers prefer following influencers who look and act like them, instead of celebrities.

Problem
There's currently no online destination where consumers can easily find social media influencers that look like them.

Methodology
Software Development Life Cycle (SDLC) Waterfall Model

Opportunity
Building a website that fills gap in the market by providing users with an easy way to find social media influencers using customizable search parameters.

Conclusion
- **Design:** all design decisions to be geared towards consumer.
- **Format:** website format best for initial MVP due to time constraints.

Methodology

Define MOE's

Establish Baseline

Market Analysis

Trade Study

Stakeholders Analysis

Requirements

Design

Development

Testing

Deployment

Maintenance

Waterfall Model

Measures of Effectiveness

Timeliness
There's currently no website where users can easily find influencers that look like you. A central MOE relies on being the first product to answer this gap in the market.

Ease of Use
Website must be intuitive and easy for users to find influencers that look like them.

Cost of Implementation
The MVP must utilize tools that are free and/or low-cost as budget is minimal.

Scale

Low   Below average   Average   Above Average   High

1   2   3   4   5

- Values given on scale from 1 to 5, 5 being greatest value
- Values given based on professional experience and industry metrics

Background

Personal

Experience

- 6+ years in influencer marketing

Motivation

- Not a standard body type, big and tall.
- Hard to shop for clothing, in-person and online.
- Personal sizing can range from XL to 4XL depending on brand, product line and/or supplier.
- **Need** for one-stop-shop to find models/influencers that looked like me using products online.

Picture: Cousin, myself, grandmother

Marketing

82%

Of consumers trust social networks to guide purchasing decisions

HUFFPOST

NEWS   POLITICS   ENTERTAINMENT   LIFE   PERSONAL   SHOPPING   VIDEO

Make This The Year You Follow Instagram Influencers Who Look Like You

49% Of consumers rely on social media influencer recommendations for these decisions

$13.8bn The amount marketers spent within the influencer space in 2021

ROI Influencer marketing campaigns earn $5.78 for every dollar spent

Source: Digital Marketing Institute, Influencer Marketing Hub, HuffPost

10

Fashion Industry

2022 metrics

36%
of total fashion sales will be online

911mm
global online consumers

$765bn
global online fashion market

30%
of all fashion related products bought online are returned compared to 8.9% when bought at brick-and-mortar stores

Source: Entrepreneur.com, MediaKix

Are there any companies currently speaking to this unique problem?
Market Gap

![True Fit logo]

TRUE FIT®

True Fit is the only platform that solves online apparel fit at scale.

Pros
- Consumer-driven
- Easy to find your size across a broad range of brands/suppliers
- Serves 17K brands, $250 billion in cross market buying behavior, and first party preference data from over 200 million registered True Fit members.

Cons
- **Zero tie-in with the growing influencer market**
- Focus on fit only, limited applications to fashion/clothing industries
- Only available to paying brands, high barrier of entry cost
- Hard to visualize fit, consumer must trust TrueFit's proprietary algorithm.

Source: TrueFit

Brand
POS
Website

Questionnaire

Suggested Sizing

13

Izea is an influencer search solution enhanced by "artificial intelligence and machine learning."

Pros

- Great for PR and brands to identify strong performing influencers within an array of broad categories (beauty, fashion, regions, etc.)
- Get content engagement, audience and performance data behind each influencer
- Campaign integration for easy influencer outreach campaigns

Cons

- **Not consumer driven**
- Only available to subscribers, high barrier of entry cost ($189/month)
- AI/ML claims questionable
- Cannot search based on appearance/fit
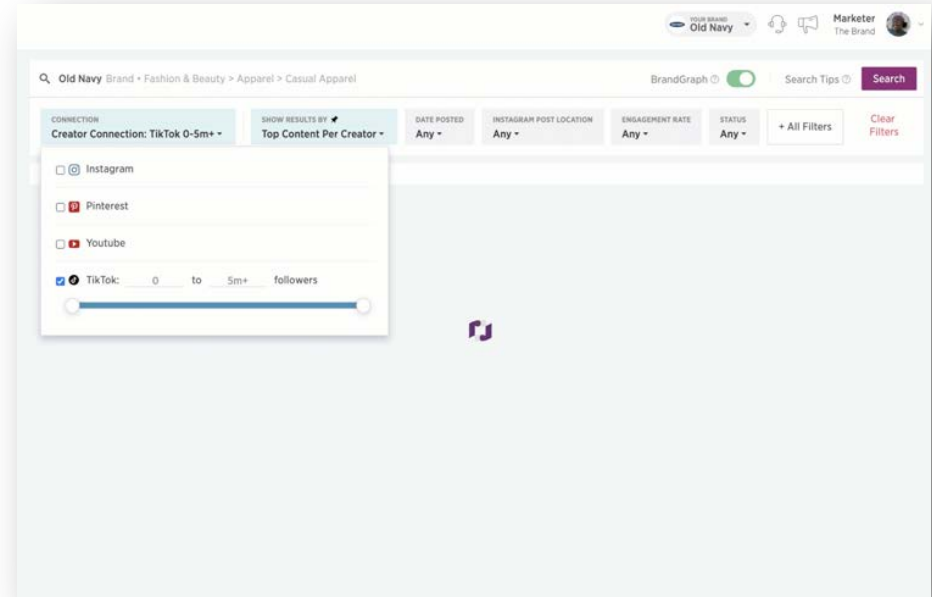
How can you currently find influencers online?
Current Baseline

Organic Discovery

When consumers find influencers through their own discovery on the internet.

