



Project 2: Prototype IR System

By

Mr.	Suppakorn	Pojsomphong	6588077
Miss	Nimmida	Maneewan	6588181
Miss	Nonraphan	Tatipanithee	6588196

**Report Submitted in Partial Fulfillment of
the Requirements for**

ITCS414 Information Retrieval and Storage

**Faculty of Information and Communication Technology
Mahidol University
2024**

Table of contents

Introduction	1
Problem We Are Trying to Solve	2
Existing relevant systems	3-4
Implementation	5-8
- Data Collection and Data Statistics	
- Example of Documents	
- Tools and Softwares	
- System Diagram	
- Screenshots of the System	
- Example Step-by-Step Search Sessions	
Discussion	
- Limitations of the System	
- Technical Difficulties and Challenges, and How to Handle	
- Lessons Learned	
- Opportunities for Future Improvements	
Conclusion	

Introduction

This project is created as a part of ITCS414 Information Storage and Retrieval which represents the search engine of Sanrio characters. We have implemented the details of each character including image, name, debut date and relevant information in the search system.

As we all know Sanrio has been popular among a wide range of ages from the past to present. There have not been any websites that can provide clear information, complete character lists, and a user-friendly interface experience. Furthermore, casual fans alike often struggle to locate comprehensive and accurate resources about their favorite Sanrio figures. This project seeks to address this gap by developing a dedicated search engine tailored specifically to Sanrio characters.

The primary goal of this project is to implement a search engine that delivers a seamless and efficient way for users to discover and explore Sanrio characters. The search engine will allow users to search by **character name, description, debut year, creator name, species, and gender**, ensuring they can quickly locate the information they need. In addition, we aim to offer a user-friendly interface, enhanced with high-quality images and intuitive navigation, that caters to both ardent and casual fans.

In conclusion, our project not only fulfills the requirements of the ITCS414 Information Storage and Retrieval course but also aims to offer a robust and user-friendly platform. Ensuring fans have easy access to accurate and complete information about their favorite Sanrio characters. We believe this will enrich the overall fan experience and help foster a deeper connection with the Sanrio universe.

We believe that this project, while initially conceptualized as an academic assignment, holds the potential to evolve into a valuable tool for the Sanrio community. It reflects our commitment to enhancing the fan experience and our passion for technology and innovation.

Problem We Are Trying to Solve

1. **Scattered Information Across Multiple Websites:** Currently, users need to search for Sanrio character information across various websites. Therefore, this makes it difficult to find all the information they need in one place.
2. **Incomplete and Inconsistent Character Data:** Many websites do not provide users with complete and updated information for all Sanrio characters. So, this causes users to often find that character profiles are missing details and are incomplete.
3. **Limited Search Functionality:** Existing search systems on platforms related to Sanrio characters often lack advanced search capabilities:
 - **Character Traits:** Users who want to search for specific character traits, such as "pink rabbit" or "a character with a red bow," encounter difficulties since general search engines and wikis do not support queries based on multiple characteristics (e.g., color, species, or accessories).
 - **Partial Names and Keywords:** If users can remember only part of a character's name, the current search systems in the websites may fail to return accurate results. For example, searching for "Kit" might not retrieve "Hello Kitty" unless the full name is entered.
 - **Advanced Queries:** The current systems, which are on the website for finding information about Sanrio characters, do not accommodate more complex searches, such as creation date or creator. For instance, a user who wishes to search for characters that are created in the 1980s or are designed by specific creators would find this type of query/ search impossible on existing platforms.
 - **Ranking and Relevance:** Even when results are returned, they are not usually ranked by how relevant they are. This means users might have to scroll through unrelated characters or results that only weakly match the search term.
4. **No Filtering or Ranking:** There are no features to filter or rank search results based on what the user prefers, making it harder to find characters quickly and easily.

Existing relevant systems

Hello Kitty Wiki offers a list of Sanrio characters, yet it does not provide a comprehensive and up-to-date list of all characters. Additionally, the search results provided do not fully meet the preference and expectations of users and fans. This makes it difficult for users to easily find detailed information about lesser-known or newer characters, resulting in a less satisfying user experience.

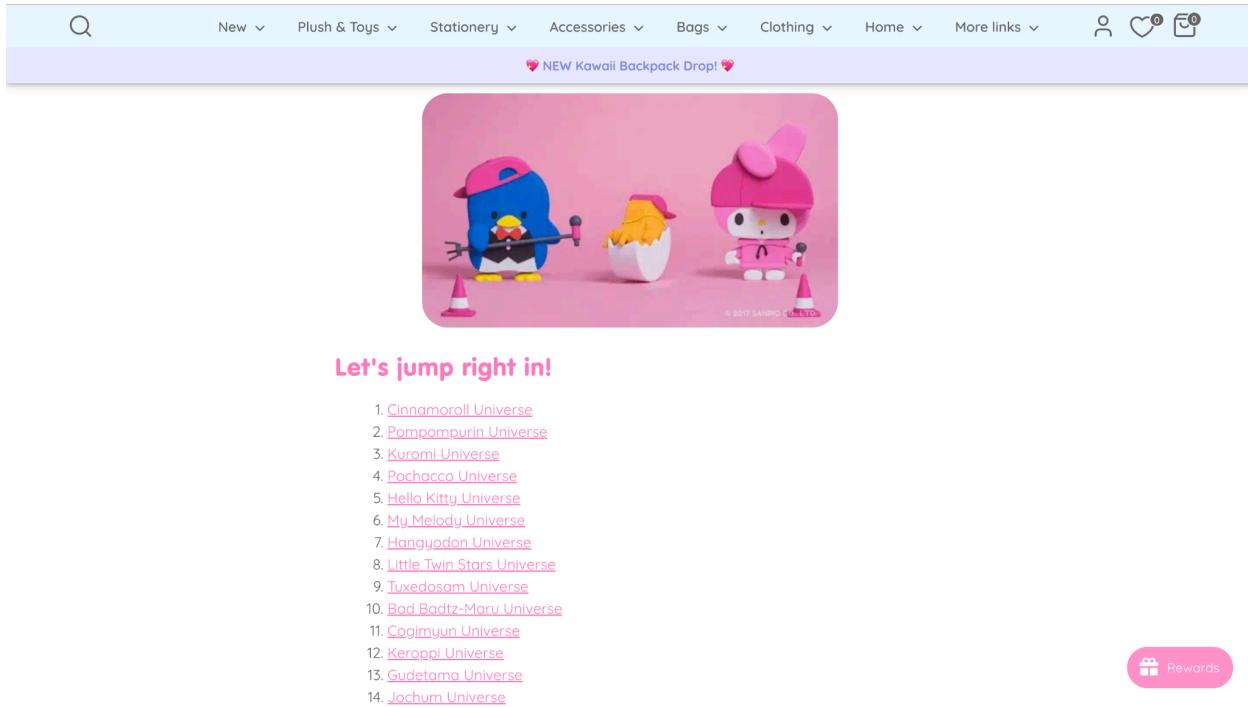
Website Link: https://hellokitty.fandom.com/wiki/List_of_Sanrio_characters

The screenshot shows the Hello Kitty Wiki homepage with a pink background featuring Hello Kitty characters. At the top right are 'SIGN IN' and 'REGISTER' buttons, and at the top center is a search bar showing '1,236 PAGES'. The main content area is titled 'Search Results' and displays a message: 'There is a page named "Kuromi" on Hello Kitty Wiki. See also the other search results found.' Below this is a search bar with 'This wiki ▾ kuromi' and a magnifying glass icon. To the right of the search bar is an 'Advanced ▾' button. The search results are categorized by type: 'Articles', 'Posts', 'Photos and Videos', 'Blogs', 'People', and 'Everything'. The 'Articles' section shows a result for 'Kuromi' with a link to the page. The 'Photos and Videos' section includes an advertisement for eBay. On the right side of the search results, there is a vertical sidebar with various links like 'FAN CENTRAL', 'GAMES', 'ANIME', 'MOVIES', 'TV', 'VIDEO', and 'WIKIS'. At the bottom right of the search results area is a promotional banner for XM, featuring two bulls and the text 'พร้อมรับโบนัสเงินฝาก สูงสุด \$10,500'.

Blippo is a website offering a comprehensive list of Sanrio characters, including images, names, and descriptions. The website is user-friendly, allowing users to browse characters by universe, making it easier to find a character if they know which universe they belong to.

