Smithsonian Learning Lab User Search

Digital Analytics Evaluation Report

Allysha A. Leonard, Jesse Ludington, Tess Porter, Arthi Sundararajan, and Wenjing Wu

Pratt Institute

Table of Contents

Table of Contents

1. Executive Summary

The [Smithsonian Learning Lab](#) is a free platform for exploring millions of digital resources held by Smithsonian museums, creating interactive learning resources, and sharing discoveries and creations with others. Faced with the increasing number of users and collections, the Lab wants to understand how users interact with the platform's primary feature: search. Our research focuses on two guiding questions: how are users using the Smithsonian Learning Lab's search, and are users finding the content they're looking for?

Findings, based on data pulled from Google Analytics, are grouped into two areas: search behavior and search results. According to our research, we suggest the following:

- **Search Behavior Recommendations:**
  - Simplify the search process outlined on the home page
  - Feature popular collections and educational resources on the home page
  - Make search refinement easier by redesigning the search results page
- **Search Results Recommendations:**
  - Optimize search results with resource relevancy and connected collections
  - Connect collection keywords with a controlled vocabulary
  - Encourage exploration with a "Part of Collection" resource filter, cross-tagging, and linked open data
  - Call for Smithsonian-Created collections based on popular search terms,
  - Support museum search through white-listed name searchability
  - Improve language and accessibility with new refine search filters

2. Introduction

The [Smithsonian Learning Lab](#) is a creation of the Smithsonian Office of Educational Technology (OET). The Lab's purpose is to facilitate the discovery and use of digital materials held by the Smithsonian Institution's wide range of museums. The Lab is geared towards learners and educators of all ages, including both formal and informal learners, and emphasizes sharing and creative use of the materials they host. Materials on the Lab are divided into two categories: resources, which are materials digitized by Smithsonian museums, and collections, which are user-created groups of resources.

Because inquiry is at the heart of the Smithsonian Learning Lab, OET is interested in developing a deeper understanding of how users are using the website's search function to find collections. They are also interested in finding out whether searches are successful or unsuccessful.
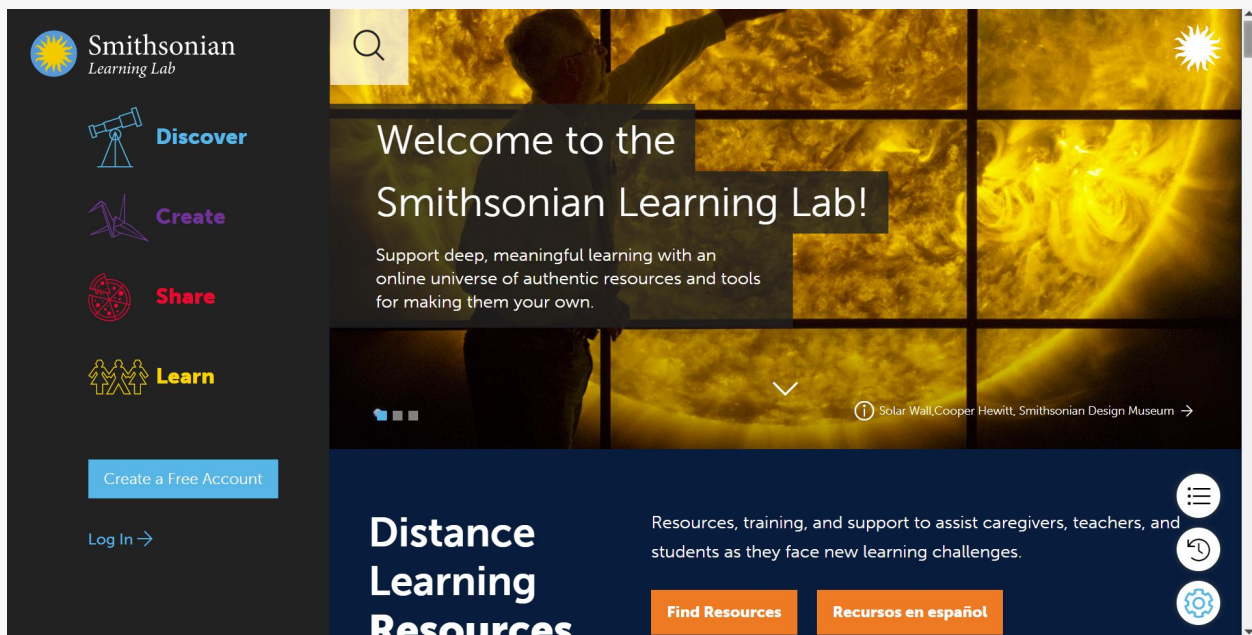

Fig. 1: Smithsonian Learning Lab Home Page on Desktop.

3. Research Objectives

Through our analysis, we strive to answer the following research objectives:

1. **Search Behavior:** How are users using the Smithsonian Learning Lab's search?
2. **Search Results**: Are users finding the content they're looking for?

Our research incorporates two user perspectives for using the Lab's search, each with its own measures of a successful search:

- **Exploratory**, where users are uncertain what exactly they are looking for.
- **Results Driven**, where users are looking for something specific.

Answering these research objectives will help the Lab's team better understand what users are looking for, how to improve the Lab's search interface, how to support Smithsonian educators in building new educational content to fill gaps, and how to improve user supports for creating and publishing collections that are easily discoverable by others.

Fig. 2: Smithsonian Learning Lab Home Page on Mobile.

4. Methodology

4.1 Methods

To better understand how users are utilizing the Smithsonian Learning Lab website, the client has provided a plethora of data from Google Analytics for the Pratt team to analyze. The analysis spans from January 1, 2020 to April 1, 2022, to thoroughly analyze trends in website usage.

4.2 Tools and Metrics

Using data collected from Google Analytics, aggregation, analysis, and visualization were performed using Google Sheets, Microsoft Excel, Google Data Studio, and Tableau. Our analysis utilizes the following metrics: users, source, page views, unique pageviews, exit pages, session duration, landing pages, sessions with search, unique searches, time after search, average search depth, search terms, search filters, search refinements, and search exits. Our analysis also filters out users who fall within the "spam" segment, as defined by the Lab team, in order to remove outlier data.

To support the Lab in moving forward from this report, the Pratt team developed an adaptable [dashboard in Google Data Studio](). The dashboard displays real-time insights on Lab search behavior on a quarterly basis and can be manipulated to show data from a specific time period.

4.3 Limitations

While a significant amount of data related to search behavior on the Lab site has been collected, it is mainly quantitative, and there is a lack of insight into user motivation and end-use of content. Dividing the Lab's audience into exploratory and results-driven user types is an assumption made based on search behavior. In addition, we have analyzed session duration metrics in order to better understand user interactions with content, but we do not have additional insight into whether these sessions are based on independent use or teaching in the classroom.

While there are goals set up in Google Analytics for the Smithsonian Learning Lab, they are not measuring meaningful and tracked target objectives for the Lab. It would be very beneficial for the Lab to set up goals connecting user search behavior with users accessing available resources on the site to continue the analysis of search success moving forward.

5. Key Findings

5.1 Search Behavior Findings

5.1.1. Search Activity

Sessions with search account for **14.98%** of all sessions by non-spam users during the time period analyzed. There were 1,471,594 new users during this time period, with **13.30%** of those users (196,196) conducting a search. 502,894 unique searches were conducted during the time period analyzed, for an **average of 17,960 unique searches** per month. The home page of the Lab website is the most popular start page for users conducting a search, with **24.47%** of users beginning their search journey there.

The percentage of sessions with search consistently peaks in **February** of 2020, 2021, and 2022 at **24.04%**, **18.36%**, and **15.53%** respectively. When compared to the overall sessions with search in Jan 1, 2020—Apr 1, 2022, it becomes clear that each of these months is well above average.
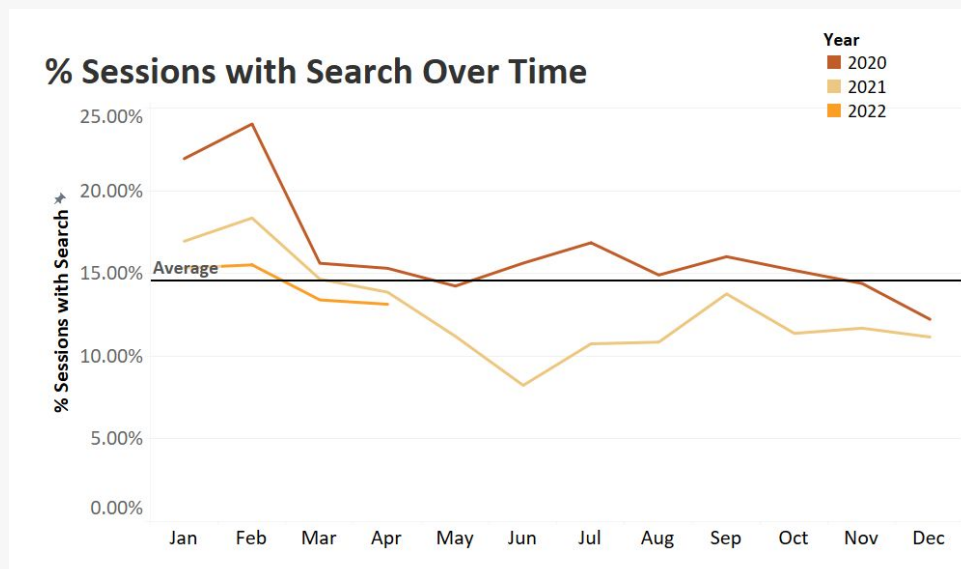

Fig. 3: Percentage of Sessions With Search Each Month for 2020, 2021, and 2022.

February has the highest percentage of sessions with the search of any single month across the entire period analyzed. Looking towards finding 5.1.2, perhaps this is because February is Black History Month. Seeing as **"black history"** is the most popular search term for the time period analyzed, this correlation is highly probable. It is possible that educators are asking students to perform themed searches during heritage months, or perhaps users simply see the Lab as a good place to find resources and collections related to cultural heritage months.