Increasingly becoming more paid, either directly through paid search or indirectly through earned media (e.g. Digital Promotions, Public Relations).

Explore Tab
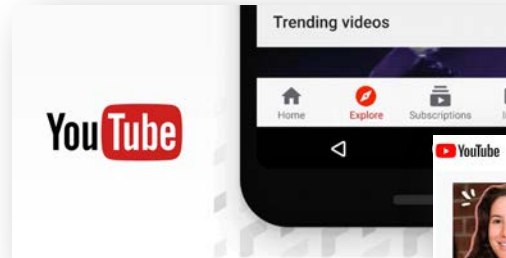
from the main page or "Explore" tab of a social network

Keyword Search
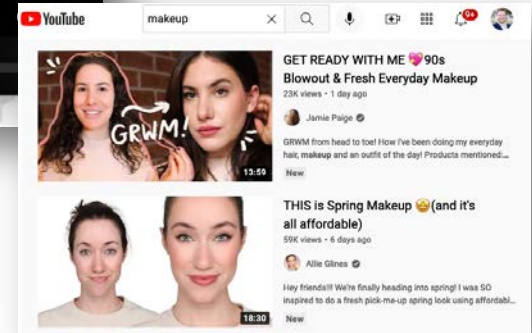
from a user's own keyword searches on the network itself or search engine like Google/Bing.

Influencer PR Lists

Searching the web for "lists" of a type of influencer.

These outlets lack search customization (if any) and are geared towards professionals within public relations and digital publicity circles, not consumers themselves.

Problem

There's currently no online destination where consumers can *easily* find social media influencers that look like them.

Solution/Objective

Building a website that fills gap in market by providing users with an easy way to find social media influencers using customizable search parameters.

Caveats and Limitations

- MVP to be consumer-driven as it answers immediate gap in market
- Limited time and resources for Spring 2022 Capstone project, choose features and tools based on development window
    - **Data:** scraped manually, wangled and cleaned for inclusion into static database, will not be dynamic
    - **Users:** to focus on general consumer and backend admin, will not support additional users (influencers, PR, brands) at this time
    - **Ecosystem:** must limit size of ecosystem due to cost, will only support one social network
    - **Format:** must choose most time/cost efficient format that the broadest consumer age demographic finds easy to use

Scope

Out of Scope

In Scope

Format

Embedded application

Website

Ecosystem/ Social Networks

Data/ API tools

Static Data Parsehub

Dynamic Data Scrapy

Frontend/ Backend Tools

Softr AirTable

Consumers Admin

Users

Influencers, PR, Brands
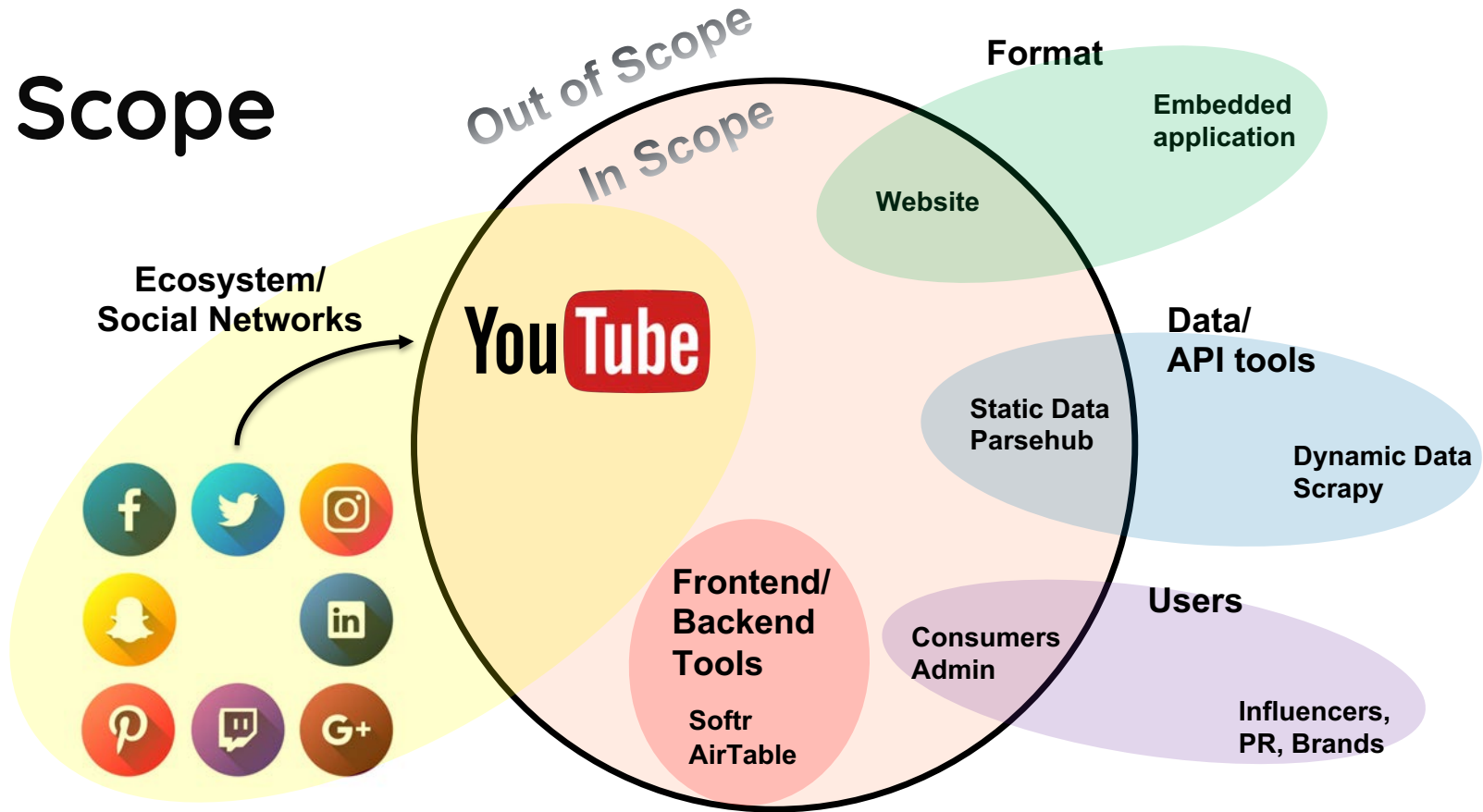
Stakeholder Analysis