Website Link:

<https://www.blippo.com/blogs/characters/sanrio-characters?srsltid=AfmBOor4vN8EOxv-KEuQYbo2uVY61V778h3ivXm4LOQO5VFenRpZA4xP>



Implementation

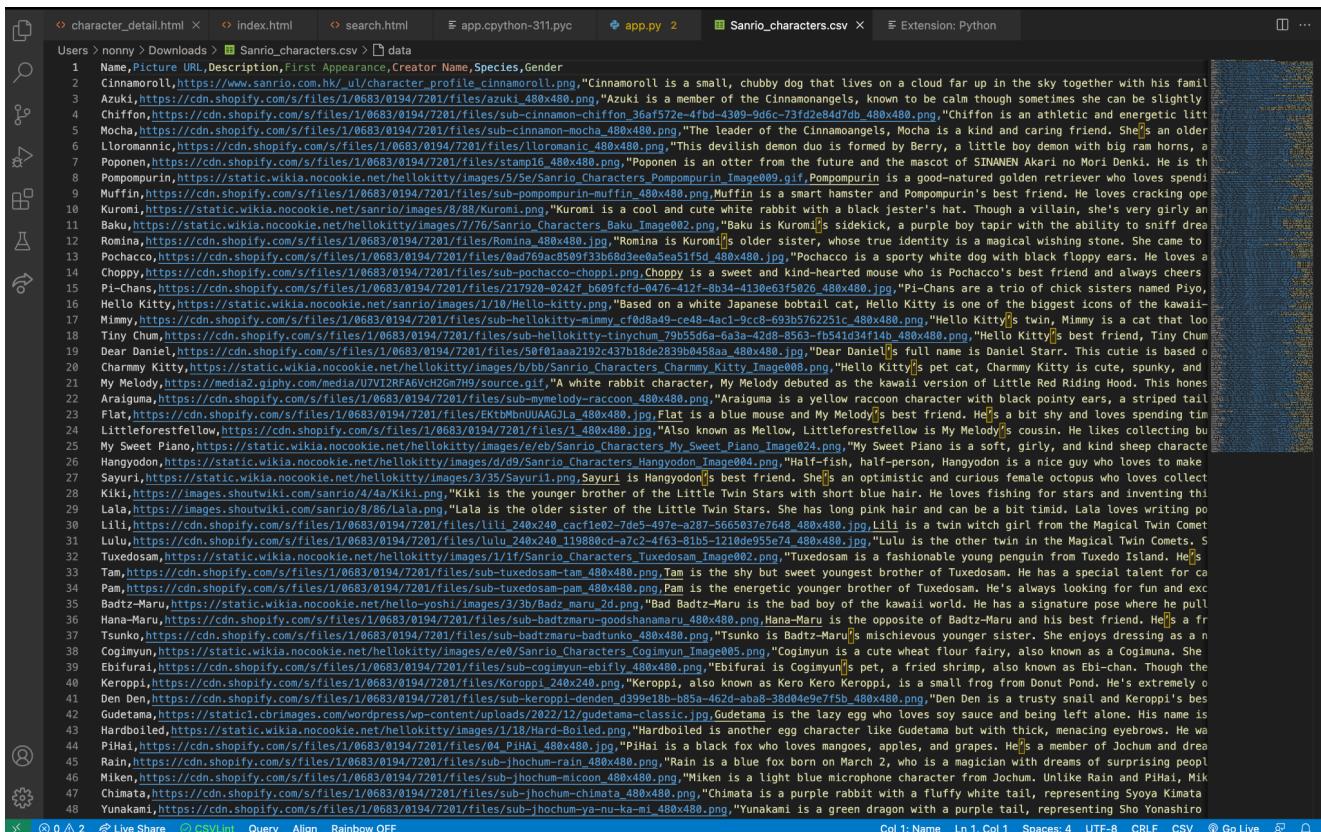
Data collection and data statistics

The information of every character is retrieved from the website “Hello Kitty Wiki” and “Blippo” which consists of 223 characters’ information which include name, image, description, debut year, creator name, species, and gender. The data is collected in the CSV format then converts the information into JSON format by using python code to normalize the data.

Link of the website: https://hellokitty.fandom.com/wiki/List_of_Sanrio_characters,
<https://www.blippo.com/blogs/characters/sanrio-characters?srslid=AfmBOor4vN8Eoxv-KEuQYbo2uVY61V778h3ivXm4LOQO5VFenRpZA4xP>