5. Key Findings

% Sessions with Search by Month

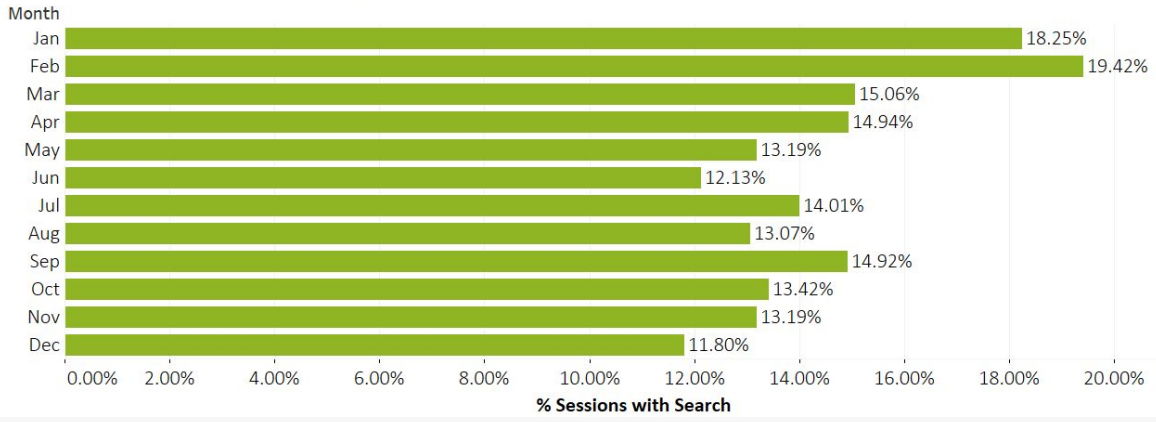| Month	% Sessions with Search
Jan	18.25%
Feb	19.42%
Mar	15.06%
Apr	14.94%
May	13.19%
Jun	12.13%
Jul	14.01%
Aug	13.07%
Sep	14.92%
Oct	13.42%
Nov	13.19%
Dec	11.80%

Fig. 4: Percentage of Sessions with Search by Month of the Year Aggregated From 2020, 2021, and 2022.

5. Key Findings

5.1.2 Search Terms

5.1.2.a. Top Search Terms

To understand patterns in search terms, we conducted a focused analysis of terms with over 100 unique searches (n= 192; 0.06% of the total terms searched), which accounted for a total of over **46,000** unique searches (14% of the total unique searches conducted) and an average of **240** unique searches for each term. To support the analysis of such a large dataset, search terms were sorted into overarching categories such as those for heritage months, school subject areas, pedagogical techniques, individuals, places, and particular topics.

The **top 10 search terms** by unique searches are: "black history," "hispanic heritage," "space," "art", "trail of tears," "slavery," "womens history," "civil war," "civil rights," and "music."
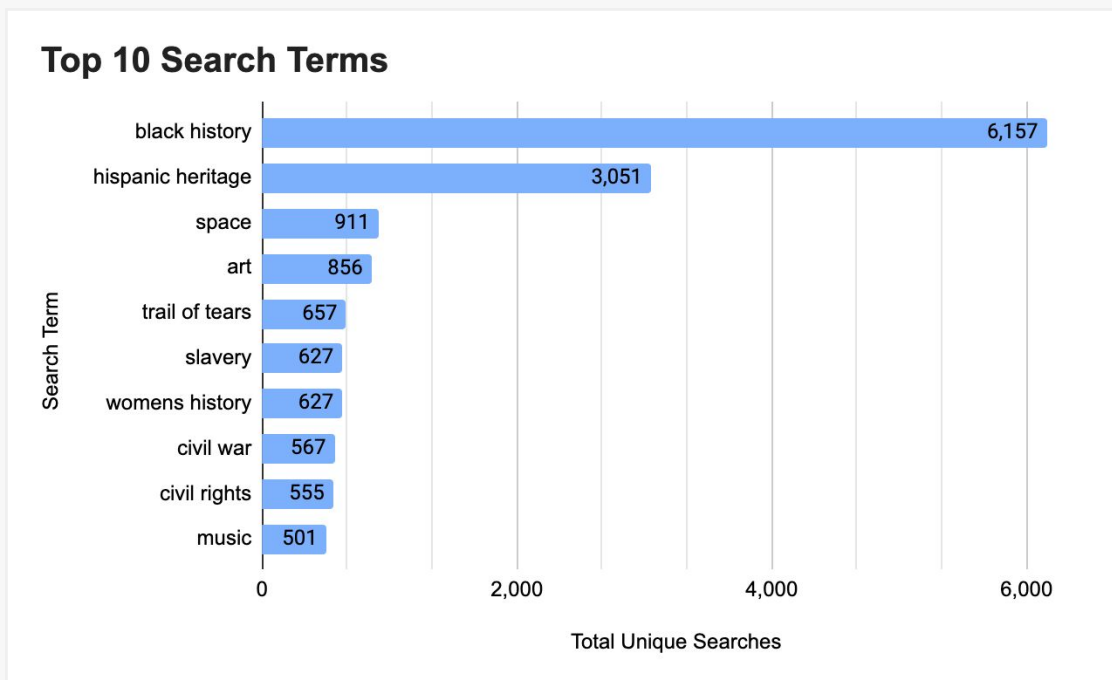

Fig. 5: Top 10 Search Terms.

5. Key Findings

The **top 10 search categories** by unique searches are: heritage months, subjects, astronomy, pedagogy, individuals, colonial America, Native Americans, ancient history, women, and places. Together, these comprise **61%** of the searches that have over 100 unique searches.
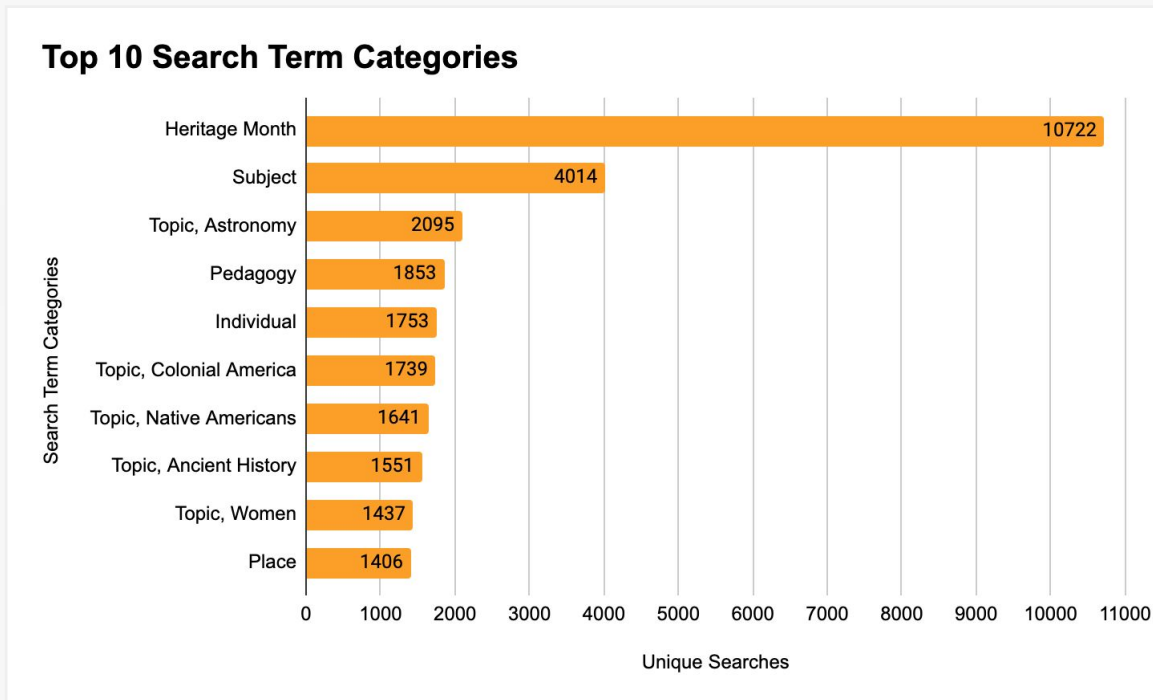
Top 10 Search Term Categories

Search Term Category	Unique Searches
Heritage Month	10722
Subject	4014
Topic, Astronomy	2095
Pedagogy	1853
Individual	1753
Topic, Colonial America	1739
Topic, Native Americans	1641
Topic, Ancient History	1551
Topic, Women	1437
Place	1406

Fig. 6: Top 10 Search Term Categories.

Searches related to heritage months are consistently the most popular. These include "black history," "hispanic heritage," "asian pacific american history," "american indian heritage," and "women's history." These comprise **23%** of the searches that have over 100 unique searches.

Searches for subject areas follow closely behind. These are highly varied and include 17 distinct search terms that received over 100 unique searches. These cover both broad subjects (art, music, math, history, social studies, and science) and more specific subjects (biology, psychology, chemistry, archaeology, astronomy, and health).

Searches for pedagogy, while only #4 in the Top 10 Total Unique Searches by Category, may be worth further attention given the Lab's target audience. Search terms in this category include: "easy pz," "lesson plans," "project zero," "games," "What Makes You Say That," "education," and "thinking routines."

5. Key Findings

5.1.2.b. Search Word Frequency

To understand patterns in search terms, we used a Word Cloud to track the frequency of particular words used in the top 5,000 searches ranked by the number of unique searches.

Fig. 7: Word Cloud Displaying Frequency of Search Terms.

The **top 10 most frequently used words**, alongside the number of their frequency, are:

1. American (n=117)

2. Art (n=94)

3. Ancient (n=78)

4. War (n=76)

5. History (n=75)

6. Black (n=59)

7. Rights (n=44)

8. Civil (n=42)

9. African (n=37)

10. World (n=34)

From this, we can deduce that **social studies topics connected to America, ancient civilizations, African Americans, and civil rights issues, as well as art topics, are among some of the most frequently searched.**

5. Key Findings

5.1.2.c. Searches for Museums

Approximately **4,900** unique searches (3% of total activity) involve searching for a museum. Of these searches, 64.5% were of hand-typed terms, while 35.5% used the Lab's pre-built filters for searching museum resources.
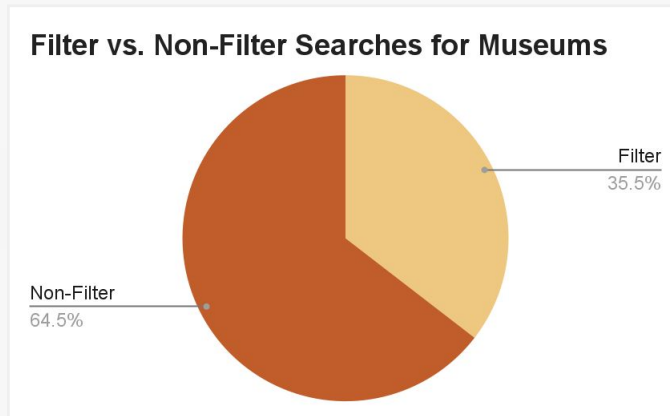

Fig. 8: Filter vs Non-Filter Searches for Museums.

Of the hand-typed terms, **69%** were of museum names, while the remainder were for museums external to the Smithsonian, the names of Smithsonian staff members, and topic terms (such as "Art Museum Summer Session").
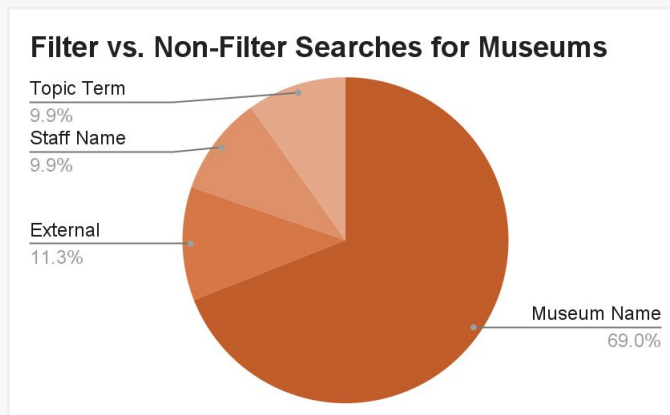
Fig. 9: Filter vs Non-Filter Searches for Museums.