1. **Consumers:** product success requires consumer adoption en masse, all design decisions to be based on value added within online customer shopping journey

2. **Admin:** viable products depends on developer/admin to create working, sustainable product

3. **Financial:** leverage key milestones to secure funding

4. **Influencers:** create space that is influencer positive, as influencers themselves can make or break platform

5. **PR/Brands:** integrate opportunities where brands can gain access to network/conversation

22

User Analysis

User (Actor)	Actions (Jobs)	Story Ending
Consumer	• Populate search parameters • Find influencers	• Click-thru to influencer on platform
Influencers	• Populate search parameters • Find influencers (or lack thereof) • Create profile, enter self-reported parameter data, contact info	• Find competition, gaps in market • Access to brand sponsorship(s)
PR/Marketing/Brands	• Populate search parameters • Find influencers • Pay for influencer introduction	• Influencer introduction • Brand spotlight within feature/slideshow
Admin	• Easily manage issues • Easily update data • Design that easily supports future build out with additional features	• Issue resolved • Data updated • Features added

MVP

23

User Analysis

Trade Study

Format

Website
Scale Value: High, 5

Pros
- Most easily used and recognized format across broadest age demographic
- Complete ownership of full experience, can easily pivot to meet consumer/market needs

Cons
- Retailers/POS as after-thought, harder to directly correlate purchase decisions as experience leaves holes for organic migration

Embedded Application
Scale Value: Below Average, 2

Pros
- Great when success metrics/MOEs directly correlated to driving purchase decisions

Cons
- Do not own full experience, as it requires third-party to embed on website/POS
- Retailer/brand driven, not consumer-focused
- Requires more time/resources to implement
- Requires participation from select online retailers
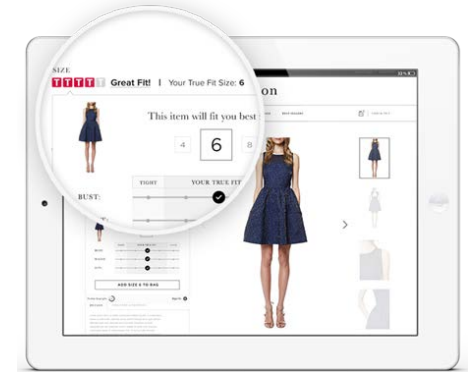
API Tools, Scraping

Scrapy
Scale Value: Below Average, 2

Pros
- Free, Open Source Tool.
- Extremely well documented.
- Easily Extensible.
- Portable Python.
- Deployment is simple and reliable.

Cons
- Time consuming to inspect and develop the crawler to simulate AJAX/PJAX requests.

Parsehub
Scale Value: Average, 3

Pros
- ParseHub allows you to collect and store data on servers automatically.
- Automatic IP rotation.
- Scraping behind logic walls allowed.
- Scaleable.

Cons
- Not free, pay service

☆ 43.3k stars
👁 1.8k watching
⑂ 9.6k forks

▲ 6 Great support
▲ 5 Easy setup
▲ 4 Complex websites
▲ 3 Native Desktop App

Scraping, Legality

Is scraping publicly accessible data legal?

Yes, it is!

On April 18, 2022, the Ninth Circuit reaffirmed its original decision and found that scraping data that is publicly accessible on the internet and not copyrighted is **not a violation of the Computer Fraud and Abuse Act,** or CFAA, which governs what constitutes computer hacking under U.S. law.