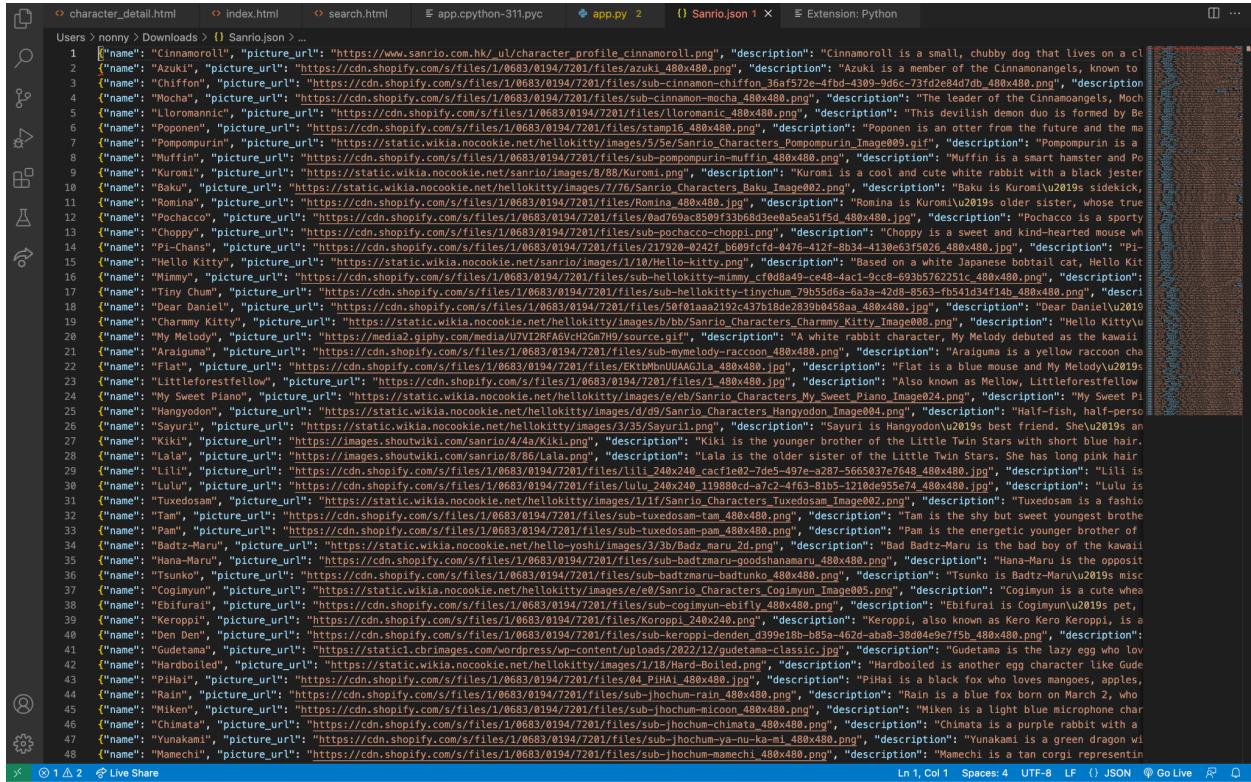
Example of Documents

CSV format



```
Name,Picture,URL,Description,First Appearance,Creator Name,Species,Gender
1,Cinnamoroll,https://www.sanrio.com.hk/\_/l/character\_profile\_cinnamoroll.png,"Cinnamoroll is a small, chubby dog that lives on a cloud far up in the sky together with his family.",2011,Naoko Ito,Cinnamon,Genderless
2,Azuki,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/azuki\_480x480.png,"Azuki is a member of the Cinnamangels, known to be calm though sometimes she can be slightly grumpy.",2011,Naoko Ito,Cinnamon,Genderless
3,Chiffon,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-cinnamon-chiffon\_36af572e-4fbd-4309-96dc-73fd2e847db\_480x480.png,"Chiffon is an athletic and energetic little girl who loves to play soccer.",2011,Naoko Ito,Cinnamon,Genderless
4,Mocha,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-cinnamon-mocha\_480x480.png,"The leader of the Cinnamangels, Mocha is a kind and caring friend. She's an older sister to Azuki and Cinnamoroll.",2011,Naoko Ito,Cinnamon,Genderless
5,Llormannic,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/llormannic\_480x480.png,"This devilish demon duo is formed by Berry, a little boy demon with big ram horns, and Popopon, a cool and cute white rabbit with a black jester's hat. Though a villain, she's very girly and loves spending time with Berry.",2011,Naoko Ito,Demon,Genderless
6,Popponen,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/stamp16\_480x480.png,"Popponen is an otter from the future and the mascot of SINANEN Akari no Mori Denki. He is the best friend of Berry.",2011,Naoko Ito,Otter,Genderless
7,Pompompurin,https://static.wikia.nocookie.net/hellokitty/images/5/5e/Pompompurin\_an\_itteyem009.gif,"Pompompurin is a good-natured golden retriever who loves spending time with Berry and Popponen.",2011,Naoko Ito,Dog,Genderless
8,Muffin,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-pompompurin-muffin\_480x480.png,"Muffin is a smart hamster and Pompompurin's best friend. He loves cracking open shells and eating them.",2011,Naoko Ito,Hamster,Genderless
9,Kuromi,https://static.wikia.nocookie.net/sanrio/images/8/88/Kuromi.png,"Kuromi is a cool and cute white rabbit with a black jester's hat. Though a villain, she's very girly and loves spending time with Berry and Popponen.",2011,Naoko Ito,Rabbit,Genderless
10,Baku,https://static.wikia.nocookie.net/hellokitty/images/7/76/Sanrio\_Characters\_Baku\_Image002.png,"Baku is Kuromi's older sister, whose true identity is a magical wishing stone. She came to Earth from another dimension to help Berry and Popponen.",2011,Naoko Ito,Wishing Stone,Genderless
11,Romina,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/romina\_480x480.jpg,"Romina is a cool and cute white rabbit with a black jester's hat. Though a villain, she's very girly and loves spending time with Berry and Popponen.",2011,Naoko Ito,Rabbit,Genderless
12,Pochacco,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/8ad769ac8509f3b68a3ee0a5ea51f5d\_480x480.jpg,"Pochacco is a sporty white dog with black floppy ears. He loves a good game of fetch and loves to play with Berry and Popponen.",2011,Naoko Ito,Dog,Genderless
13,Choppo,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-pochacco-choppo.png,"Choppo is a sweet and shy mouse who is Pochacco's best friend and always cheers him on during games.",2011,Naoko Ito,Mouse,Genderless
14,Pi-Chans,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/217920-0242f\_b609fffd-0476-412f-8b34-413be63f5026\_480x480.jpg,"Pi-Chans are a trio of chick sisters named Piyo, Hello Kitty, and Mimmy.",2011,Naoko Ito,Chick,Genderless
15,Hello Kitty,https://static.wikia.nocookie.net/sanrio/images/1/10>Hello\_kitty.png,"Based on a white Japanese bobtail cat, Hello Kitty is one of the biggest icons of the kawaii culture.",2011,Naoko Ito,Cat,Genderless
16,Mimmy,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-hellokitty-mimmy\_cfdd8a49-c4e8-4ac1-9ec8-693b5762251c\_480x480.png,"Hello Kitty's twin, Mimmy is a cat that looks just like her but is a bit more laid back and loves to sleep in.",2011,Naoko Ito,Cat,Genderless
17,Tiny Chum,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-hellokitty-tinychum\_7955d6a-6a3a-42d8-8563-f6541d34f14b\_480x480.png,"Hello Kitty's best friend, Tiny Chum is a tiny hamster who loves to play with Berry and Popponen.",2011,Naoko Ito,Hamster,Genderless
18,Dear Daniel,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/5f0f01aa2192c437b2de2839b0458aa\_480x480.jpg,"Dear Daniel's full name is Daniel Starr. This cutie is based on the character of the same name from the movie 'The Polar Express'. He is a half-fish, half-person, Hangyodon is a nice guy who loves to make things.",2011,Naoko Ito,Fish,Genderless
19,Charmy Kitty,https://static.wikia.nocookie.net/hellokitty/images/6/6b/Sanrio\_Characters\_Charmy\_Kitty\_Image008.png,"Hello Kitty's pet cat, Charmy Kitty is cute, spunky, and loves to play with Berry and Popponen.",2011,Naoko Ito,Cat,Genderless
20,My Melody,https://media2.giphy.com/media/UTV12RFAWCH2Gm7H9/source.gif,"A white rabbit character, My Melody debuted as the kawaii version of Little Red Riding Hood. This hones her artistic abilities and loves to draw.",2011,Naoko Ito,Rabbit,Genderless
21,Araiguma,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-mymelody-raccoon\_480x480.png,"Araiguma is a yellow raccoon character with black pointy ears, a striped tail and loves to collect shiny objects.",2011,Naoko Ito,Raccoon,Genderless
22,Flat,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/EKTHbnUAG1a\_480x480.jpg,"Flat is a blue mouse and My Melody's best friend. He's a bit shy and loves spending time with her.",2011,Naoko Ito,Mouse,Genderless
23,Littleforestfellow,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/1\_480x480.jpg,"Also known as Mellow, Littleforestfellow is My Melody's cousin. He likes collecting bugs and loves to play with Berry and Popponen.",2011,Naoko Ito,Insect,Genderless
24,My Sweet Piano,https://static.wikia.nocookie.net/hellokitty/images/e/eb/Sanrio\_Characters\_My\_Sweet\_Piano\_Image024.png,"My Sweet Piano is a soft, girly, and kind sheep character who loves to play the piano and loves to play with Berry and Popponen.",2011,Naoko Ito,Sheep,Genderless
25,Handyodon,https://static.wikia.nocookie.net/hellokitty/images/d/d9/Sanrio\_Characters\_Handyodon\_Image004.png,"Half-fish, half-person, Hangyodon is a nice guy who loves to make things and loves to play with Berry and Popponen.",2011,Naoko Ito,Fish,Genderless
26,Sayuri,https://static.wikia.nocookie.net/hellokitty/images/3/35/Sayuri.png,"Sayuri is Hangyodon's best friend. She's an optimistic and curious female octopus who loves to collect shiny objects.",2011,Naoko Ito,Octopus,Genderless
27,Kiki,https://images.shoutwiki.com/sanrio/4/4/Kiki.png,"Kiki is the younger brother of the Little Twin Stars with short blue hair. He loves fishing for stars and inventing things.",2011,Naoko Ito,Elf,Genderless
28,Lala,https://images.shoutwiki.com/sanrio/8/86/Lala.png,"Lala is the older sister of the Little Twin Stars. She has long pink hair and can be a bit timid. Lala loves writing poems and loves to collect shiny objects.",2011,Naoko Ito,Elf,Genderless
29,Lili,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/lili\_240x240\_cacf1e-2d65-497e-a287-5665037e7648\_480x480.jpg,"Lili is a twin with girl from the Magical Twin Comet who loves to collect shiny objects and loves to play with Berry and Popponen.",2011,Naoko Ito,Elf,Genderless
30,Lulu,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/lulu\_240x240\_l1988cd-7e2-4f63-81b5-1210d955e74\_480x480.jpg,"Lulu is the other twin in the Magical Twin Comets. She loves to collect shiny objects and loves to play with Berry and Popponen.",2011,Naoko Ito,Elf,Genderless
31,Tuxedosam,https://static.wikia.nocookie.net/hellokitty/images/1/f1/Sanrio\_Characters\_Tuxedosam\_Image002.png,"Tuxedosam is a fashionable young penguin from Tuxedo Island. He's a bit of a ladies' man and loves to collect shiny objects.",2011,Naoko Ito,Penguin,Genderless
32,Tam,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-tuxedosam-tam\_480x480.png,"Tam is the shy but sweet youngest brother of Tuxedosam. He has a special talent for making things and loves to play with Berry and Popponen.",2011,Naoko Ito,Penguin,Genderless
33,Pam,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-tuxedosam-pam\_480x480.png,"Pam is the energetic younger brother of Tuxedosam. He's always looking for fun and exciting things to do.",2011,Naoko Ito,Penguin,Genderless
34,Badtz-Maru,https://static.wikia.nocookie.net/hellokitty/images/3/b/Badz\_maru\_2d.png,"Badtz-Maru is the bad boy of the kawaii world. He has a signature pose where he pulls his hair and makes a face.",2011,Naoko Ito,Dragon,Genderless
35,Hana-Maru,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-badtzmaru-goodshamaru\_480x480.png,"Hana-Maru is the opposite of Badtz-Maru and his best friend. He's a friend of the magical twin comet.",2011,Naoko Ito,Dragon,Genderless
36,Tsunko,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-badtzmaru-badtunk\_480x480.png,"Tsunko is Badtz-Maru's mischievous younger sister. She enjoys dressing up as a nai nai and loves to play with Berry and Popponen.",2011,Naoko Ito,Dragon,Genderless
37,Cogimyun,https://static.wikia.nocookie.net/hellokitty/images/e/eb/Sanrio\_Characters\_Cogimyun\_Image005.png,"Cogimyun is a cute wheat flour fairy, also known as a Cogimuna. She loves to collect shiny objects and loves to play with Berry and Popponen.",2011,Naoko Ito,Fairy,Genderless
38,Ebifurui,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-cogimyun-ebifuri\_480x480.png,"Ebifurui is Cogimyun's pet, a fried shrimp, also known as Ebi-chan. Though he loves to collect shiny objects and loves to play with Berry and Popponen.",2011,Naoko Ito,Shrimp,Genderless
39,Keroppi,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/koroppi\_240x240.png,"Keroppi, also known as Kero Kero Keroppi, is a small frog from Donut Pond. He's extremely cute and loves to play with Berry and Popponen.",2011,Naoko Ito,Frog,Genderless
40,Den Den,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-keroppi-denden\_j399e18b-8b5a-462d-ab8-38d04e9e775b\_480x480.png,"Den Den is a trusty snail and Keroppi's best friend.",2011,Naoko Ito,Snail,Genderless
41,Gudetama,https://static1.cbrimages.com/wp-content/uploads/2022/12/gudetama-classic.jpg,"Gudetama is the lazy egg who loves soy sauce and being left alone. His name is Gudetama.",2011,Naoko Ito,Egg,Genderless
42,Hardboiled,https://static.wikia.nocookie.net/hellokitty/images/1/18/Hard\_Boiled.png,"Hardboiled is another egg character like Gudetama but with thick, menacing eyebrows. He loves to collect shiny objects and loves to play with Berry and Popponen.",2011,Naoko Ito,Egg,Genderless
43,PiHai,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/04\_PiHai\_480x480.jpg,"PiHai is a black fox who loves mangoes, apples, and grapes. He's a member of Jochum and loves to collect shiny objects.",2011,Naoko Ito,Fox,Genderless
44,Rain,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-rain\_480x480.png,"Rain is a blue fox born on March 2, who is a magician with dreams of surprising people.",2011,Naoko Ito,Fox,Genderless
45,Miken,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-miccon\_480x480.png,"Miken is a light blue microphone character from Jochum. Unlike Rain and PiHai, Miken loves to collect shiny objects and loves to play with Berry and Popponen.",2011,Naoko Ito,Mic,Genderless
46,Chimata,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-chimata\_480x480.png,"Chimata is a purple rabbit with a fluffy white tail, representing Soya Kimata.",2011,Naoko Ito,Rabbit,Genderless
47,Yunakami,https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-yu-na-ka-mi\_480x480.png,"Yunakami is a green dragon with a purple tail, representing Sho Yonashiro.",2011,Naoko Ito,Dragon,Genderless
```

