

Retro Sounding Progress Report

Austin Cari
Daniel Belousov
Steven Arbuckle
austin.cari@wsu.edu
daniel.belousov@wsu.edu
slarbuckle@hotmail.com
Washington State University
Vancouver, USA

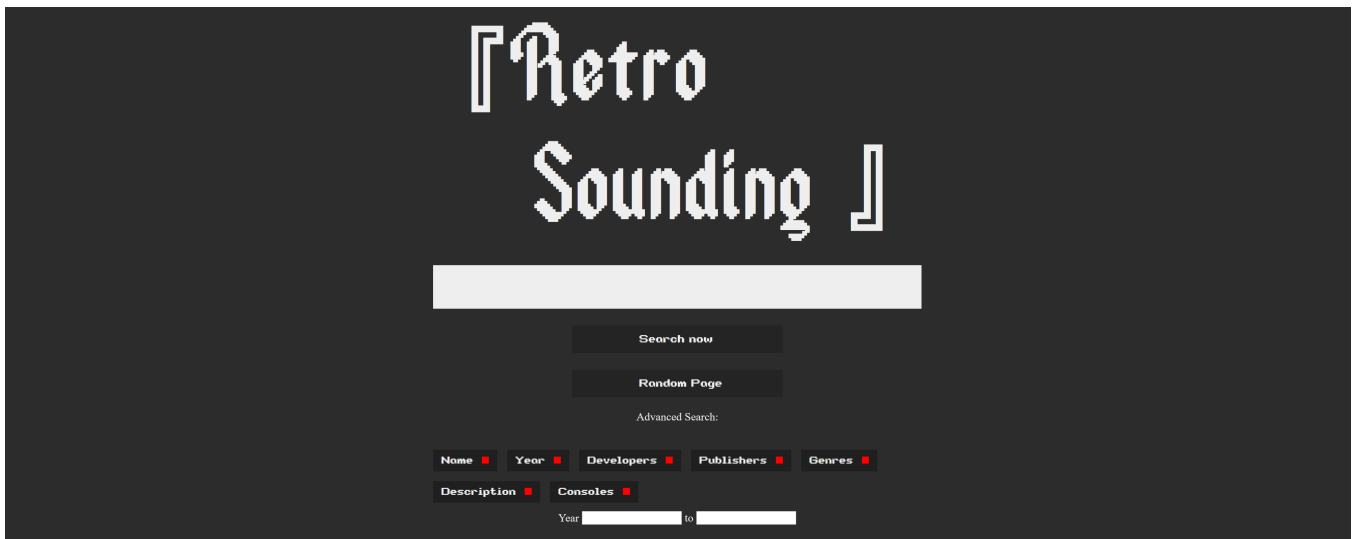


Figure 1: The homepage of Retro Sounding.

ACM Reference Format:

Austin Cari, Daniel Belousov, and Steven Arbuckle. 2020. Retro Sounding Progress Report. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.114x/1122445.11224xx>

1 INTRODUCTION

Retro Sounding is a collaborative team effort to build a web application that searches over a large index of retro-style video games as early from 1972, up until 2002. Many different websites and archives exist to document these games, but there is no exhaustive source on the topic. This application will function as a search engine that runs in the browser and can search over a database of 11,692 titles. Users will be able to look through pages of results and read more

information on every game, as well as linking to external sources like Wikipedia.

Retro Sounding is designed to look minimalist and compact but also retain the style that was iconic about video games in the 80's and 90's. The name Retro Sounding was the impromptu name of the project until an actual graphic was created using the name, and it stuck. The main goal of this project is to provide a simple user interface that users can learn quickly and be able to look for their favorite games and enjoy some nostalgia.

1.1 The data

There are many common components to a video game that are saved with each record. Information that is searched over includes the title, publishers, genres, a description, and platforms the game was released on. Information not searched over includes the images of the game and URLs to external sources.