These findings suggest that while searches for museums comprise a fraction of the total searches on the Lab, more users are hand-typing museum names than using the Lab's existing filters. This may be influenced by the fact that these filters are only available in a resource search – not a search for collections.

5. Key Findings

5.1.3 Subject Search Filters

5.1.3.a. Subject Search Filters

Approximately 41,153 unique searches involved the Lab's Refine Search Filters for collection subjects, which comprises **59%** of all searches for collections.
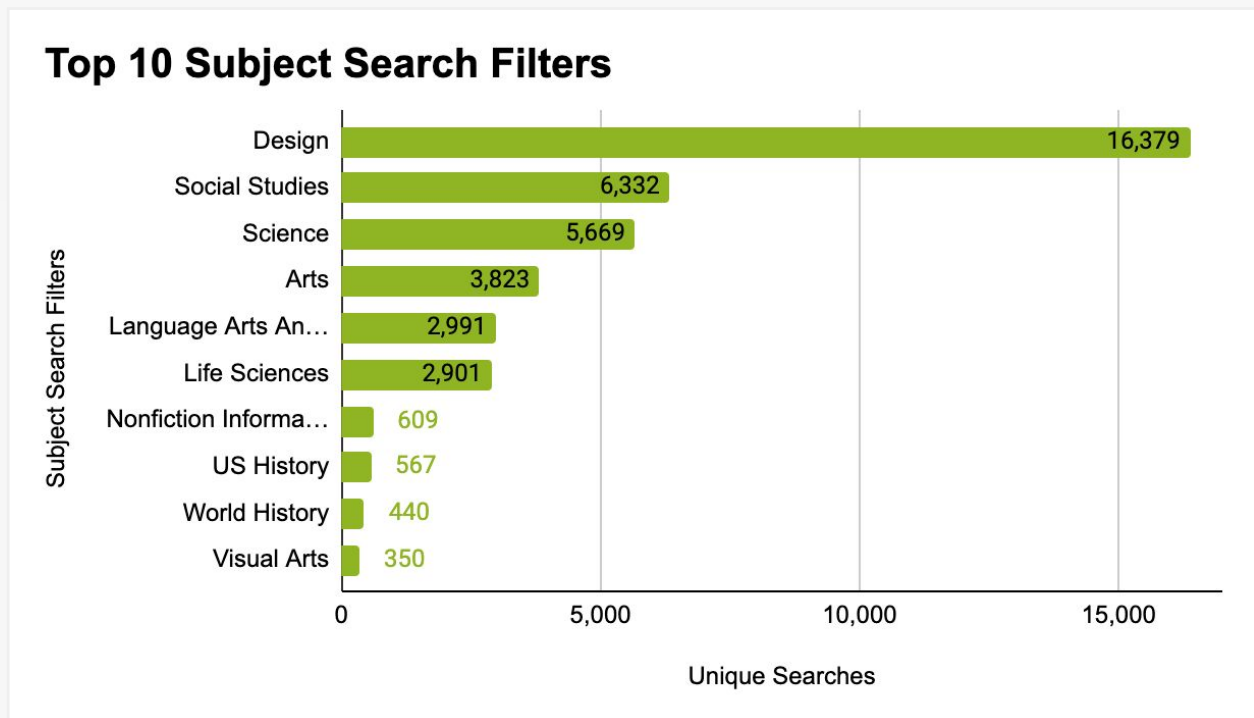

Fig. 10: Top 10 Subject Search Filters.

By far, the most popular subject filter is Design. The popularity of the following filters – Social Studies, Science, Arts, and Language Arts, and English – may at least be partially attributed to them being highlighted and linked on the Lab's home page.
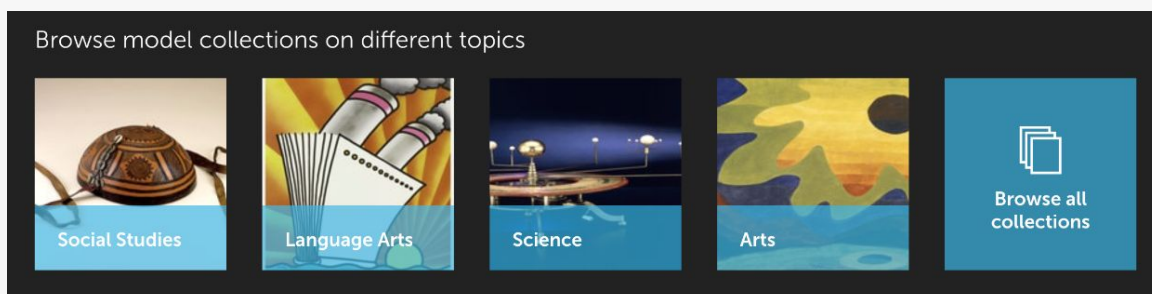

Fig. 11: Screenshot of Smithsonian Learning Lab Home Page Where Subject Filters are Linked.

5. Key Findings

5.1.3.b. Age Filters

Approximately 10,080 unique pageviews* involved the Lab's Refine Search filters for the collection age level. **Of these filters, the most popular is Elementary**, holding 31% of the total unique pageviews for age level filters, **followed by Primary**, holding 25%.
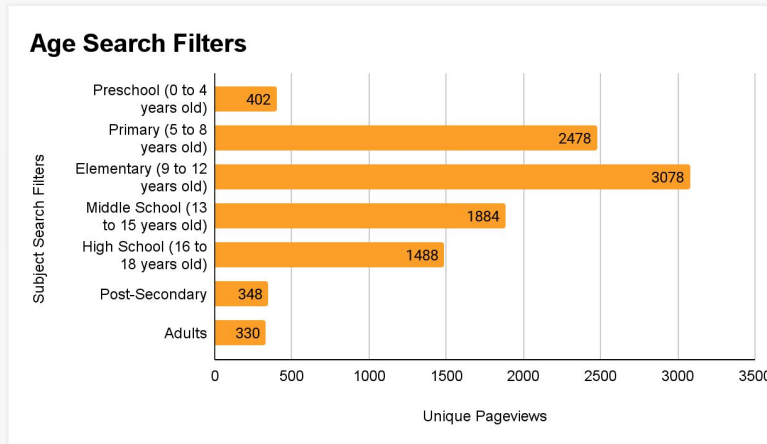

Fig. 12: Age Search Filters.

The popularity of filters does not align with the total amount of collections available in the Lab. The chart below aligns each age level filter with the total amount of published collections available in the Lab with that age level assigned.
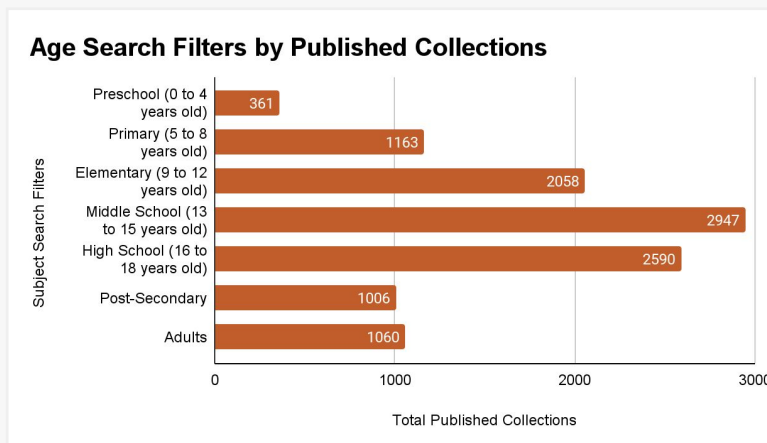

Fig. 13: Age Search Filters by Published Collections.

*Age Level Filters are not tracked in Google Analytics as search terms, so this analysis uses unique pageviews of results pages with these filters as a metric rather than unique searches.

5. Key Findings

5.1.3.c. Search Depth

The majority of search terms with the highest number of refinements are subject filters. This suggests that perhaps users are using subject filters as a jumping-off point before further refining their searches. Because users can apply a subject filter at any point in their search, it is impossible to determine whether the majority of these users are using filters in an exploratory or results-driven manner. We must therefore assume users applying filters have a mixture of motivations. It is also worth noting that **the majority of searches with high numbers of refinements are related to art and design.**
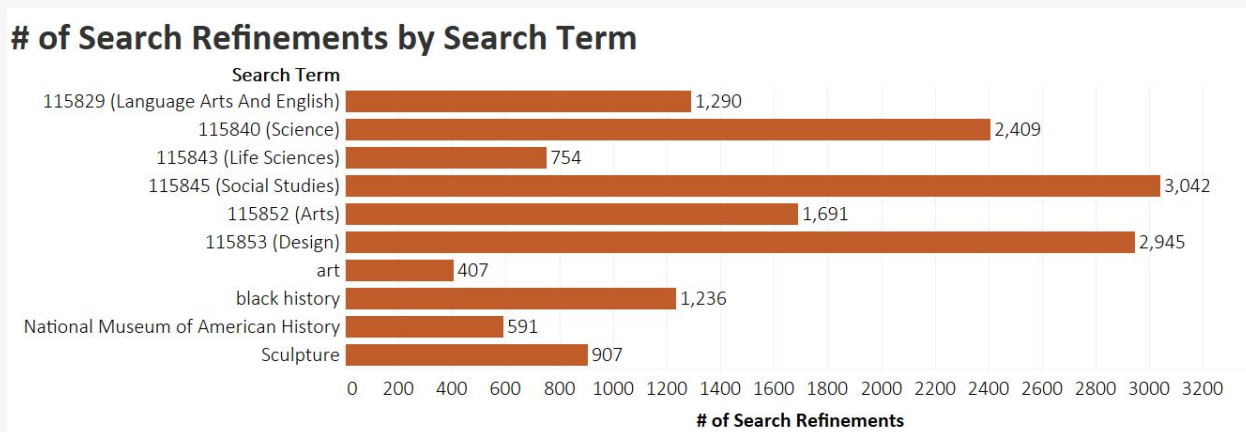

Fig. 14: Number of Search Refinements by Search Term.

5. Key Findings

5.1.4 Ethnographic Analysis of Search Tool Users

In analyzing the diverse ethnographic backgrounds of Smithsonian Learning Lab users and their ability to achieve successful searches, we can familiarize ourselves with their behaviors to better meet their needs. The findings are based upon the metrics of country, percent of sessions with a search, language, time after search, results pageviews/search, and the percentage of search exits.