JSON format



```
character_detail.html index.html search.html app.py 2 Sanrio.json 1 Extension: Python
Users > nonny > Downloads > Sanrio.json > ...
1 {"name": "Cinnamoroll", "picture_url": "https://www.sanrio.com.hk/_ul/character_profile_cinnamoroll.png", "description": "Cinnamoroll is a small, chubby dog that lives on a cloud. He is a member of the Cinnamangals, known to be very friendly and kind-hearted."}
2 {"name": "Azuki", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/azuki_480x480.png", "description": "Azuki is a member of the Cinnamangals, known to be very friendly and kind-hearted."}
3 {"name": "Chifon", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-cinnamon-chifon_36af572e-4fbda-3309-9d6c-73fd2e0ad7db_480x480.png", "description": "The leader of the Cinnamangals, Chifon is a wise and experienced character who loves to eat."}
4 {"name": "Mocha", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-cinnamon-mocha_480x480.png", "description": "The leader of the Cinnamangals, Mocha is a wise and experienced character who loves to eat."}
5 {"name": "Lloromannic", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/lloromannic_480x480.png", "description": "This devilish demon duo is formed by Besty and Popopon."}
6 {"name": "Popopon", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/stamp16_480x480.png", "description": "Popopon is an otter from the future and the main character of the Lloromannic duo."}
7 {"name": "Pompompurin", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/5/5e/Sanrio_Characters_Pompompurin_Image009.gif", "description": "Pompompurin is a white cat with a pom-pom tail and a bow tie."}
8 {"name": "Muffin", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-pompompurin-muffin_480x480.png", "description": "Muffin is a smart hamster and Po's best friend."}
9 {"name": "Kuromi", "picture_url": "https://static.wikia.nocookie.net/sanrio/images/8/89/Kuromi.png", "description": "Kuromi is a cool and cute white rabbit with a black jester hat and a bow tie."}
10 {"name": "Baku", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/7/76/Sanrio_Characters_Baku_Image002.png", "description": "Baku is Kuromi's 2019s sidekick."}
11 {"name": "Romina", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/Romina_480x480.jpg", "description": "Romina is Kuromi's older sister, whose true name is Romina."}
12 {"name": "Pochacco", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-pochacco-choppo.png", "description": "Pochacco is a sporty mouse who loves to play soccer."}
13 {"name": "Choppo", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-pochacco-choppo.png", "description": "Choppo is a sweet and kind-hearted mouse who loves to play soccer."}
14 {"name": "Pi-Chans", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/stamp16_480x480.jpg", "description": "Pi-Chans is a yellow pomeranian dog who loves to play soccer."}
15 {"name": "Hello Kitty", "picture_url": "https://static.wikia.nocookie.net/sanrio/images/1/10>Hello-kitty.png", "description": "Based on a white Japanese bobtail cat, Hello Kitty is a popular character."}
16 {"name": "Mimmy", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-helloskitty-mimmy_cf08da49-cea8-4a1-9c8-693b762251c_480x480.png", "description": "Mimmy is a small white cat with a bow tie."}
17 {"name": "Tiny Chum", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-helloskitty-tinychu_79b55d6a-6a3a-4208-8563-fb541d34f14b_480x480.png", "description": "Tiny Chum is a tiny white rabbit with a bow tie."}
18 {"name": "Dear Daniel", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/50f01aa2192c437618de2839b0458a_480x480.jpg", "description": "Dear Daniel is a 2019s character who loves to play soccer."}
19 {"name": "Charmy Kitty", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/b/bb/Sanrio_Characters_Charmy_Kitty_Image008.png", "description": "Charmy Kitty is a small white butterfly who loves to play soccer."}
20 {"name": "My Melody", "picture_url": "https://media2.giphy.com/media/UV12RFaGVCHGmPH/source.gif", "description": "My Melody debuted as the kawaii character in 1975."}
21 {"name": "Araguma", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-my melody-raccoon_480x480.png", "description": "Araguma is a yellow raccoon character."}
22 {"name": "Flat", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/EKtMBmuUAAGLa_480x480.jpg", "description": "Flat is a blue mouse and My Melody's 2019s sidekick."}
23 {"name": "Littleforestfellow", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/217920-0242f_b689fcfd-0476-412f-9b34-4138e63f5026_480x480.jpg", "description": "Also known as Mellow, Littleforestfellow is a 2019s character."}
24 {"name": "My Sweet Piano", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/e/eb/Sanrio_Characters_My_Sweet_Piano_Image024.png", "description": "My Sweet Piano is a 2019s character."}
25 {"name": "Hangyodon", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/d/d9/Sanrio_Characters_Hangyodon_Image004.png", "description": "Hangyodon is a 2019s character."}
26 {"name": "Sayuri", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/3/35/Sayuri1.png", "description": "Sayuri is Hangyodon's 2019s best friend. She is a 2019s character."}
27 {"name": "Kiki", "picture_url": "https://images.shoutwiki.com/sanrio/4/4a/Kiki.png", "description": "Kiki is the younger brother of the Little Twin Stars with short blue hair."}
28 {"name": "Lala", "picture_url": "https://shoutwiki.com/sanrio/8/86/Lala.png", "description": "Lala is the older sister of the Little Twin Stars. She has long pink hair."}
29 {"name": "Lili", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/lili_240x240_carfe02-7de5-497e-a287-5665037e7648_480x480.jpg", "description": "Lili is a 2019s character."}
30 {"name": "Lulu", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/lulu_240x240_119880cd-a7c2-4f63-81b5-1210de55e74_480x480.jpg", "description": "Lulu is a 2019s character."}
31 {"name": "Tuxedosam", "picture_url": "https://static.wikia.nocookie.net/images/1/17/Sanrio_Characters_Tuxedosam_Image002.png", "description": "Tuxedosam is a fashio...n character."}
32 {"name": "Tam", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-tuxedosam-tam_480x480.png", "description": "Tam is the shy but sweet youngest brother of the Tuxedosam brothers."}
33 {"name": "Pam", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-tuxedosam-pam_480x480.png", "description": "Pam is the energetic younger brother of the Tuxedosam brothers."}
34 {"name": "Badtz-Maru", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/3/b/Badz_maru_20.png", "description": "Badtz-Maru is the bad boy of the kawaii characters."}
35 {"name": "Hana-Maru", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-badtzmaru-goodshamaru_480x480.png", "description": "Hana-Maru is the opposite of Badtz-Maru."}
36 {"name": "Tsunku", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-badtzmaru-badtunko_480x480.png", "description": "Tsunku is Badtz-Maru's 2019s sidekick."}
37 {"name": "Cogimyun", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/e/e0/Sanrio_Characters_Cogimyun_Image005.png", "description": "Cogimyun is a cute whea...t character."}
38 {"name": "Ebiturai", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-cogimyun-ebitury_480x480.png", "description": "Ebiturai is Cogimyun's 2019s pet."}
39 {"name": "Keroppi", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/Keroppi_240x240.png", "description": "Keroppi, also known as Kero Kero Keroppi, is a 2019s character."}
40 {"name": "Dee Den", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-keroppi-deend_399e180-85a-462d-ab8-38d04e9e7f5b_480x480.png", "description": "Dee Den is a 2019s character."}
41 {"name": "Gudetama", "picture_url": "https://static.cbrimages.com/wordpress/wp-content/uploads/2022/12/gudetama-classic.jpg", "description": "Gudetama is the lazy egg who loves to sleep."}
42 {"name": "Hardboiled", "picture_url": "https://static.wikia.nocookie.net/helloskitty/images/1/18/Hard-Boiled.png", "description": "Hardboiled is another egg character like Gude...tama."}
43 {"name": "PiHA", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/04_PIHA_480x480.jpg", "description": "PiHA is a black fox who loves mangoes, apples, and watermelons."}
44 {"name": "Rain", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-rain_480x480.png", "description": "Rain is a blue fox born on March 2, who loves rain."}
45 {"name": "Miken", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-micken_480x480.png", "description": "Miken is a light blue microphone character."}
46 {"name": "Chimata", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-chimata_480x480.png", "description": "Chimata is a purple rabbit with a bow tie."}
47 {"name": "Yunakami", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-ya-nu-ka-mi_480x480.png", "description": "Yunakami is a green dragon who loves to play soccer."}
48 {"name": "Mamechi", "picture_url": "https://cdn.shopify.com/s/files/1/0683/0194/7201/files/sub-jhochum-mamechi_480x480.png", "description": "Mamechi is a tan corgi representing the color tan."}
```