Management Tools

Frontend, Softr
Scale Value: Above Average, 4

Pros
- easy backend integration with Backend/Airtable via RESTful API calls
- responsive design

Cons
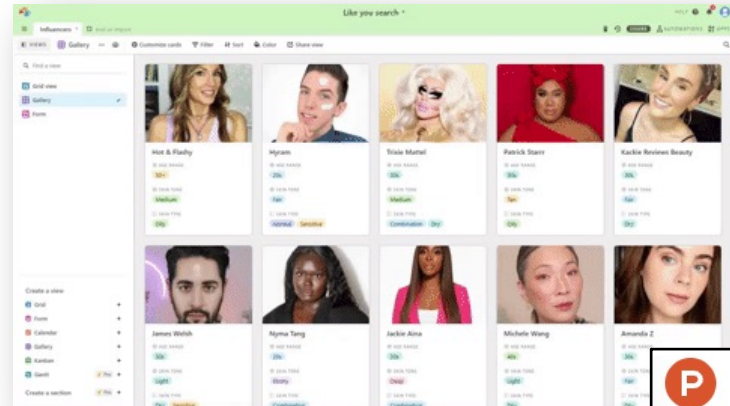- limited design features
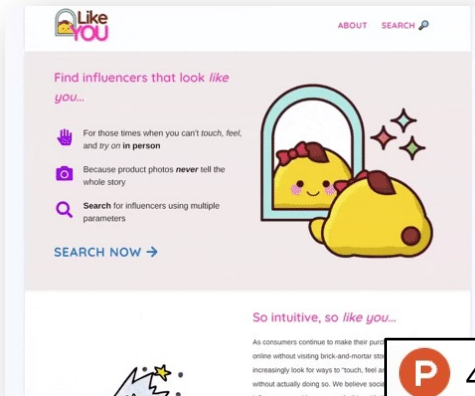
Backend, Airtable
Scale Value: Above Average, 4

Pros
- easy database management
- multi-user capabilities
- easy integration with other applications via RESTful API calls

Cons
- does not auto-refresh

P 4.5

Source: Product Hunt

P 4.5

Solution Space

	Techniques or Tools Used	Pros	Cons	Recommendation
Format	**Website** *High, 5* **vs.** **Embedded App** *Below Avg, 2*	**Website** easy start, responsive design that lends itself to desktop/mobile **Embedded app** retailer-driven	**Website** consumer-driven, retailer as after-thought **Embedded app** best for later-stage retailer-driven rollouts	**Website** • Most easily used and recognized format across broadest age demographic • Complete ownership of full experience, can easily pivot to meet consumer/market needs
API Tools	**Scrapy** *Below Avg, 2* **vs.** **Parsehub** *Avg, 3*	**Scrapy** ● Free, Open Source Tool. ● Extremely well documented. ● Easily Extensible. ● Portable Python. ● Deployment is simple and reliable. **Parsehub** ● ParseHub allows you to collect and store data on servers automatically. ● Automatic IP rotation. ● Scraping behind logic walls allowed.	**Scrapy** Time consuming to inspect and develop the crawler to simulate AJAX/PJAX requests. **Parsehub** Not free, pay service	**Parsehub** • Scraping behind logic walls allowed, needed • Scaleable, best for future
Frontend/ Backend Management	**Softr** *Above Avg, 4*	**Frontend, Softr** easy backend integration with Airtable via RESTful API calls responsive design	**Frontend, Softr** limited design features	**Softr** *and* **AirTable** • Due to time restrictons, suggest using as both tools integrate nicely together • Plays well with Parsehub for future dynamic data rollout
	Airtable *Above Avg, 4*	**Backend, AirTable** easy database management, multi-user capabilities, easy to scale in future, easy integration with other applications via RESTful API calls	**Backend, AirTable** does not auto-refresh	

30

Requirements

Requirements

Software Requirements Specifications (SRS)
The SRS outlines project purpose, scope, definitions, acronyms, abbreviations, and references. The aim here is to gather and analyze requirements while giving an in-depth playbook on how I intend to plan, design, implement and test the website project.

- Website itself – an easy-to follow GUI geared towards general web users
    - Website shall provide a uniform look for all webpages
    - Website shall provide a digital image of each influencer and links to their respective social media platforms
    - Types of GUI interfaces we will use:
        - Menus
        - Dialog boxes for seach parameters
- Backend – Database Interface for administrators
    - Form-fill Interface
- Hardware/Software Interface
    - Requires internet connection.
    - Operating System: Windows (Vista/Windows 7) or MAC OS
    - Web Browser: Internet Explorer (8.0 and above), Mozilla Firefox (3.0 and above), or Google Chrome

Requirements

Functional Requirements

ID: User Story	Epic	As a <type of user>	I want to <perform some task>	So that I can <achieve some goal>	Acceptance Criteria
1	Website	web user	visit a website where I can describe how I look	find influencers that look like me	- build working website - build all pages that meet design guidelines - ensure website connects to backend database
1.1	Search Parameters	web user	use easy search parameters	be easily guided through my LIKE YOU search	- search parameters successfully guides user throughout search process - if user does not complete search parameters, user will be notified of missing fields - user can successfully answer each question with check box(es)
1.1.1	Search Parameters	web user	check boxes that correspond to tags within influencer database	easily retrieve results with those tags	- Each individual check box corresponds with a unique tag in database - each check box successfully retrieves influencer listings in database with respective tag/parameter
1.1.2	Search Parameters	web user	be pointed towards an advanced search when there are no results	tailor how broad/narrow my search should be	- when there are no results on results page, user is directed to advanced search page