The pie chart in Fig. 15 displays percentages of sessions with search in the countries with the highest amounts of unique searches. **This chart shows that the search tool is most frequently used in the United States, Canada, United Kingdom, India, Australia, Mexico, Germany, and France**. Other countries average around 9% of sessions with search.
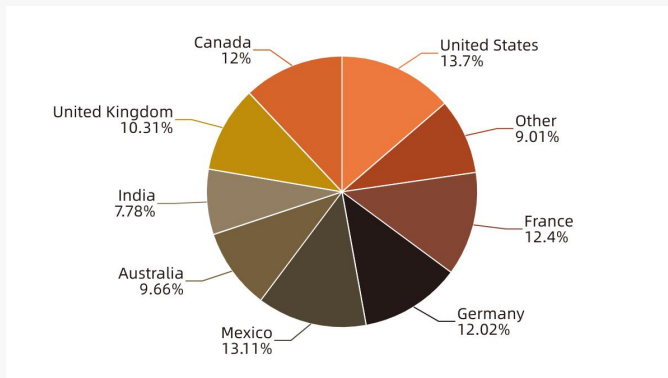

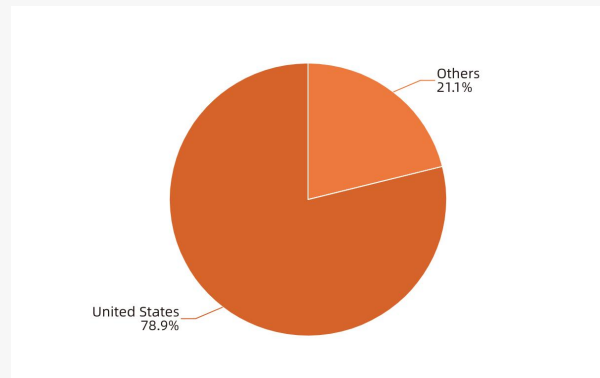Fig. 15: Percentage of Sessions with Search By Country.

Fig. 16: US Users vs. Other Countries.

The findings on the percentage of sessions with a search are of course relative to the number of users from each country. Fig. 15 demonstrates that the search tool is evenly used among top countries, even with the great disparity in users between the US and other countries shown in Fig. 16.

Furthermore, we must take into account language barriers in the use and access to the Smithsonian Learning Lab search tool. As the country dimension above shows, it is mainly English-speaking audiences that currently use the Lab. It is also very important to understand that Google Analytics uses the website browser's settings of language to represent the user, which does not necessarily reflect the true language the user speaks.
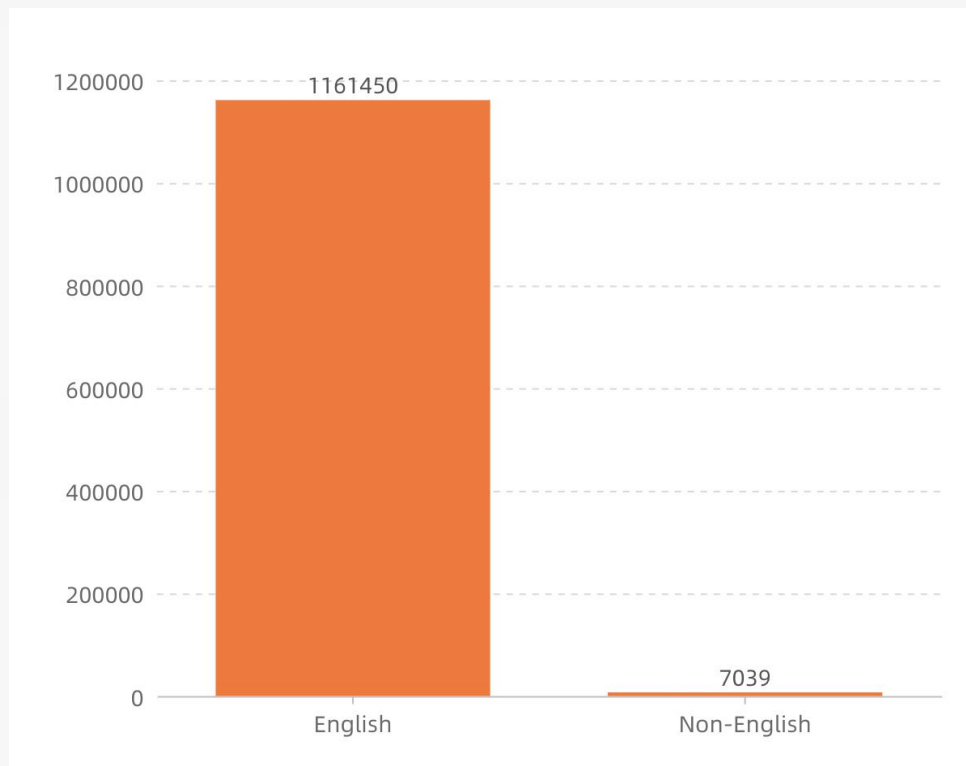
5. Key Findings

Fig. 17: English vs. Non-English Users.

From what the analytics tells us, only a small portion of users currently have their browsers set in another language. While users may speak a first language that is not English, their browser only represents them in the English-speaking group. Removing assumptions about people's ethnographic backgrounds is important for making a website and work practices ethical to its users. Allowing online spaces for people who do not speak English, or who do speak English as a second language, promotes accessibility globally.

Finally, we sorted the languages with the most time after a search by those with the most users, so that the Smithsonian Learning Lab team can prioritize which languages are the best candidates for translation to meet user needs.

5. Key Findings

LANGUAGE	USERS	TIME AFTER SEARCH	RESULTS PAGEVIEWS AFTER SEARCH
RUSSIAN	4,069	07:30	25
FRENCH	2,596	20:00	44
DUTCH	594	07:06	7
INDONESIAN	579	08:10	1
POLISH	552	16:48	4
GERMAN	175	06:32	3
ARABIC	18	12:43	17

Fig. 18: Number of Users, Time After Search, and Results Pageviews After Search With Search Terms.

Distributing a survey about user motivations that is linked to the search terms will support greater insight into user behavior in non-English speaking groups. We also recommended further research into the language of Lab users compared with the metrics of time after search, results in pageviews/search, and % of search exits. This deeper investigation could help us determine the usability of the search tool to speakers of other languages.

5. Key Findings

5.2 Search Results Findings

5.2.1. Traffic Sources and Search Results Pageviews

In order to better interpret behavior on the Lab site, the source of traffic needs to be identified. This metric helps establish the start of a user's journey on the site and gives insight into their intent. **In *Figure 19*, it can be seen that only 13% of sessions on the site that utilize search are direct traffic.** From GA, it is evident that 87% of the traffic to the site is indirect; these users are brought to the site via mainly search engine result pages (SERPs). It can be assumed that this indirect traffic does not consist of exploratory users. Based on this conjecture, we can determine that the aim for these users is to find answers to their questions quickly and thus should have a low number of result pageviews per search.

MEDIUM	SESSSIONS	%SESSIONS WITH SEARCH
(NONE)	719,395	13.05%
REFERRAL	646,942	19.27%
ORGANIC	629,986	13.14%
EMAIL	24,496	8.67%

Fig. 19: Medium of Sessions and Percentage of Sessions with Search.

By analyzing the behavior flow of indirect traffic in *Figure 20*, it is evident that almost 50% of users who organically end up on the search and search results pages end up dropping off and abandoning the search, while the other 50% continue their exploration. The first interaction that users who continue pursuing in their search are exploring collection and resource pages. The drop-off percentages at this interaction is much lower, as most users tend to explore at least one resource or collection before exiting. The largest percentage of users on their second interaction end up conducting a second search or search refinement.

5. Key Findings

Fig. 20: Behavior Flow for Sessions with Search.

The current result pageviews per search is over 3, meaning that users are viewing at least 3 resource or collection pages on the Lab search results page before finding content that they are looking for or abandoning the search. This statistic could indicate that the organization of search results on the site is not optimal for users seeking specific content. This could also mean that many irrelevant search results are being displayed and disrupting the search experience.

5. Key Findings

5.2.2. Search Results Accuracy

We selected the top 10 search terms according to the amount of total unique searches. "Black History" and "Hispanic Heritage" are the most searched terms, with 6,198 and 3,080 unique searches respectively. The number of unique searches of other terms is all below 1,000. On average, users stay for 3 minutes and 52 seconds after a search.

Top 10 Search Terms

Search Term	Total Unique Searches	Time after Search
black history	6,198	00:03:43
hispanic heritage	3,080	00:02:16
art	849	00:04:13
space	768	00:03:40
womens history	611	00:01:08
slavery	610	00:08:04
black history month	599	00:03:07
National Museum of American History	599	00:05:26
trail of tears	595	00:03:12
civil war	577	00:03:58

Fig. 21: Total Unique Searches and Time After Search with Search Terms.

In order to further study whether current search results meet users' expectations, we analyzed the result pages of the top 10 search terms and calculated metrics such as recall results, precision rate, and Avg. % Relevant on 1st page of these terms. These metrics are defined as follows:

- Shown results: If the search term turns up results, enter 1, otherwise, 0.
- Precision rate: The percentage of collections in the search result set that is relevant to the search term.
- % Relevant on first page: The percentage of collections on the first result page set that is relevant to the search term.

We found that the top 10 search terms all successfully recalled contents. The average precision rate is 70%, and the average percent relevant on the first page is 92%. Compared with search terms with a large number of collections, search terms with a small number of collections perform more accurately and better.

5. Key Findings

In addition, collections' titles including the search term is more likely to be recalled. An interesting fact is that the average time after search with terms that have low relevance result page is 319 seconds, while that of other terms is 195 seconds.