Tool and Softwares

1. Python

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. In this project, we have implemented python to create a web server and convert data from csv format to JSON format.

2. Elastic

Elasticsearch is an open source distributed, RESTful search and analytics engine, scalable data store, and vector database capable of addressing a growing number of use cases. In this project, we implement Elasticsearch to indices, inverted index, score the document, and send a method request to retrieve the data based on the query.

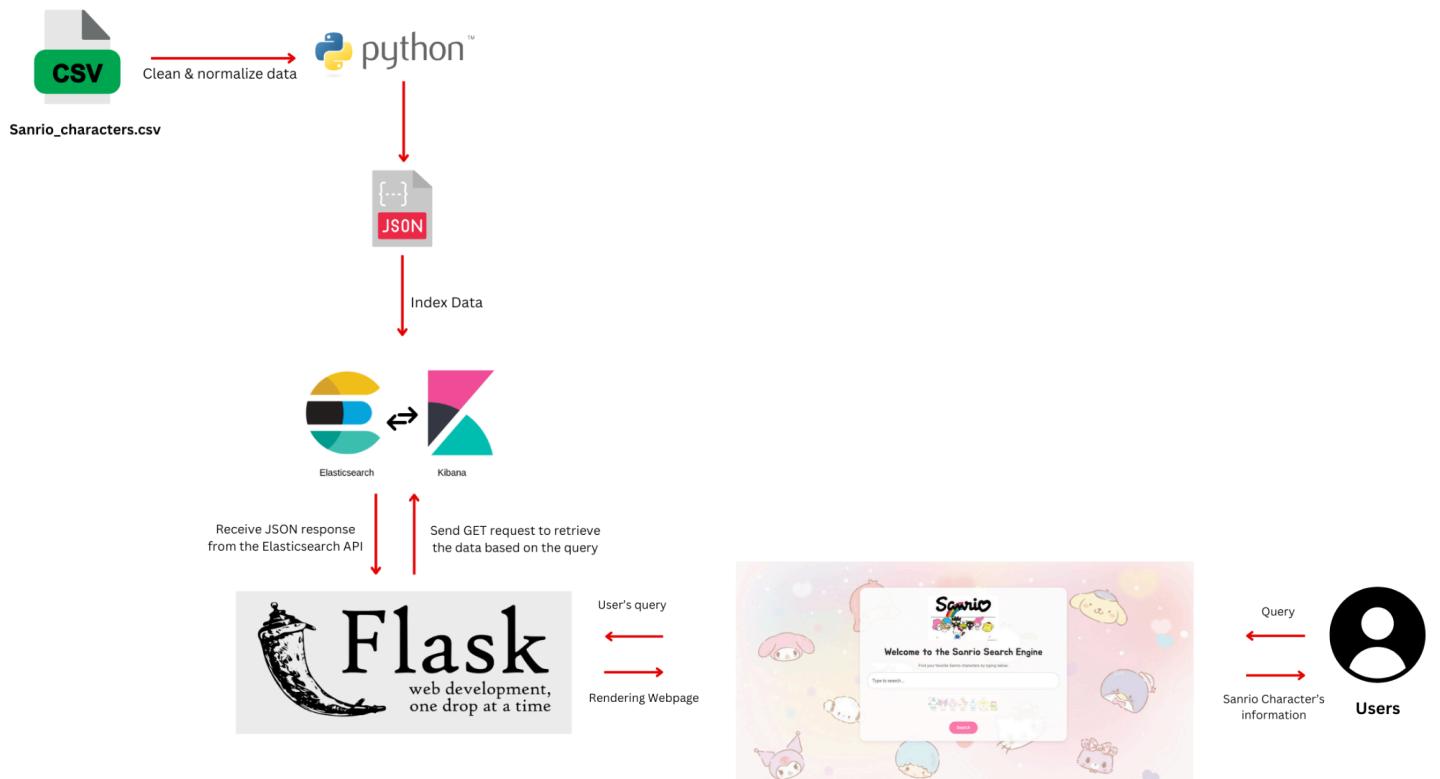
3. Kibana

Kibana is a data visualization and management tool for Elasticsearch that provides real-time histograms, line graphs, pie charts, and maps. It lets you visualize your Elasticsearch data and navigate the Elastic Stack.