The data is within the whoosh! index, and the the data is stored in the following schema:

```
schema = Schema(GID=ID(stored=True, unique=True),
name=TEXT(stored=True, analyzer=SimpleAnalyzer()),
release_year=NUMERIC(stored=True),
developers=KEYWORD(stored=True, commas=True, lowercase=True),
```

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference'17, July 2017, Washington, DC, USA

© 2020 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.114x/1122445.11224xx>

```

publishers=KEYWORD(stored=True,commas=True,lowercase=True),
images=KEYWORD(stored=True,commas=True),
consoles=KEYWORD(stored=True,commas=True,lowercase=True),
sources=KEYWORD(stored=True,commas=True),
genres=KEYWORD(stored=True,commas=True,lowercase=True),
description=TEXT(stored=True)
)

```

Video game titles were scraped from Wikipedia and individual scrapers were made for each specific page. A large hub page existed that linked to comprehensive lists of video games organized by categories[16]. Articles that grouped games by console/platform were the primary pages we wanted to scrape, since these games made for a certain console were naturally grouped together relative to their release years. The list of consoles we concerned ourselves with was the Apple II[4], the Classic Mac OS[11], Microsoft DOS[3], arcade cabinets[5], Atari 2600[6], Famicom[7], NES[13], SNES[15], Nintendo 64[12], Gameboy[10], Gameboy Advance[8], Gameboy Color[9], the Sega Genesis[14] and the ZX Spectrum computer[17].

1.2 Web scraping techniques

The data was scraped one console at a time in an arbitrary order. A scraper written in python was made specifically for each page, one page for each console.

Each page will require its own scraper since the format of Wikipedia pages are not uniform. Beautiful Soup and the lxml etree module are used in the different scrapers to parse through HTML. Any title that was released after 2002 is not saved. The data scrapes the name, release year, developers, producers from the list page. If the title is also a link to another page then it visits that page and grabs the genre, and image from the infobox if it exists, and uses the first few paragraphs as the description.

Once all the information is gathered, it is written into a CSV file that will later be imported into a Whoosh! index.

2 FEATURES

The main feature of Retro Sounding is the ability to search for different video games using text queries. Users can look through a list of relevant results and select a game to see more info. Filters are also available to search based on only certain attributes like publisher or publication year.

2.1 Images

An image, most likely the box art or cover, will be displayed beside each result in the results page. This helps each result to stand out and be easily identified. Not every title has images available, and a default placeholder image will be used instead and displayed on the timeline.

2.2 External Links

Links to our data source will be included with every result that has one. This will encourage further reading from users wish to find more in-depth information. If an entry does not have a link present no external link will be displayed.

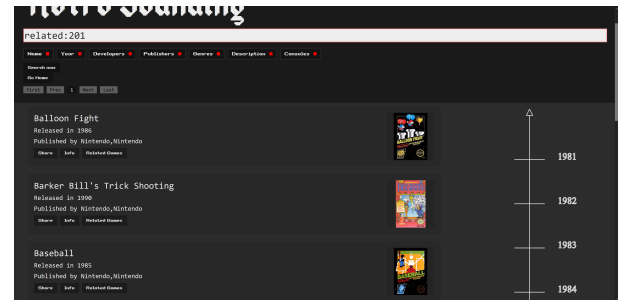


Figure 2: Results page for related games to 'Duck Hunt' with an ID of 201.

2.3 Related Results

Users can look for games similar to any other entry by pressing the "Related Games" button. They will be brought back to the results page with a specialized query that searches over the database looking for similar attributes like console, genre, developer and publisher. This feature helps the user experience by executing a specially weighted query that contains four attributes of the previous title, all at the click of a button.

Each attribute is given a particular weight in the specialized query. Games with a similar console are weighted at 0.7, genres at 1.3, developers at 1.0 and publishers at 0.9. Titles that have many similar attributes are weighted even higher than games with just one. When a user is viewing information about a particular game, they can click on attributes like platform and genre to view other games specifically with that particular attribute, instead of a related search.