SEARCH TERM	SHOWN RESULTS	RELEVANCE	% RELEVANT ON 1ST PG
BLACK HISTORY	1(145)	YES	100%
HISPANIC HERITAGE	1(26)	YES	100%
ART	1(4341)	NO	75%
SPACE	1(495)	NO	66.7%
WOMENS HISTORY	1(2)	YES	100%
SLAVERY	1(144)	NO	87.5%
BLACK HISTORY MONTH	1(25)	YES	100%
NATIONAL MUSEUM OF AMERICAN HISTORY	1(335)	YES	100%
TRAIL OF TEARS	1(10)	YES	100%
CIVIL WAR	1(226)	YES	100%

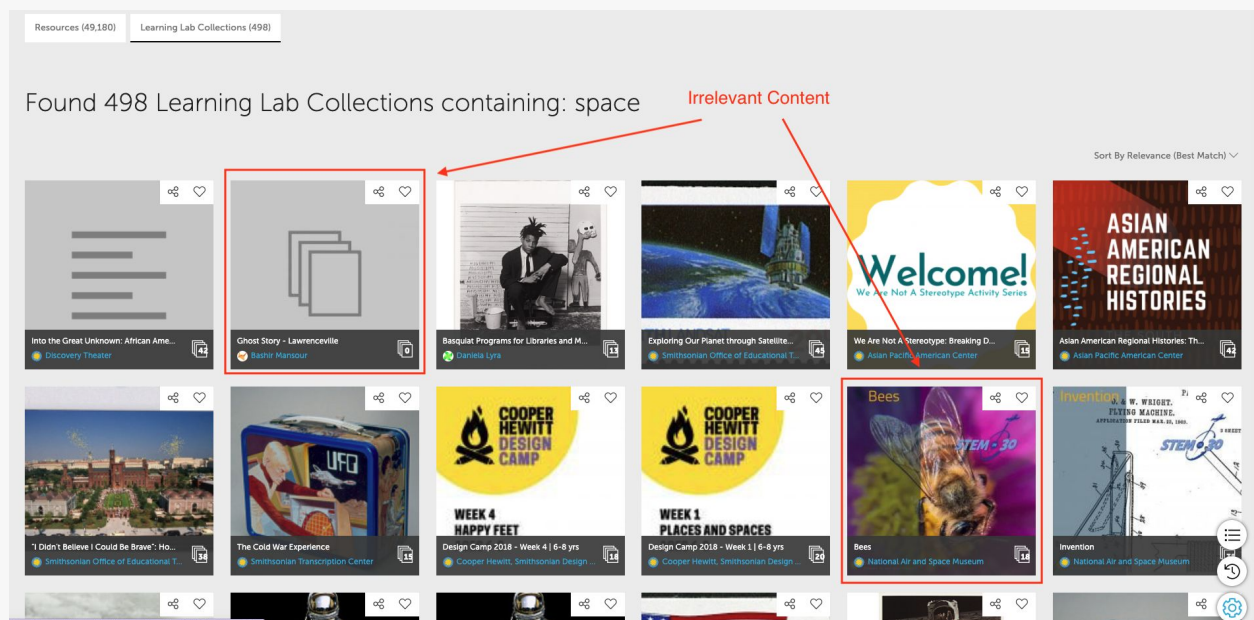*Fig. 22: Shown Results, Relevance, and Percent Relevant on First Page with Search Term.*

Fig. 23: Results Page of Search Term "space."

5. Key Findings

5.2.3. Search Results Layout Effectiveness

Besides accuracy, usability is a major contributing factor to the experience of a results-driven user on the Smithsonian Learning Lab. It is difficult to come to conclusions with respect to usability based on analytics alone because of the lack of feedback from the users. According to the analytics of the percentage of users who arrive at a search results page, however, **less than 50% are utilizing search refining filters**. Given that there are millions of resources within the Lab's database, the search and refining process could be very arduous for those not utilizing these features. Sorting also seems to lack prominence on the site as less than 1% of users are using any type of sorting as shown in *Figure 24*. Of the sort filters, those related to "Date Modified" are the most popular, with around 0.99% sorting by most recent modification and 0.5% sorting by oldest modification. Apart from how many users are utilizing these UI features that are being tracked, some qualitative findings are also apparent after exploration of the search results page.

Top 10 Search Terms		
Search Term	**Total Unique Searches**	**Time after Search**
black history	6,198	00:03:43
hispanic heritage	3,080	00:02:16
art	849	00:04:13
space	768	00:03:40
womens history	611	00:01:08
slavery	610	00:08:04
black history month	599	00:03:07
National Museum of American History	599	00:05:26
trail of tears	595	00:03:12
civil war	577	00:03:58

Fig. 24: Usage of Sorting Filters.

By conducting a cursory heuristic evaluation of the search results page, some findings related to the layout's effectiveness become apparent.

Lack of visibility in metadata without changing the layout or hovering over resources and collections significantly reduces the scannability of the page. Although there are alternative layout options including list view, grid view, and grid view with metadata, because the default layout option is just

5. Key Findings

image-based, users are forced to hover over each result to identify what it is. This becomes an even larger issue with collections since an image alone does not give the user any context into what other resources may be included in the collection.

The prominence of browsing features such as search refinement also impacts the effectiveness of the current search layout. Having search filters hidden by default forces the user to take an extra step in order to refine their search. Given that most users are clicking through more than one result after their search, empowering them to utilize refinements can make their search process more fruitful and efficient. Given that not all search filter categories are visible at once due to the "Subject" filter being expanded, users are not immediately aware of all ways in which they can refine their search.
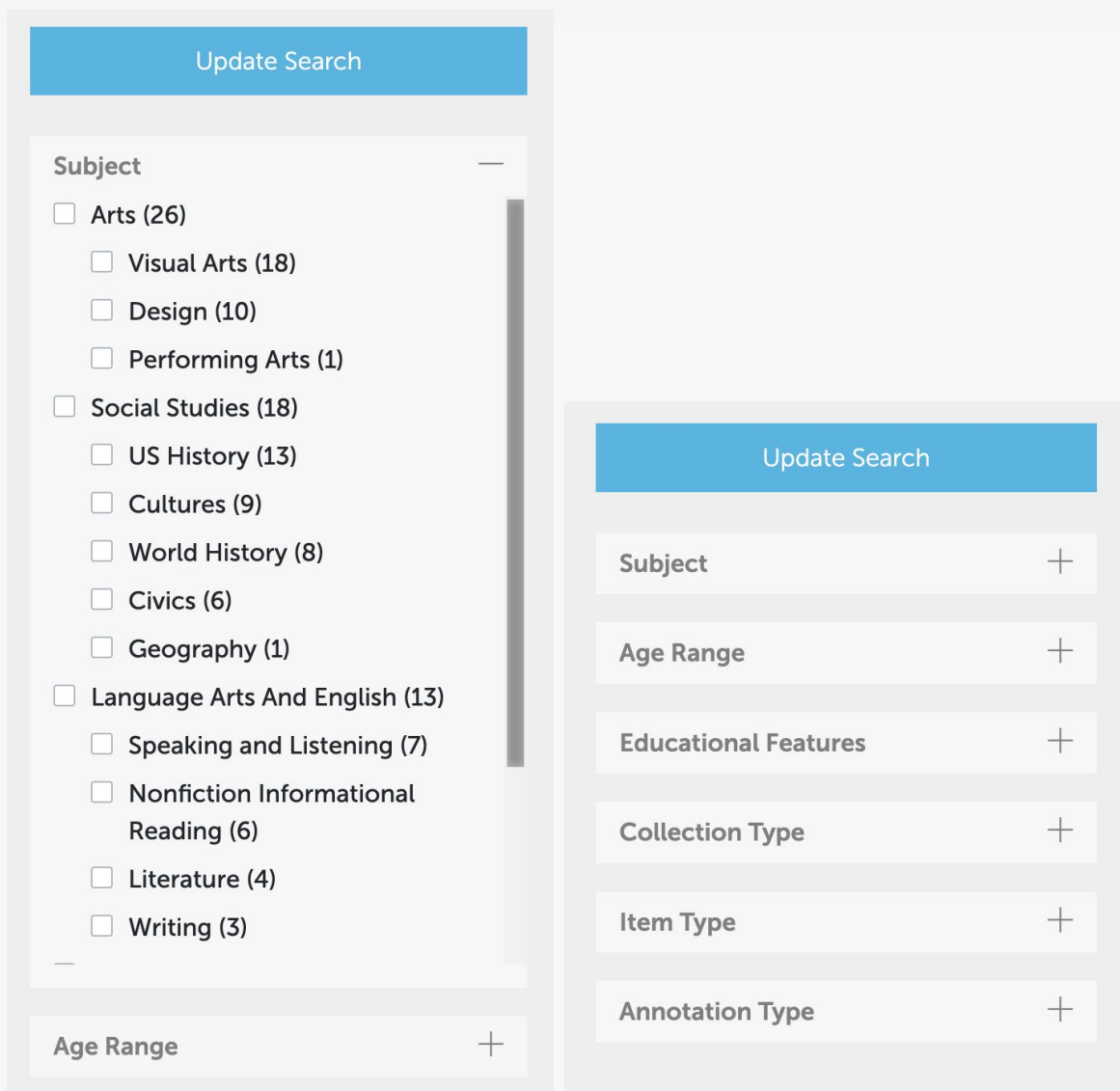
Fig. 25: Default Filter Preview vs Collapsed Filter Preview.

5. Key Findings

Once a search is refined using said filters, there appears to be a lack of visual distinction between tabs (resources, collections) and applied filters. Given that both are indicated by text in white rectangles, users may not realize that the first two are navigational elements whereas any following ones would be refinements applied to their search. The lack of clear visual distinction may prevent users from seeing collections pertaining to their search. In *Figure 26*, it is evident that the only distinction between tabs is an underline visually indicating the current page.
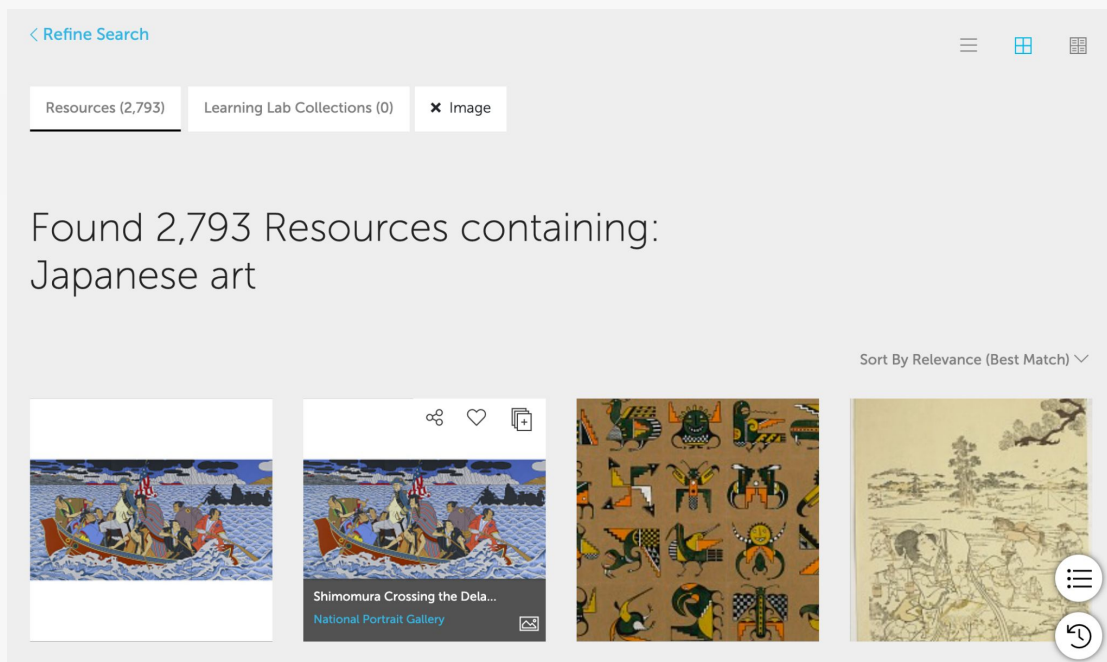
Fig. 26: Search Results Page for "Japanese Art" Filtered by Images.