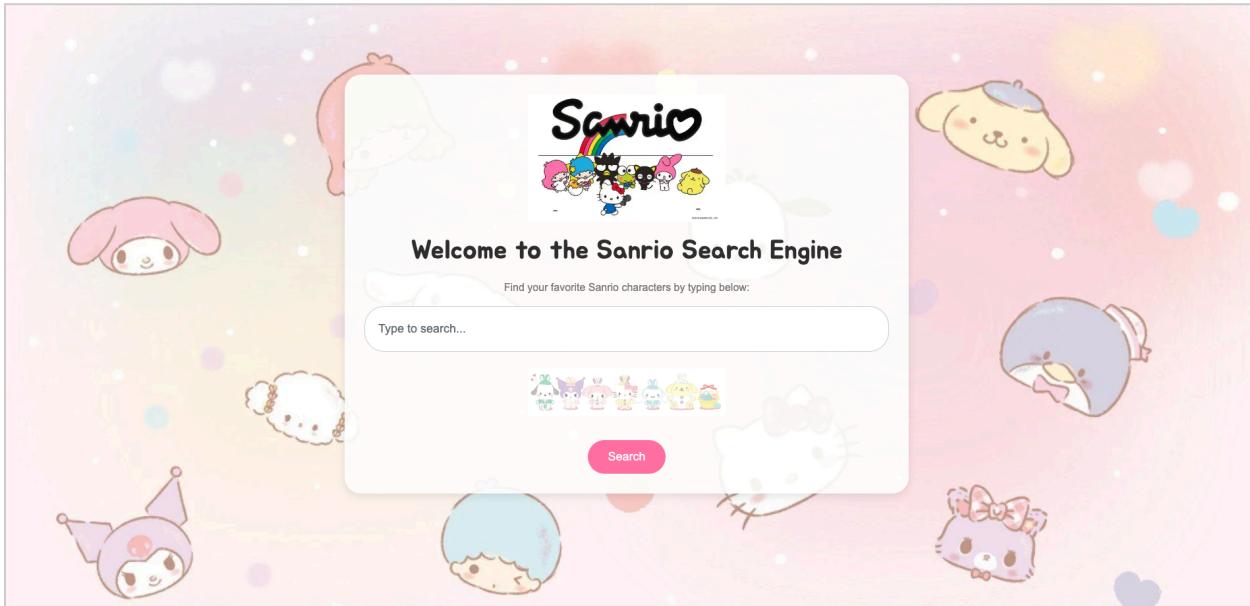
4. Flask

Flask is a lightweight Python web application framework. It was designed to be simple and flexible, allowing developers to easily build web applications. Flask provides the necessary tools for routing, handling requests, and rendering templates, but it leaves a lot of the structure and selection of tools up to the developer. In this project, Flask is implemented in helping to run the web application.

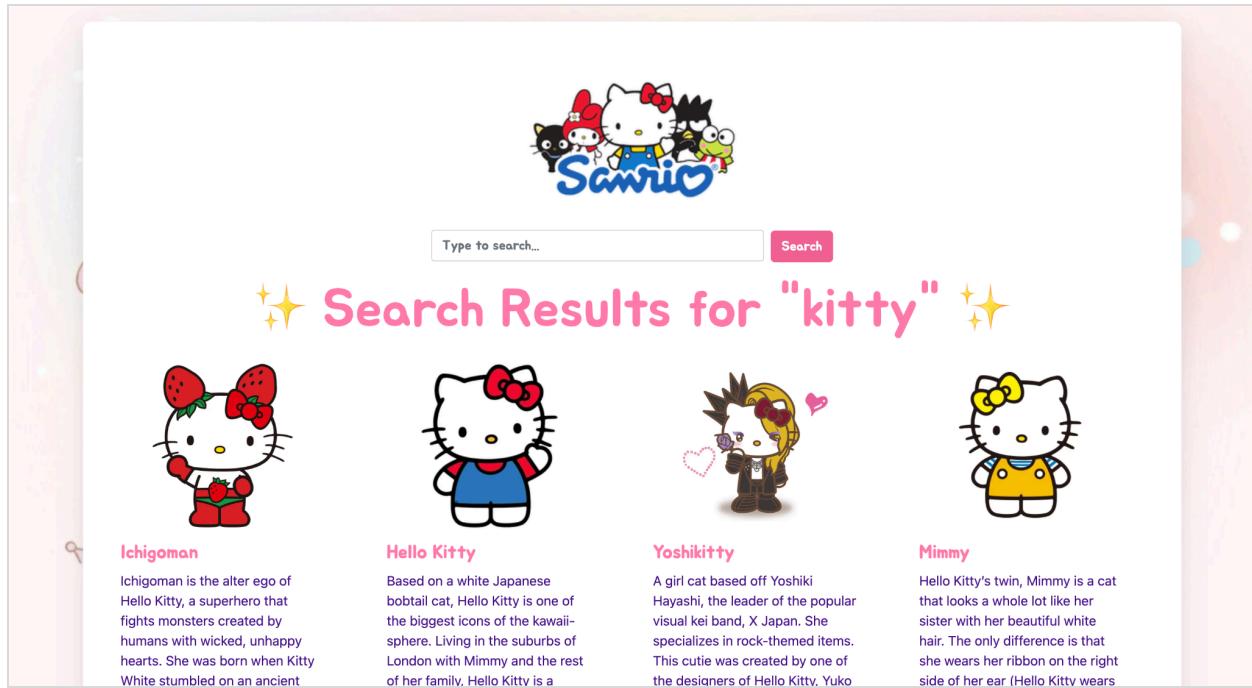
System Diagram



Screenshots of the System



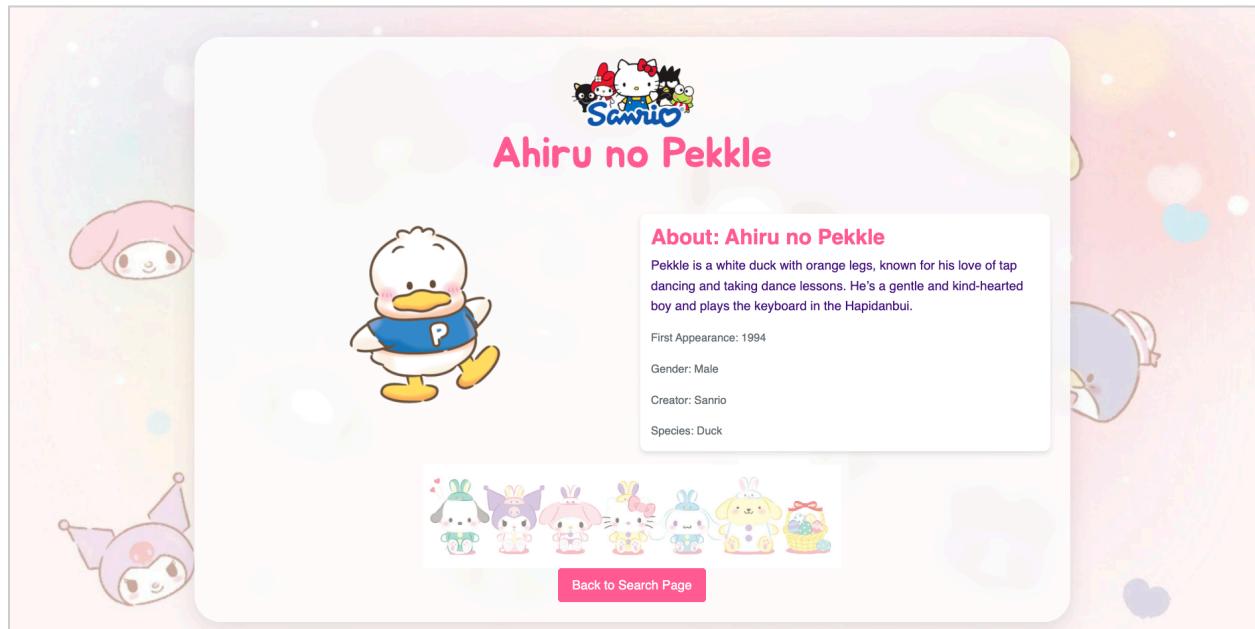
A screenshot of the Sanrio Search Engine search results page. The background is a pink gradient with floating Sanrio characters. At the top, there's a search bar with the placeholder "Type to search..." and a pink "Search" button. Below the search bar, the text "Search Results" is displayed in a large, pink, stylized font with sparkles on either side. Four search results are shown in a grid: 1. "Adorozatorumary" (a blue cat-like character) with a description about being created in collaboration with Japanese singer Ado. 2. "Ahiru no Pekkle" (a white duck with orange legs) with a description about his love of tap dancing. 3. "Alfredaloha" (a brown elephant wearing a hat and shirt) with a description about him being a posh character who likes wearing shirts and hats. 4. "Araiguma" (a yellow raccoon with black ears) with a description about him being a cleanliness lover. Each result includes a small character icon and a detailed text description below it.