2.4 Advanced Search

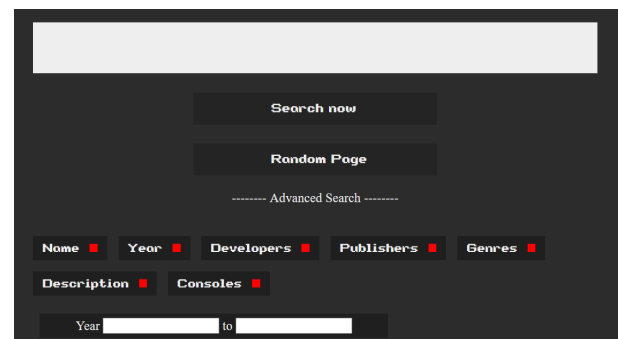


Figure 3: Filters that are available to turn on/off when searching.

A list of filters is displayed on the search page where users can turn off certain attributes to be searched for. These filters can help narrow down your results. For instance, if you searched for Nintendo, you would get many games made and published by Nintendo, but also other games where Nintendo is mentioned in the description. If you want to find only the games that Nintendo has published, you would only want the publisher/developer filters checked. Other

attributes that can be turned off are name, consoles, genres, year and description.

If the user likes, they can also specify a range of release years to search over. If the user only describes a minimum then the query will look at games that were released afterwards. If the user only describes a maximum year, then the query will only look at games that were released before.

The search bar and filters are still available to the user on the results page, where they can enter a new query or refine their existing query. An advanced search feature will help users who know what they are searching for narrow down their search and ignore unnecessary attributes.

2.5 Random page search

Users will be able to hit the "Random page" button. Inspired after the very similar feature on Wikipedia, this will grab a random game from the Whoosh! Index and bring the user straight to the entry's page without having to enter a search. This lets users find games they would never be able to find other ways. Leave it up to chance! The user will be sent to a random page if they attempt to search with an empty query. This feature helps the user experience because sometimes users are not on your website to find information, sometimes they are just bored. And this lets users find games they would have never tried to search for without having known about it.

2.6 Timeline

Displayed next to the list of search results will be a timeline that has each game in the results list. The user can select an image from this list to go to that game's entry page, as an alternative to using the main list of results.

This feature is a visual aid that helps users see interesting trends as games are released over the years. Watching the transition of a franchise over multiple decades or seeing the increasing frequency of games in a certain genre can lead users to draw interesting and meaningful conclusions about trends in the release of video games.

2.7 Sharing results

If you find an entry that you find particularly interesting, you can click the share button which copies the url pointing to that game's entry page to your clipboard. You can share the link with others so they can see that page too. This feature is a tried-and-true method that every other website offers, a simple way to share articles and pages with friends. If people are sharing the website with each other, we will have more visitors come to the site and enjoy the work we've done.

3 SEARCH ENGINE

Retro Sounding uses Whoosh! to index and search over our data. All of the CSV files are uploaded into the whoosh index before it is searched using a special script that compares each new entry against all existing entries to make sure that duplicate results are discarded. However, some games have multiple entries because of subtle differences in attributes. For instance, Street Fighter was originally developed by Capcom for the arcade cabinet. In 1988 however, a professional port was published by US Gold! a year

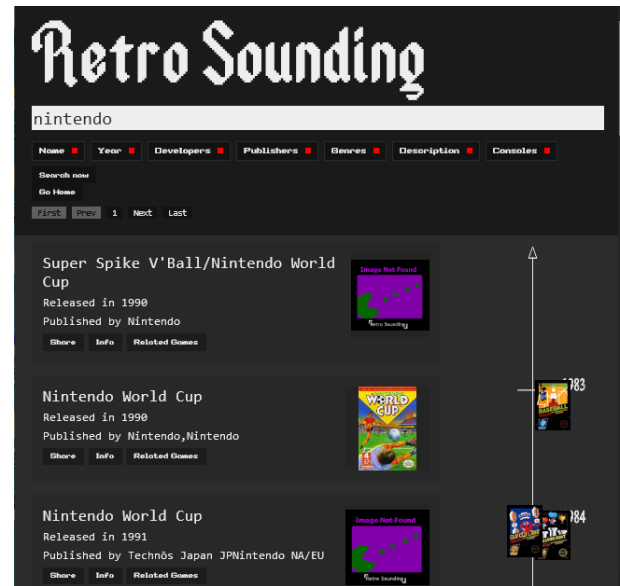


Figure 4: Results page for the query 'Nintendo'.

after the original release of Street Fighter which was licensed from Capcom. Franchises that have strict control over their game (like Pokémon) will only appear once since these games were made specifically for a very particular platform and not ported to other consoles. Despite having similar names, images and description, these are treated as different entries in the database.