6. Recommendations

6.1. Search Behavior Recommendations

6.1.1. Simplifying the Search Process

It is apparent that the audience of the site consists of both exploratory and results-driven users. In order to improve the experience for both user types, the structure and layout of the home page can be updated to make search and browsing simpler. **Since the primary function of this site is being able to search, the search bar needs to be accessible and easy to find. Although it is currently located at the top of the home page, due to the fact that there are four different search bars on this page, users may be confused as to which to use. It is recommended to consolidate the search to one search bar.** As seen in *Figure 27*, the consolidated search bar will have the ability to narrow search results by allowing the user to select what they are searching for from a drop-down.

Fig. 27: Redesign of Smithsonian Learning Lab Home Page.

While this makes search simpler, users who are looking to explore existing resources more freely need a starting point to their journey. Highlighting resources and collections can support this process.

6. Recommendations

6.1. Search Behavior Recommendations

6.1.2. Featuring Popular Collections

Based on our analysis of search activity and search terms (findings 5.1.1. and 5.1.2.a.), it is clear that users of the Smithsonian Learning Lab often turn to the Lab's search function to find materials related to cultural heritage months. **In order to give users what they're looking for and drive increased traffic to the Lab's search, we propose the creation of a module on the Smithsonian Learning Lab home page that showcases featured collections for users.** This feature will benefit exploratory users who may have a topic in mind but would enjoy browsing different suggested collections.

This module could be updated every season, each month, or at selected points throughout the year (whichever model the Lab staff determines is most feasible for existing time and resources). An optimal location for this module would be below the "Discover" search bar, as this is the first search bar users encounter when scrolling through the website.

No matter the update frequency, special attention should be paid to including content related to cultural heritage months. Findings 5.1.1. and 5.1.2.a. make it clear that February is a key month for Smithsonian Learning Lab searches and that users are interested in Black History Month content, so using the proposed module to feature collections related to Black History Month in February each year should be a priority. Other popular heritage months users search include Hispanic Heritage Month and Women's History Month. Additional cultural heritage months to consider featuring in the proposed module include Asian American and Pacific Islander Heritage Month, Native American Heritage Month, and LGBTQ Pride Month.
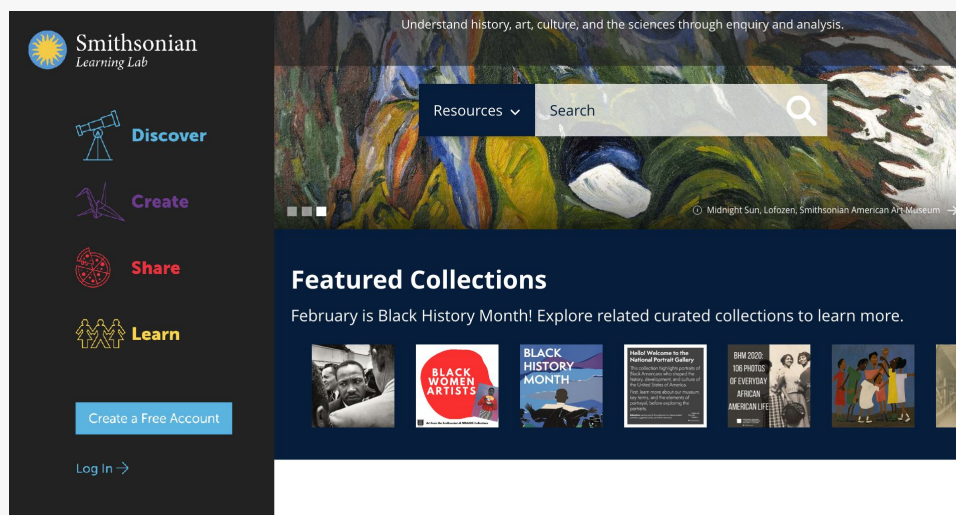

Fig. 28: Mockup of Proposed Featured Collections Module.

6. Recommendations

6.1. Search Behavior Recommendations

6.1.3. Featuring Resources For Teachers

As shown in finding 5.1.3.b.,10,080 unique page views during the period analyzed involved usage of the Lab's age level refinement filters. Educators find these age level filters useful for accessing collections about their students' age groups. **To best serve this segment of the Lab's user base, we propose the creation of a module on the home page that would allow immediate access to collections filtered by the appropriate age level.** This module would be modeled after and would ideally be located directly below the featured collections module. Since this location is so close to the first search bar users encounter on the Lab's home page, it is the perfect spot for such a module.

This feature would benefit results-driven users, particularly educators, who need to refine their searches quickly and efficiently to access age-appropriate content for their students. Since the featured collections module is geared towards exploratory users and the age level module is geared towards results-driven users, situating these modules next to each other provides search solutions for both motivations.
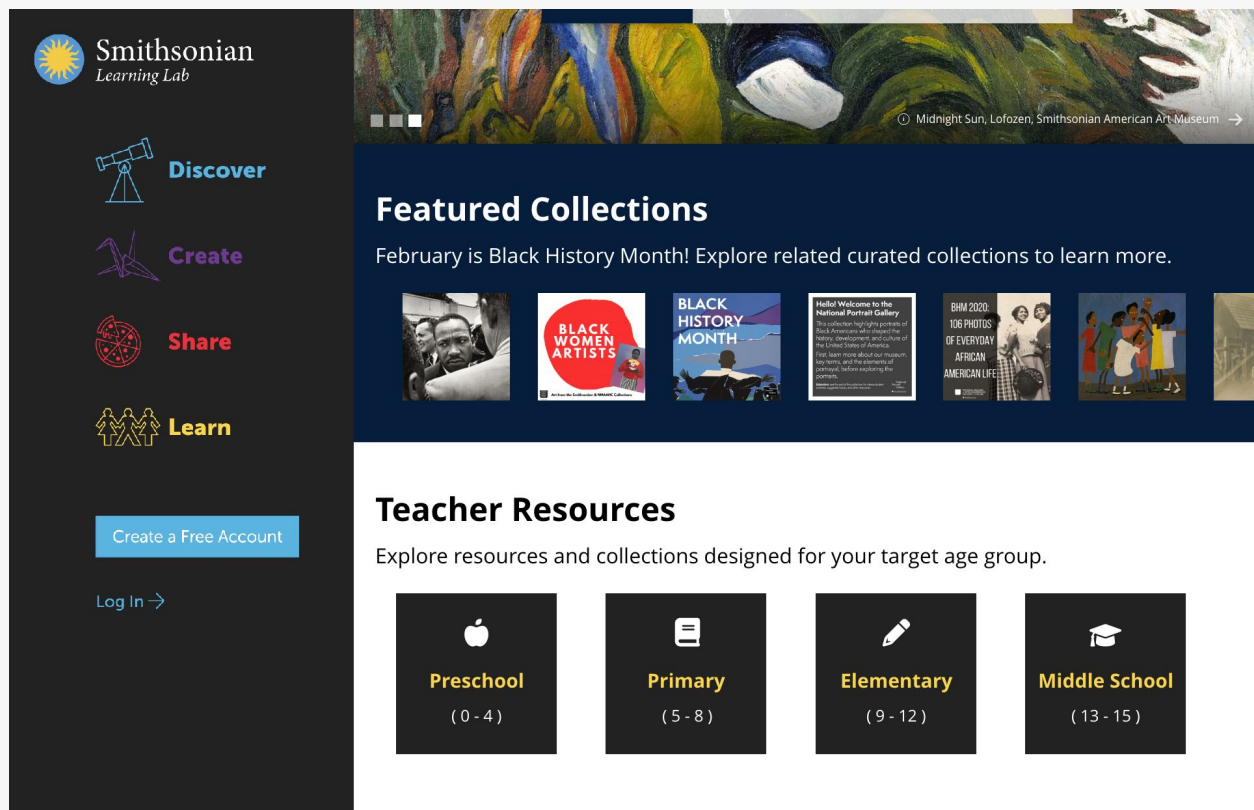

Fig. 29: Mockup of Proposed Age Level Module.

6. Recommendations

6.1. Search Behavior Recommendations

6.1.4. Make Search Refinement Easier

The Lab houses an incredible number of resources and collections. **To support users in their search for specific content, we recommend a redesign of the search results page,** as shown in *Figure 30*. Many aspects of the page that support search refinement could be improved.
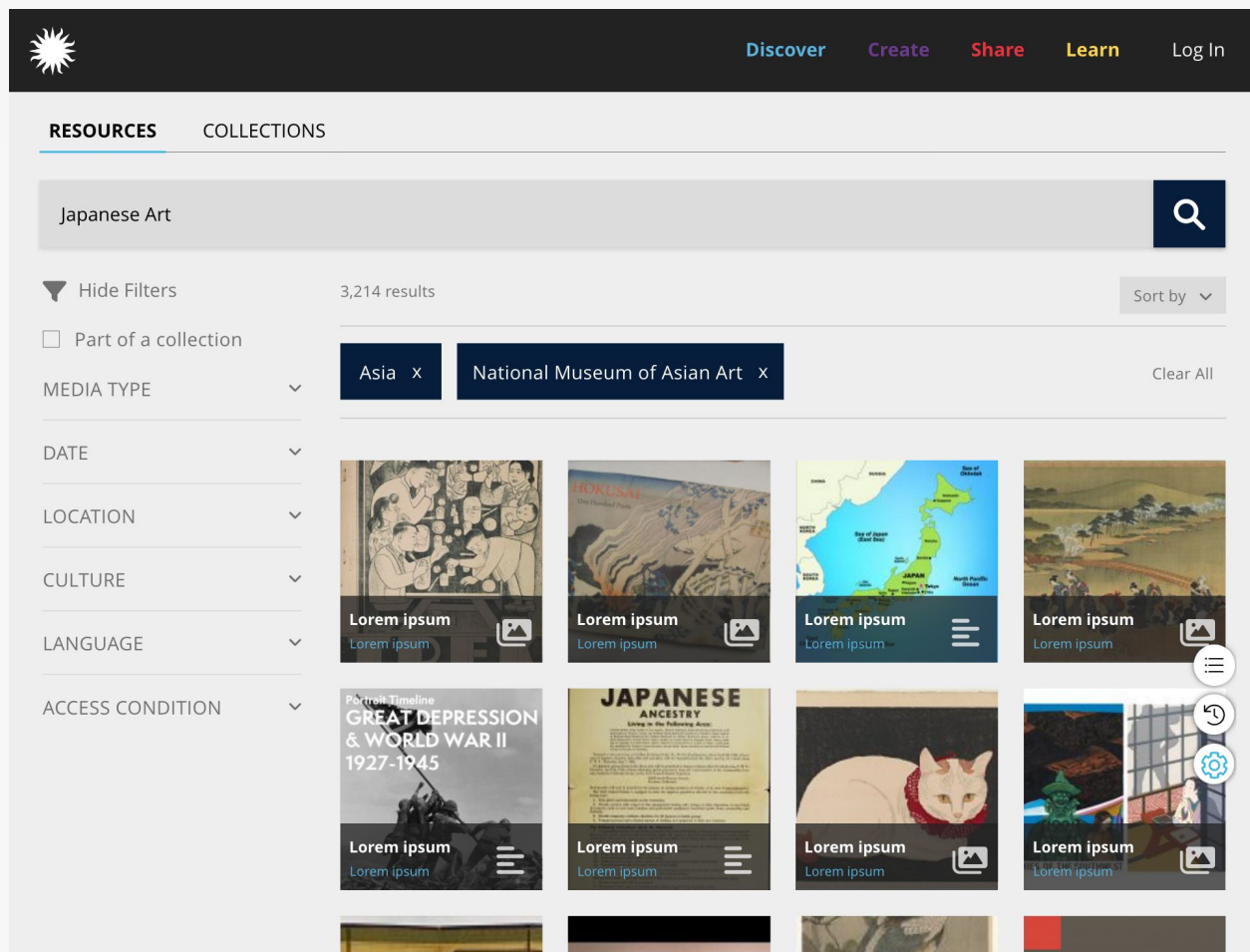

Fig. 30: Mockup of Search Results Page Redesign.