Requirements

Functional Requirements

ID: User Story	Epic	As a <type of user>	I want to <perform some task>	So that I can <achieve some goal>	Acceptance Criteria
2	Search	web user	search for influencers	retrieve a list of influencers based on a series of search tags/parameters	- uses SQL AND operator exclusively - if multiple tags, search results would show only those that meet tag1 AND tag2 AND tagN
2.1	Search	web user	search based on social media platform (e.g. YouTube)	find platform-specific influencers	- ensure check box in questionnaire successfully matches with tag/parameter in influencer database
2.2	Search	web user	search based on genre/industry (e.g. beauty, fashion, etc.)	find genre/industry specific influencers	- ensure check box in questionnaire successfully matches with tag/parameter in influencer database
2.3	Search	web user	search based on skin tone (e.g. fair, light, medium, deep, etc.)	find influencers that have the same skin tone as me	- ensure check box in questionnaire successfully matches with tag/parameter in influencer database
2.4	Search	web user	search based on skin type (e.g. dry, oily, combo, sensitive, etc.)	find influencers that have the same skin type as me	- ensure check box in questionnaire successfully matches with tag/parameter in influencer database
2.5	Search	web user	search based on eye color (e.g. blue, brown, hazel, green, etc.)	find influencers that have the same eye color as me	- ensure check box in questionnaire successfully matches with tag/parameter in influencer database
2.6	Search	web user	search based on hair color (e.g. (blonde, brown, red, black)	find influencers that have the same hair color as me	- ensure check box in questionnaire successfully matches with tag/parameter in influencer database
2.7	Search	web user	search based on age range (teens, 20s, 30s, 40s, 50s, 60+)	find influencers that have the same age range as me	- ensure check box in questionnaire successfully matches with tag/parameter in influencer database

Requirements

Functional Requirements

ID: User Story	Epic	As a <type of user>	I want to <perform some task>	So that I can <achieve some goal>	Acceptance Criteria
3	Search Parameters	web user	know the number of influencers that meet my search tags/parameters	tailor how broad/narrow my search should be	- with each check box/tag clicked, refresh # of influencers that meet those tags/parameters at the top of search page - ensure that # provided correctly corresponds with those tagged in database - this would be located on search parameters page AND advanced search page
4	Search	web user	have an advanced search feature	custom tailor my search without the search parameters pipeline	- advanced search page which all grouped parameters located on one page
5	Backend	admin	easily update database	keep data as fresh as possible	- data can be changed/updated easily on backend via web interface