3.1 User Interface

Users can search the database by providing text queries into the search bar and simply pressing enter. Whoosh! will return results that match the specified text, and prioritise games that have multiple matching attributes, while still allowing results that are missing some. Results are displayed on top of each other, with the most relevant results at the top of the list. The search bar and filters are still accessible at the top of the page so users can modify their queries or search for something else (see Fig. 4). However, if they need to, they can go back to the homepage by clicking Go Home.

Results are divided up into sets of 10, and users can navigate between pages using the First, Prev, Next and Last buttons located at the top and bottom of the search results. Each page generates its own page and the timeline located at the right of the page will only show results for the current page, not the entire query. Every result will have a title, a release year, the publisher and developer along with the image to the right. The Share, Info and Related Games button are also next to each entry and is unique to that entry (see Fig. 5). When a user finds an interesting game that they want to know more about, they can click on the info button and they will go to that game's entry page, which contains all the information the index has about that game. The console and genre attribute will be links that will search the database for those specific terms, with a special query in the search bar.

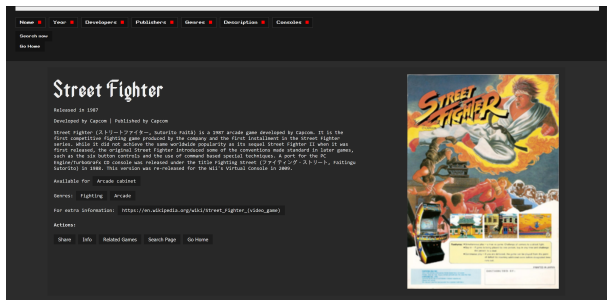


Figure 5: Entry page for the popular title Street Fighter.

3.2 Performance

Queries are executed instantaneously and users do not have to wait at all for results to be returned. The relevance of the results is also extremely high, with almost all queries returning upwards of 90 percent relevant results. Individual titles that contain multiple entries due to multiple releases, ports to other platforms or licensing agreements tend to clump up together on the results list together and push down other relevant results. The sheer number of results returned is also not very helpful, since relevance starts to drop off between pages 5 and 10 for general terms, and even into page 2 for very specific terms or genres. However very popular franchises and publishers will have no shortage of results to choose from, like Nintendo or Pokemon.

4 CONCLUSION AND FUTURE WORK

The current scope of the project will prove to be a use full tool for finding information about retro-style video games for hobbyists. As the project continues to grow, the main objective will be to populate the database with games that do not already exist on other websites, and to clean data that was scraped in error. Some entries refer to the wrong page, or have bad descriptions. Errors like these are hard to search for, and so could not be fixed in a timely manner.

Potential features for the application would be to incorporate fuzzy search and N-grams, change between ranking methods (TF-IDF and BM25) and create even more stylish pages that pay homage to the unique art style of the era.

Creating an image gallery that can support multiple images would be useful to show different box arts released in different regions and segments of the game.

There are other consoles that were considered to be added, but were either too hard to scrape or did not contain enough games. These consoles were the Sony Playstation, the Atari 5600 and 7800, PC-FX home console, the Neo-Geo, and homebrew/indie games developed for desktop environments.

Extra styling needs to be done with the elements on the page, like adding a pixelated cursor in the shape of a left-facing chevron (<) on the search bar. The site does not do enough to draw any contrast in color to the eye, and it does not draw the user's attention as much as preferred. A background image that is drawn in the style of pixels would keep consistent with the already existing style. Another stylish feature that would be considered is adding a light mode alternative to the pages for those who do not like the default dark theme.

The current implementation works well and is a good example of what the team had in mind. While there exists some ambiguity between entries, queries are always satisfied and show relevant results. Users can navigate the website easily and modify their search to their liking as well as find similar games. The website's design is simple so that users do not feel overwhelmed with images and moving parts. It was also a very conscious decision to use a dark theme for the application, and it turned out great.