Overall, the hierarchy of the page could be restructured to reflect expected browsing behavior. Retaining the navigation bar on the search results page and moving the logo to the left-hand side to match the home page not only allows for free navigation through the site but also reduces the cognitive load for the user to find the home page. Since the tab selected (resources or collections) is the limiting factor in executing a search, it makes the most sense to visually indicate which tab is selected and move all tabs above the search bar. Giving these tabs strong visual contrast removes the need to repeat what sort of content is in the results. Being that entering a search term is a vital aspect

6. Recommendations

of this page, having the bar span the entire width of the page gives it the prominence it deserves. While the number of results that are produced from the search is an important measure of recall, it does not need to take up so much screen real estate.

Having filters displayed by default encourages search refinement. Collapsing filter categories will increase the visibility of options and paths for search narrowing. We also recommend reordering filter categories by priority and frequency of use in addition to including a filter to identify whether a resource is part of a collection, as this could indicate the value, trustworthiness, or popularity of a resource.
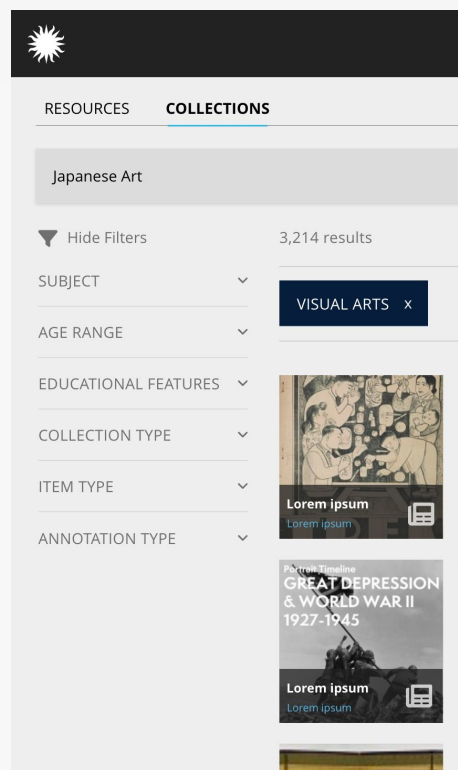
Fig. 31: Mockup of Search Page Collections Tab Redesign.

Since utilizing filters improves the search experience, it needs to be evident to the user which filters they have already applied. Making the filter tags visually different from tabs make it clear to the user how they have narrowed their search. Including ways to remove the filters, both one-by-one or by clearing all, allows them to have more freedom in their search process.

6. Recommendations

When looking through resources or collections, since just imagery is not sufficient to identify and determine a user's interest in content, metadata including content title and authors should be displayed by default. Though inspiration seekers may be content just browsing through image thumbnails, the primary audience of the Lab site is more likely to be curious about content details upfront. This being said, actions such as favoriting or adding to a collection could be hidden until a user hovers over a piece of content as shown in *Figure 32*.
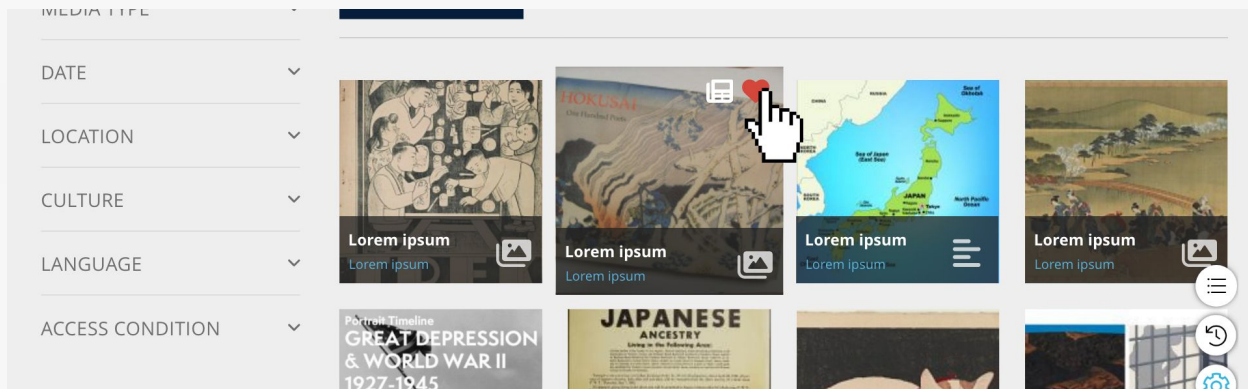
Fig. 32: Mockup of Search Result Interaction.

6. Recommendations

6.2. Search Results Recommendations

6.2.1. Optimizing Recall Strategy and Results Layout

In terms of resource recall, organizing resources by scores weighted by the popularity and the relevance of the resource will be more efficient to fetch information that users would like to see on the first page of results. In addition, search terms with typos may not be able to successfully recall all the relevant content. **In** the case of a small number of resources, we suggest **providing a message that reads "Are you looking for (corrected search term)" on the result page and including relevant content.** Meanwhile, to help exploratory users discover related content while looking at a specific resource, we suggest **adding a tab to the left navigation from which users can view resources and collections with similar topics or keywords.** This will reduce unnecessary steps and make content more intuitive. Finally, we suggest **conducting user research to find out why users stay longer in terms with low relevance than those with high relevance**. Research could include testing the page dwell time of new, different layout styles to see which layout style should be implemented to the Lab.
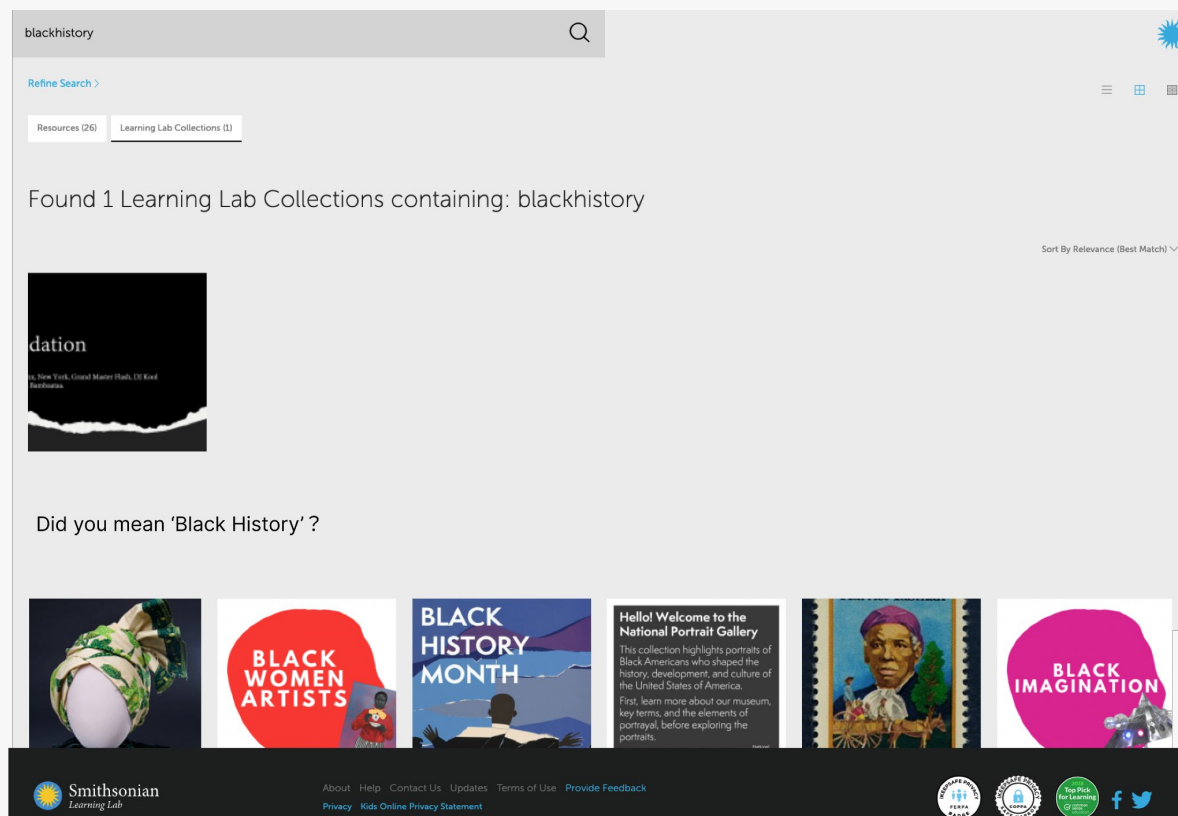
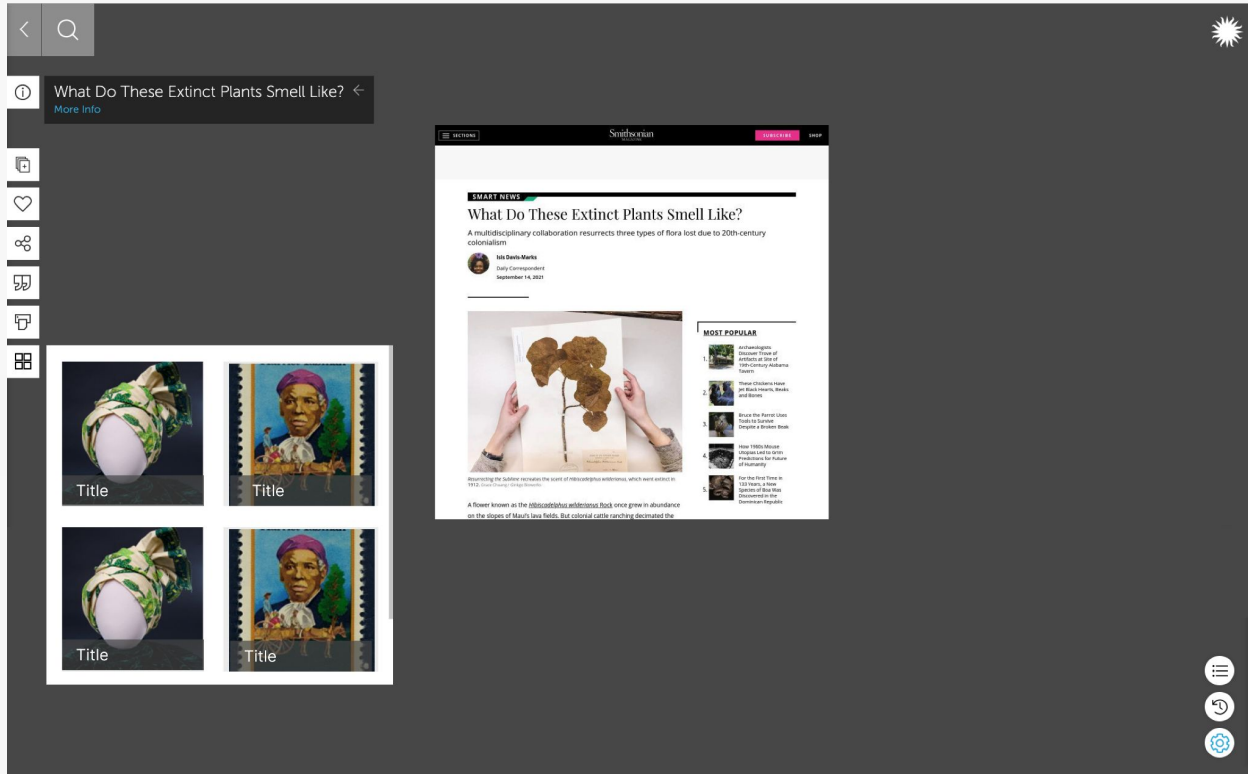*Fig. 33: Mockup of "Did You Mean..?" Example After Search.*