Non-Functional Requirements
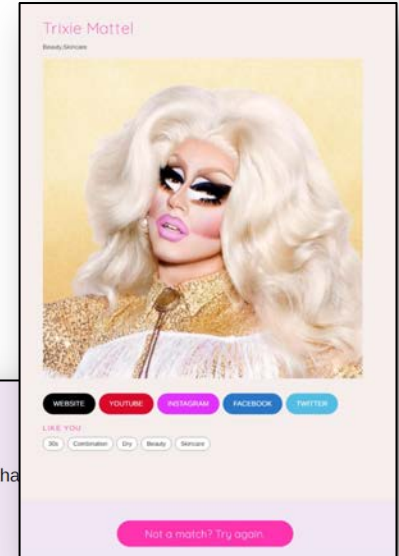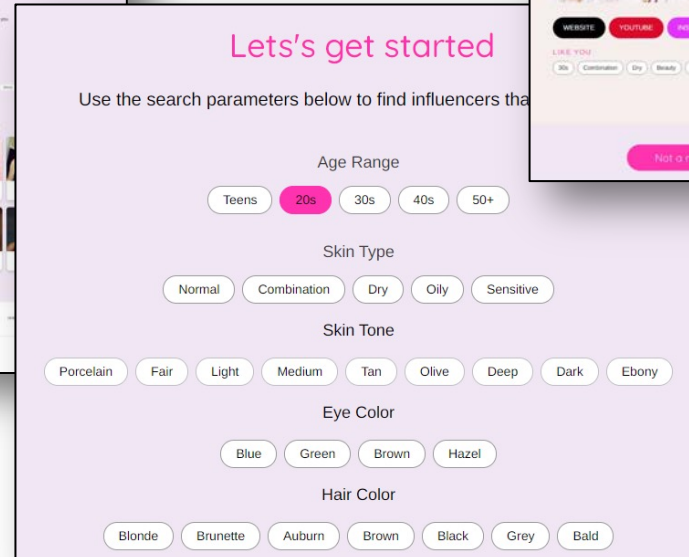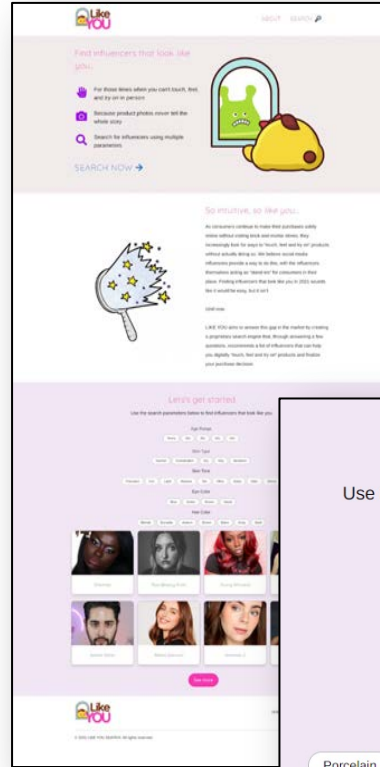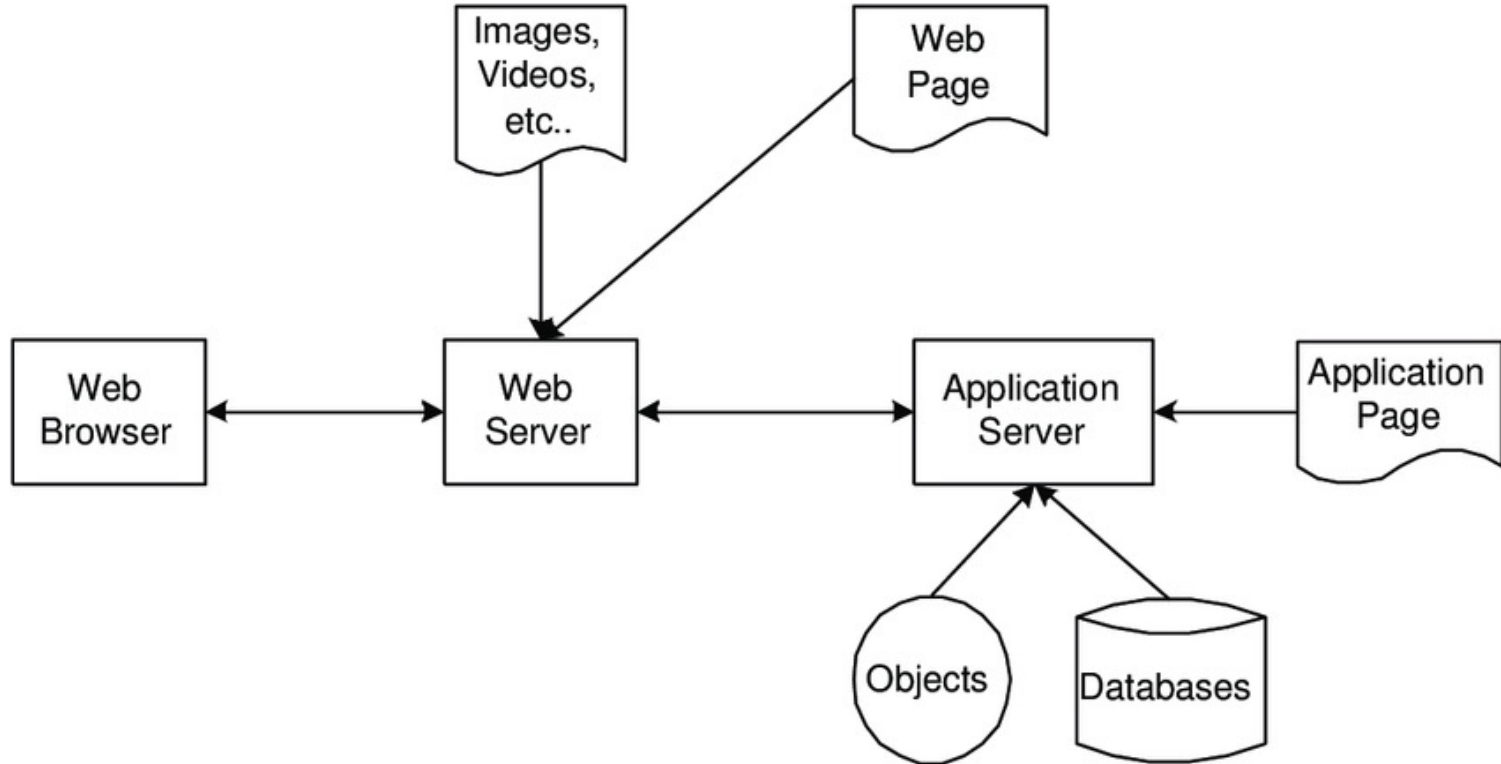
- Secure access

- 24x7 Availability

Design

Intuitive Design

The website is an intuitive design housing:
- intro with call-to-action
- about section
- search/filters with real-time results
- Each influencer result can be expanded for item_detail