The Sanrio website features a search bar at the top with the text "Type to search..." and a "Search" button. Below the search bar, the text "Search Results for 'kitty'" is displayed in a pink, stylized font with starburst icons on either side. Four character cards are shown below the title:

- Ichigoman**: A Hello Kitty character dressed as a superhero, holding a red flower.
- Hello Kitty**: The original white Hello Kitty character wearing a blue dress.
- Yoshikitty**: A girl cat with spiky hair, wearing a black and gold outfit, holding a heart.
- Mimmy**: A white cat twin of Hello Kitty, wearing a yellow dress.

Each character card includes a small image, the character's name, and a brief description.



The Sanrio website displays a character profile for Ahiru no Pekkle. At the top, the Sanrio logo is shown above the text "Ahiru no Pekkle". Below this, there is a large image of Pekkle, a white duck with orange legs, wearing a blue shirt with a 'P' on it. To the right of the image is a box containing the following information:

About: Ahiru no Pekkle

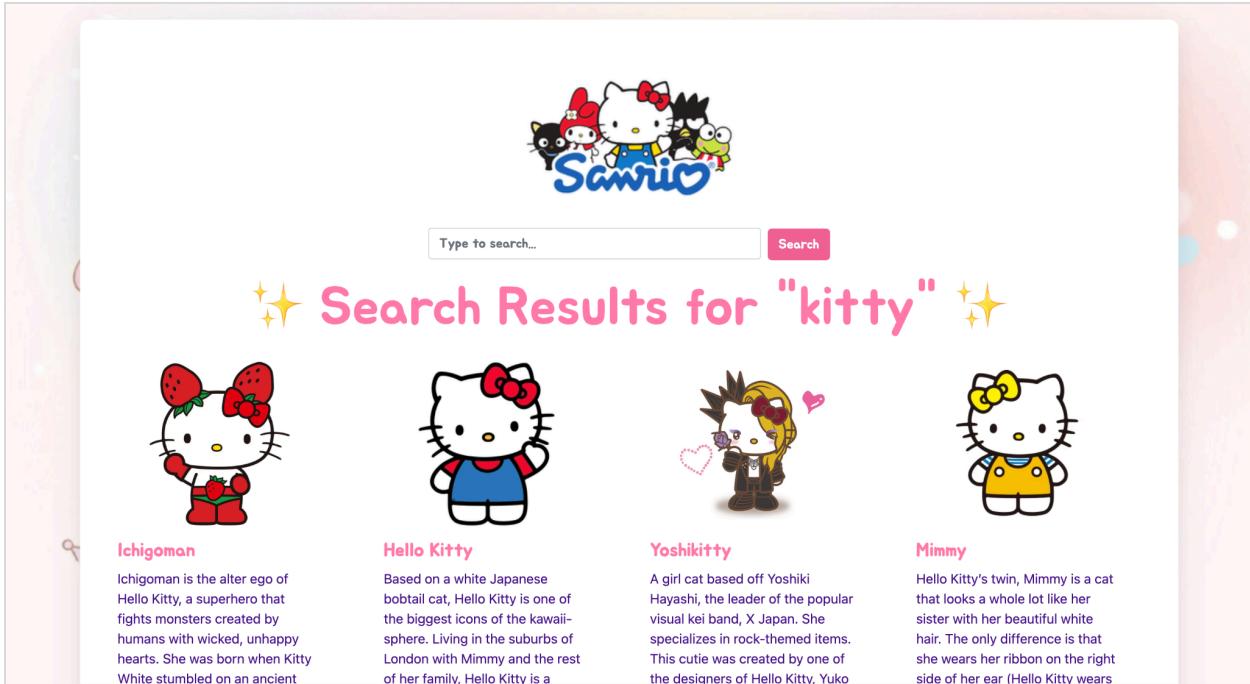
Pekkle is a white duck with orange legs, known for his love of tap dancing and taking dance lessons. He's a gentle and kind-hearted boy and plays the keyboard in the Hapidanbu.

First Appearance: 1994
Gender: Male
Creator: Sanrio
Species: Duck

At the bottom of the page, there is a row of smaller Sanrio character icons and a "Back to Search Page" button.

Example Step-by-Step Search Sessions:

1. One word query

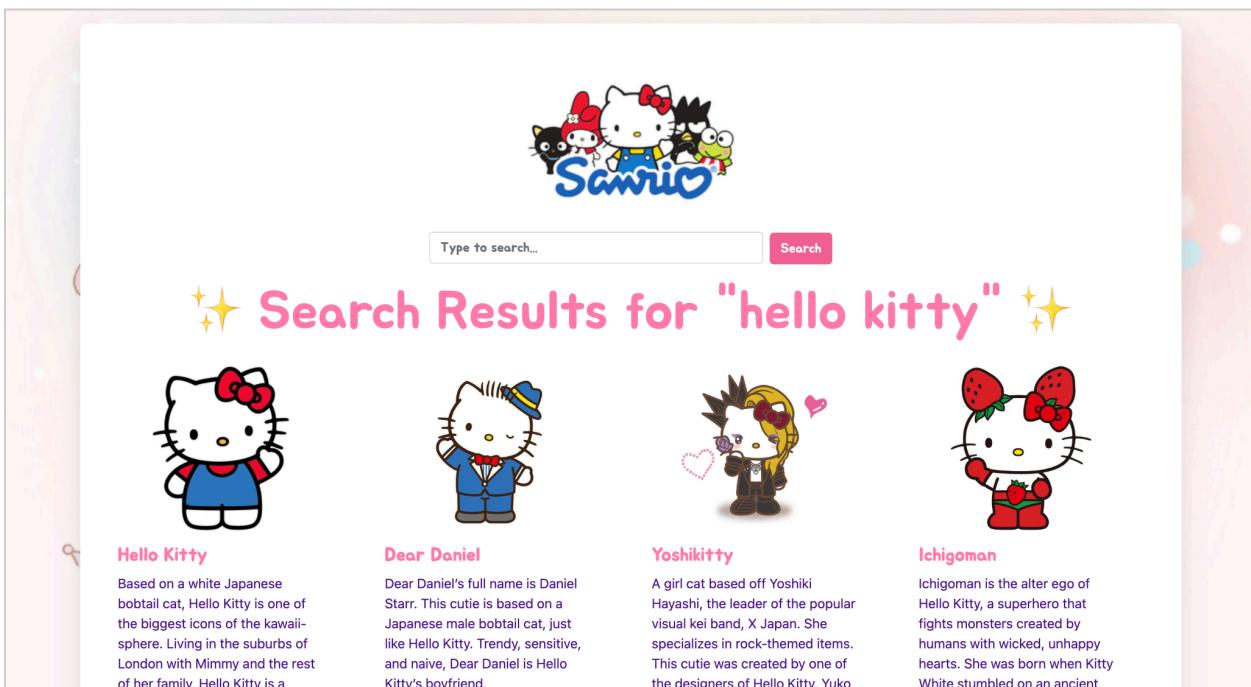


The screenshot shows a search results page for the term "kitty". At the top, there is a Sanrio logo featuring Hello Kitty and her friends. Below the logo is a search bar with the placeholder "Type to search..." and a pink "Search" button. The main title "Search Results for 'kitty'" is displayed in a large, pink, stylized font with sparkles. The results are presented in four cards, each containing an illustration of a Hello Kitty character and its name.