We can get missing data from sites such as Retrogames.gz[2], and Metacritic.com[1] who have complicated descriptions, and extra pictures for games on many consoles that we have already described. Adding a feature to handle a gallery of images would help show multiple segments of a game, and add different box art from different regions.

As our team continues to work on Retro Sounding, we will keep these ideas in mind as we try to create a robust and complete database of video games from the golden era of gaming.

REFERENCES

- [1] [n.d.]. Action Games by Metascore. <https://www.metacritic.com/browse/games/genre/metascore/all>
- [2] [n.d.]. Game consoles. <https://www.retrogames.cz/konzole.php>
- [3] Wikipedia contributors. [n.d.]. Index of DOS games — Wikipedia, The Free Encyclopedia. Retrieved December 5, 2020 from https://en.wikipedia.org/wiki/Index_of_DOS_games
- [4] Wikipedia contributors. [n.d.]. List of Apple II games — Wikipedia, The Free Encyclopedia. Retrieved December 5, 2020 from https://en.wikipedia.org/wiki/List_of_Apple_II_games
- [5] Wikipedia contributors. [n.d.]. List of Arcade Video games — Wikipedia, The Free Encyclopedia. Retrieved December 5, 2020 from https://en.wikipedia.org/wiki/List_of_arcade_video_games
- [6] Wikipedia contributors. [n.d.]. List of Atari 2600 Games — Wikipedia, The Free Encyclopedia. Retrieved December 4, 2020 from https://en.wikipedia.org/wiki/List_of_Atari_2600_games
- [7] Wikipedia contributors. [n.d.]. List of Family Computer Games — Wikipedia, The Free Encyclopedia. Retrieved December 5, 2020 from https://en.wikipedia.org/wiki/List_of_Family_Computer_games
- [8] Wikipedia contributors. [n.d.]. List of Game Boy Advance games — Wikipedia, The Free Encyclopedia. Retrieved December 5, 2020 from https://en.wikipedia.org/wiki/List_of_Game_Boy_Advance_games
- [9] Wikipedia contributors. [n.d.]. List of Game Boy Color games — Wikipedia, The Free Encyclopedia. Retrieved December 4, 2020 from https://en.wikipedia.org/wiki/List_of_Game_Boy_Color_games
- [10] Wikipedia contributors. [n.d.]. List of Game Boy Games — Wikipedia, The Free Encyclopedia. Retrieved December 1, 2020 from https://en.wikipedia.org/wiki/List_of_Game_Boy_games
- [11] Wikipedia contributors. [n.d.]. List of Macintosh games — Wikipedia, The Free Encyclopedia. Retrieved December 6, 2020 from https://en.wikipedia.org/wiki/List_of_Macintosh_games
- [12] Wikipedia contributors. [n.d.]. List of Nintendo 64 Games — Wikipedia, The Free Encyclopedia. Retrieved November 29, 2020 from https://en.wikipedia.org/wiki/List_of_Nintendo_64_games
- [13] Wikipedia contributors. [n.d.]. List of Nintendo Entertainment System Games — Wikipedia, The Free Encyclopedia. Retrieved November 29, 2020 from https://en.wikipedia.org/wiki/List_of_Nintendo_Entertainment_System_games
- [14] Wikipedia contributors. [n.d.]. List of Sega Genesis Games — Wikipedia, The Free Encyclopedia. Retrieved December 3, 2020 from https://en.wikipedia.org/wiki/List_of_Sega_Genesis_games
- [15] Wikipedia contributors. [n.d.]. List of Super Nintendo Entertainment System Games — Wikipedia, The Free Encyclopedia. Retrieved December 3, 2020 from https://en.wikipedia.org/wiki/List_of_Super_Nintendo_Entertainment_System_games
- [16] Wikipedia contributors. [n.d.]. List of Video Games — Wikipedia, The Free Encyclopedia. Retrieved December 6, 2020 from https://en.wikipedia.org/wiki/Lists_of_video_games
- [17] Wikipedia contributors. [n.d.]. List of ZX Spectrum games — Wikipedia, The Free Encyclopedia. Retrieved December 5, 2020 from https://en.wikipedia.org/wiki/List_of_ZX_Spectrum_games