Lets's get started

Use the search parameters below to find influencers tha

Age Range

Teens | 20s | 30s | 40s | 50+

Skin Type

Normal | Combination | Dry | Oily | Sensitive

Skin Tone

Porcelain | Fair | Light | Medium | Tan | Olive | Deep | Dark | Ebony

Eye Color

Blue | Green | Brown | Hazel

Hair Color

Blonde | Brunette | Auburn | Brown | Black | Grey | Bald

Architectural View

Work Plan

Task Name	Duration	Start	Finish	% Complete
☐ LIKE YOU Website	94d	12/16/21	04/26/22	100%
☐ Phase 1 Initial Research	22d	12/16/21	01/14/22	100%
Resarch project idea	22d	12/16/21	01/14/22	100%
Competitor analysis	5d	01/10/22	01/14/22	100%
☐ Phase 2 Data Acquisition	21d	01/18/22	02/15/22	100%
Acquire data	21d	01/18/22	02/15/22	100%
☐ Phase 3 Build Website	15d	02/23/22	03/15/22	100%
Design	5d	02/23/22	03/01/22	100%
Implement	5d	03/02/22	03/08/22	100%
Test	5d	03/09/22	03/15/22	100%
☐ Phase 4 - Deliverables	16d	03/16/22	04/06/22	100%
Improvements	5d	03/16/22	03/22/22	100%
Project Report	11d	03/23/22	04/06/22	100%
☐ Phase 5 - Feedback/Correcti	10d	04/13/22	04/26/22	100%
Corrections	5d	04/13/22	04/19/22	100%
Finalized Project Report	5d	04/20/22	04/26/22	100%

Recommendations to Client

- Most important stakeholder is consumer
    - Market ready for consumer-driven influencer search hub
    - Value resides in consumer trust in influencers
    - Majority of consumers prefer following influencers who look and act like them, instead of celebrities.

- Ease of use key MOE
    - As consumers are key stakeholders, adoption relies heavily on ease of use across broad age demographic

- Suggest clients complement existing services with consumer-driven features

Demonstration

Future Work

Future Work

Data
- Add additional social network data for inclusion into Like You
- Expand search parameters to include fit
- Make data dynamic, real-time search

Users
- **Influencers:** Add capability for influencers to add themselves to database, claim profiles, add self-reported data and contact info
- **PR/Brands:** Monetize database with PR/marketing/brand teams by facilitating influencer sponsorship and product releases/campaigns

Future Work

Embedded Application, Brand/POS Focus

- Create embedded application (similar to TrueFit) for brands to embed within POS websites
- After taking a quick questionnaire consumers are shown images of pre-selected brand influencers wearing/using product in suggested size.

Brand POS Website

Questionnaire

Suggested Sizing, Influencer Images

Future Work

Influencer & PR/Brand Monetization
- Build out opportunities for LIKE YOU to connect influencers with PR/Brands
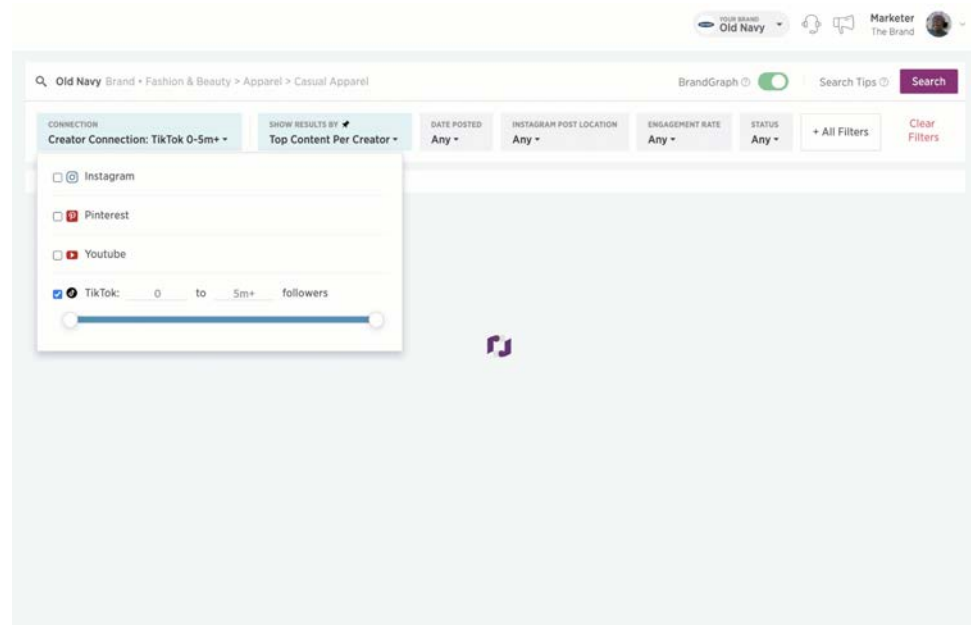
Scholarly References

1. "The Best New Products in Tech." *Product Hunt*, https://www.producthunt.com/.

2. Crudo, Ben. "The 'Return Apocalypse' Is Coming. How Do We Prepare?" *Entrepreneur*, Entrepreneur, 23 Mar. 2021, https://www.entrepreneur.com/article/366997.

3. DMI, Simon @. "20 Surprising Influencer Marketing Statistics: Blog: Online Digital Marketing Courses." *Digital Marketing Institute*, Digital Marketing Institute, 11 Apr. 2022,   https://digitalmarketinginstitute.com/blog/20-influencer-marketing-statistics-that-will-surprise-you.

4. Fit, True. "Global Leader True Fit Raises $30M on 85% Revenue Growth." *Global Leader True Fit Raises $30m On 85% Revenue Growth*, 7 Dec. 2021, https://www.prnewswire.com/news-releases/global-leader-true-fit-raises-30m-on-85-revenue-growth-301437605.html.

5. "Influencer Marketing Survey Results: 2019 Industry Benchmarks." *Mediakix*, https://mediakix.com/influencer-marketing-resources/influencer-marketing-industry-statistics-survey-benchmarks/.