Character	Description
Ichigoman	Ichigoman is the alter ego of Hello Kitty, a superhero that fights monsters created by humans with wicked, unhappy hearts. She was born when Kitty White stumbled on an ancient
Hello Kitty	Based on a white Japanese bobtail cat, Hello Kitty is one of the biggest icons of the kawaii-sphere. Living in the suburbs of London with Mimmy and the rest of her family, Hello Kitty is a
Yoshikitty	A girl cat based off Yoshiki Hayashi, the leader of the popular visual kei band, X Japan. She specializes in rock-themed items. This cutie was created by one of the designers of Hello Kitty, Yuko
Mimmy	Hello Kitty's twin, Mimmy is a cat that looks a whole lot like her sister with her beautiful white hair. The only difference is that she wears her ribbon on the right side of her ear (Hello Kitty wears

If the inputted search query matches with the words that exist in the character_name and description, Elasticsearch will list the characters that match with the result and show in the search result page.

2. Multiple word query



The screenshot shows a search results page for the term "hello kitty". The layout is identical to the previous one, with the Sanrio logo at the top, a search bar, and a main title "Search Results for 'hello kitty'" with sparkles. The results are shown in four cards.

Character	Description
Hello Kitty	Based on a white Japanese bobtail cat, Hello Kitty is one of the biggest icons of the kawaii-sphere. Living in the suburbs of London with Mimmy and the rest of her family, Hello Kitty is a
Dear Daniel	Dear Daniel's full name is Daniel Starr. This cutie is based on a Japanese male bobtail cat, just like Hello Kitty. Trendy, sensitive, and naive, Dear Daniel is Hello Kitty's boyfriend.
Yoshikitty	A girl cat based off Yoshiki Hayashi, the leader of the popular visual kei band, X Japan. She specializes in rock-themed items. This cutie was created by one of the designers of Hello Kitty, Yuko
Ichigoman	Ichigoman is the alter ego of Hello Kitty, a superhero that fights monsters created by humans with wicked, unhappy hearts. She was born when Kitty White stumbled on an ancient

If the inputted search query matches with the multiple words that exist in the character_name and description, Elasticsearch will list the characters that match with the result and show in the search result page.

3. Partial match

The screenshot shows a search results page for "white cat". At the top, there is a logo featuring Hello Kitty and her friends, followed by a search bar with the placeholder "Type to search..." and a "Search" button. Below the search bar, the text "Search Results for \"white cat\" " is displayed in a pink, stylized font, flanked by two small star icons. The search results are presented in a grid format with four items:

- Little Lovin**: A boy dog with brown fur and droopy black ears. Littlelovin loves eating ogura toast and wears a white hat.
Score: 8.97
- Pururun Kyupi**: Pururun Kyupi is a white water fairy that can only be seen by people with beautiful hearts. Kyupi enjoys fortune-telling and can make everyone happy with her presence.
Score: 8.87
- Boo Gie Boo**: Boo Gie Boo is a naughty white dog who loves adventures and wears a blue hat. This retro character debuted in 1978.
Score: 8.71
- Kuromi**: Kuromi is a cool and cute white rabbit with a black jester's hat. Though a villain, she's very girly and enjoys romantic stories and good food.
Score: 8.23

If the inputted search query matches with some of the characters that exist in the character's information, Elasticsearch will list the information of that character on the search results page.

4. Ranking

```
else:
    # Define search body with multi_match and match_phrase for exact sequence matching
    body = {
        "size": page_size,
        "from": page_size * (page_no - 1),
        "query": {
            "bool": {
                "should": [
                    # Exact phrase match for the name field
                    {
                        "match_phrase": {
                            "name": {
                                "query": keyword,
                                "boost": 5 # Give the 'name' field a higher boost (more important)
                            }
                        }
                    },
                    # Exact match or partial match (wildcard) for the name field
                    {
                        "wildcard": {
                            "name": {
                                "value": f"{keyword.lower()}*",
                                "boost": 4 # Boost partial matches (lower priority than match_phrase)
                            }
                        }
                    },
                    # Match query across multiple fields with a focus on name
                    {
                        "multi_match": {
                            "query": keyword,
                            "fields": ["name^3", "description^2", "creator_name", "species"],
                            "type": "best_fields",
                            "fuzziness": "AUTO"
                        }
                    }
                ],
                "minimum_should_match": 1
            }
        }
    }
```

The Ranking system on which data is going to appear first is based on its hit scores that is done by the query that's interacting with the Elasticsearch.

Discussion

Limitation of the System

1. Lack of Autocomplete or Suggestions

Currently, the search engine does not offer an autocomplete or suggestion feature. This can affect the user experience, especially for those who might not be sure about the exact spelling of a character's name. Adding a real-time suggestion or autocomplete function would help users find what they are looking for more easily and quickly. It would also make the search results more relevant and user-friendly, guiding users as they type.

2. Image Handling and Validation

The search engine retrieves image URLs directly from the Elasticsearch database, but if any of these URLs are broken or lead to missing images, it negatively impacts the user experience. To fix this, the system needs a better way of checking whether the images are valid before displaying them. Implementing a fallback image (like a default image for missing or invalid URLs) would also ensure a smoother experience for users, even if the images cannot be found.

3. Security Concerns

At present, sensitive credentials (such as the ELASTIC_PASSWORD) are directly included in the code. This is a security risk, particularly in production environments where data protection is critical. It's recommended to use more secure methods, like storing credentials in environment variables or a dedicated secrets management system. This would ensure that sensitive information is protected and not exposed in the source code.

4. Limited Character Information

While the character detail pages show basic information like the character's name, description, and first appearance, there is potential to enrich these pages. Providing more context, such as the character's background, related characters, or links to related media (e.g., TV shows, movies, or games), could improve the user's understanding and engagement with the content. This additional information would create a richer experience for users interested in learning more about the characters.

5. Inconsistent Handling of Search Input

The system currently lacks clear rules for how it handles search input, especially regarding case sensitivity, punctuation, and stop words. This inconsistency could result in incomplete or inaccurate search results, especially when users search for characters with names that include special characters, hyphens, or spaces. Clearer guidelines for handling these variations would improve search accuracy and make the system more robust.

6. Absence of Caching Mechanisms

The application does not currently implement caching for repeated or frequent searches. Caching is important because it helps improve performance by storing and reusing data that has already been retrieved. For common searches, caching could significantly reduce the load on the Elasticsearch server and improve response times, making the application more efficient for users.

Technical Difficulties and Challenges, and How to Handle

The main challenge of this project was utilizing Elasticsearch effectively, as it stops running when the terminal session is terminated. For instance, if we shut down or restart the desktop, the terminal session ends, and we need to reinstall Kibana and run the terminal again to get the Elasticsearch token and localhost connection from Kibana.

Due to Elasticsearch and Kibana being new tools for us, we spent most of our time studying and experimenting with them. As a result, we had less time to develop and design the webpage, so it may lack some functionality and visual appeal that could attract more users. However, the search function is fully operational.

Lessons Learned

Working on this project provided us with valuable insight about search engines. We studied how search engines work and how to integrate them into websites. Additionally, We explored a new method—Elasticsearch—and learned how to implement this tool to create an efficient search system and optimize it for web servers.

Through this project, We learned skills that will help me develop more effective and user-friendly search systems in the future. Furthermore, these skills are valuable for our career path, as understanding search technology and tools like Elasticsearch can be highly relevant in fields such as web development, data science, and information technology.

Opportunities for Future Improvements

Facing the challenges of doing this project gives us several opportunities for future improvements. First, We could establish Elasticsearch and Kibana as services, allowing them to start automatically when the machine boots, eliminating the need for manual reinstallation and setup on every reboot. It could save time and ensure that the system functions smoothly, even after restarting the desktop.

To improve the user experience, we might focus more effort on improving the website's design and function. This could involve creating a user-friendly interface with interactive elements, and introducing additional features such as filters, and sorting choices to make search results more customizable and accessible. Moreover, we may expand the web application to support more languages.

As we obtain more Elasticsearch experience, we can focus on improving query performance and indexing, which allows the search system to handle larger datasets and deliver faster, more accurate results.

In conclusion, All of these improvements will make the web application more robust, scalable, and user-friendly, providing a better user experience and positioning the project for larger use cases.

Conclusion

This project focuses on "Sanrio Character", a popular game in Thailand, and provides an extensive database with detailed information on game characters, their statistics, and skills. The system stands out for its comprehensive content, user-friendly design, and intuitive navigation. We believe that this project initially conceptualized for academic purposes, has the potential to become a valuable resource for the broader gaming community. It exemplifies their commitment to enhancing the gaming experience and reflects their passion for technology and innovation. The project demonstrates an impressive integration of technical skills and a clear understanding of the needs of its intended audiences.