6. "Influencer Search." *IZEA*, 18 Feb. 2022, https://izea.com/influencer-search/.

7. Meena, Satish, et al. "ECommerce Will Account for 36% of Global Fashion Retail Sales by 2022." *Forrester*, 16 Nov. 2018, https://www.forrester.com/blogs/ecommerce-will-account-for-36-of-global-fashion-retail-sales-by-2022/.

8. Santora, Jacinda. "Key Influencer Marketing Statistics You Need to Know for 2022." *Influencer Marketing Hub*, 29 Mar. 2022, https://influencermarketinghub.com/influencer-marketing-statistics/.

9. Scrapy. "Scrapy/Scrapy: Scrapy, a Fast High-Level Web Crawling & Scraping Framework for Python." *GitHub*, https://github.com/scrapy/scrapy.

10. "True Fit for Retailers." *True Fit for Retailers*, https://www.truefit.com/business.

11. Whittaker, Zack. "Web Scraping Is Legal, US Appeals Court Reaffirms." *TechCrunch*, TechCrunch, 18 Apr. 2022, https://techcrunch.com/2022/04/18/web-scraping-legal-court/.

12. Wong, Brittany. "Make This the Year You Follow Influencers Who Actually Look like You." *HuffPost*, HuffPost, 23 Aug. 2021, https://www.huffpost.com/entry/instagram-influencers-who-look-like-you_l_5fff2c4ac5b6c77d85ec470c.
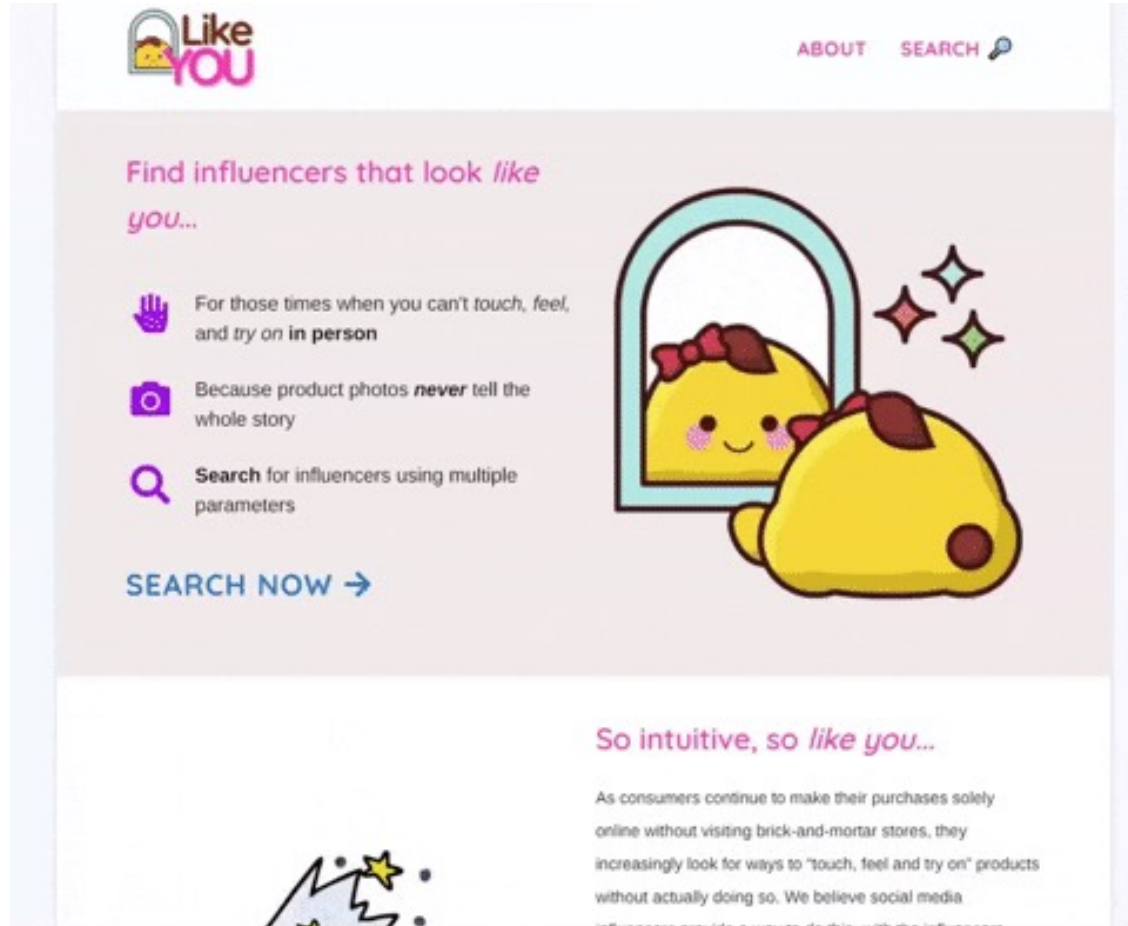
Appendix

Design

Frontend

Softr was used because:

- easy backend integration with Airtable via RESTful API calls
- responsive design

Design

Backend

Airtable was chosen because:

- easy database management
- multi-user capabilities
- easy integration with other applications via RESTful API calls