6. Recommendations

Fig. 34: Mockup of New Linked Icon on Left-Hand Side of Item Page to Show Related Collections and Other Linked Data.

6. Recommendations

6.2. Search Results Recommendations

6.2.2. Connecting Collection Keywords

The Smithsonian Learning Lab recently launched a new Keywords tool that allows users to add searchable terms to their collections, using an unrestricted text input field, to increase the collection's potential for discovery by other users.
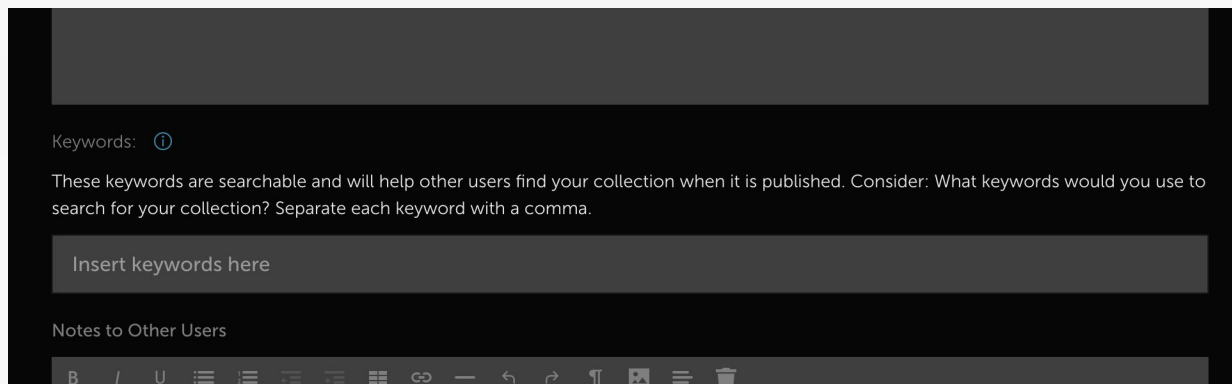

Fig. 35: Screenshot of Current Keyword-Tagging Feature for Collections.

To increase the discoverability of collections and increase the recall and precision rates of collection searches, the team should consider building a controlled vocabulary into the Keywords tool. This controlled vocabulary should be populated from the most frequently searched terms and updated every month with existing development releases. This change will regulate the terms used to tag collections, ensure that collections are tagged with the most common words users search with, prevent tagging typos, and provide collection-creators with a starting place to brainstorm keywords. By building this into the process of user-creation of collections, this change would also save the Lab team time and effort in regulating collection keywords and increasing the recall rate and precision rate of searches themselves, like the recent change to require users to add alt text when uploading resources to the Lab.

Keywords added through this tool should be made visible to collection viewers when they click the Information button. They should also be hyperlinked to searches for the keyword term—like resource keywords already are—creating opportunities for further exploration and discovery.

When users search for terms to add to their collections as keywords, consider displaying the number of other collections that also contain this keyword after the term—a technique used by the blogging platform Medium—to assist users in the keyword selection process.

6. Recommendations

Consider, too, allowing users to type their own unique words into the field to submit them for review by the Lab team and inclusion in the larger controlled vocabulary. This will enable the Lab team to be quickly responsive to user needs, and could be built into the existing processes used to manage Terms of Use reviews of published collections.

6. Recommendations

6.2. Search Results Recommendations

6.2.3. Encouraging Exploration

Building on the possibilities for successful search results and reaching wider audiences, we recommend adding a "Part of a Collection" refinement feature to resource searching as a "yes or no" option. This will allow users to view assets that have been added to collections and have been deemed useful by other users in the past. It is important to note that the Lab currently has the option to see if the resource has been added to a collection, but only when viewing the resource in detail. Allowing the user to search within collected items alone creates a new level of knowledge networking that would better benefit students and teachers to fit their projects and curriculum. Adding this as a search refinement for searching collections would be useful for cross-referencing assets between collections as well.

Cross-Collection tagging is another important feature we recommend adding because it can help lead the user to more relevant material. Allowing collection creators to cross-tag collections enables gaps in access from the search tool to be filled by users and educators. In this way, users are helping Smithsonian Learning Lab do the work of mapping assets in relevancy by including topics of collections and other assets in the object description. This will allow the site's search engine to generate even more relevant searches from the repetitive topics and keywords used in these popularly saved items. Furthermore, cross-collection tagging opens up future opportunities to organize collections on the collections search page by popularity, with favorited items and analysis of clicks to understand third-party use of one's created collection.Creating a way for similar collections and similar items to be displayed on an item page will allow for easier exploration. This generates predictions on what a user would like to view next.

Using Linked Open Data in sources such as Wikidata will increase the discoverability of a museum asset through Google and other search sites, increase Search Engine Optimization, allow other websites that are relevant to the item to be connected to the Lab's resource pages, and enable deeper connections to discussions of an object on the web. Keep in mind that these steps towards SEO are also directly relevant to meeting users' needs in other countries and languages by making the item more findable. See Fig. 36 on how Smithsonian Learning Lab could connect their assets to current discussions regarding relevant people, places, and things, allowing for the Smithsonian to be one main point for information regarding these objects.

6. Recommendations

Fig. 36: Screenshot of Wikidata Linked Open Data Entry and Smithsonian Learning Lab Item "Amelia Earhart" (See Resources Appendix).

6. Recommendations

6.2. Search Results Recommendations

6.2.4. Call for Smithsonian-Created Collections

Using findings on search result accuracy and the top search terms, categories, words, and filters (Sections 5.2.2. and 5.1.2.), **develop a call for content from Smithsonian educators for new Smithsonian collections**. This approach will enable the creation of new educational resources based on user needs. It is important to note that, given the findings on search refinements (Section 5.1.3.), Smithsonian-created collections in the top subject areas will likely facilitate further exploration by users who begin their searches hoping to learn more and find educational content within specific subject areas.

Preliminary focus areas, based on findings from this report's analysis, include:

SUBJECT AREAS	AGE LEVELS	HERITAGE MONTHS	TOPICS AND TERMS
ARTS (especially Design, Visual Arts, and Music)	**#1: ELEMENTARY** (9 to 12 years old)	**BLACK HISTORY MONTH**	**#1: ASTRONOMY** (including space, moon, solar system, etc.)
SOCIAL STUDIES (especially US History, World History, and Archaeology)	**#2: PRIMARY** (5 to 8 years old)	**HISPANIC HERITAGE MONTH**	**#2: INDIVIDUALS** (such as Rosa Parks, Martin Luther King Jr., etc.)
SCIENCE (especially Life Sciences, Astronomy, Biology, and Chemistry)	**#3: MIDDLE SCHOOL** (13 to 15 years old)	**ASIAN PACIFIC AMERICAN HERITAGE MONTH**	**#3: COLONIAL AMERICA** (including American Revolution, Jamestown, etc.)
LANGUAGE ARTS AND ENGLISH (especially Nonfiction Informational Reading)	**#4: HIGH SCHOOL** (16 to 18 years old)	**AMERICAN INDIAN HERITAGE MONTH**	**#4: NATIVE AMERICANS** (including Trail of Tears, Removal Act, etc.)
	#5: PRESCHOOL (0 to 4 years old)	**WOMEN'S HISTORY MONTH**	**#5: ANCIENT HISTORY** (including Greece, Egypt, Rome etc.)
	#6: POST-SECONDARY		**#6: WOMEN** (including suffrage, etc.)
	#7: ADULTS		**#7: PLACES** (Egypt, Alaska, Japan, Africa etc.)
			#8: AFRICAN AMERICAN (including slavery, Civil rights, etc.)

6. Recommendations

6.1. Search Results Recommendations

6.2.5. Searching for Museums

As the team considers how best to feature collections created by Smithsonian museums, the team should consider that the majority of users looking for museum-related content are using typed terms—rather than the Lab's search filters—to do so. Users are also attempting to search for content related to museums external to the Smithsonian, which do have content in the Lab, the same way.

The team should consider making users and groups marked as white-listed in the Lab's administrative back-end searchable in the Lab, just as titles, descriptions, and keywords can currently be searched. The controlled vocabulary is recommended in Section 6.2.2. may support this solution; consider pre-populating the keyword fields in collections created by white-listed users with the user or group name. Because these solutions will only increase the discoverability of collections, making these user and group names searchable should exist in tandem with other solutions for enabling easy browsing of existing institutional profiles, such as featuring profiles on the homepage or creating a third tab on the search results page that displays profile results.

6. Recommendations

6.1. Search Results Recommendations

6.2.6. Improving Language and Accessibility

A key recommendation for the search tool redesign suggested above is to add language refinements to searches, allowing for accessibility among ESL and non-English speakers in their search. Furthermore, inclusivity can be enhanced in collection access and use by allowing translations from multiple languages to describe each object. This is especially useful for the articles that are published only in English. Having translations of these would allow for the site search to be used and resources to be added to collections created by speakers of multiple languages. With India, Mexico, Germany, and France as the top countries for their high percentages of sessions with a search, we recommend translations to **Hindi, Spanish, German, and French** as considerations for beginning translations.
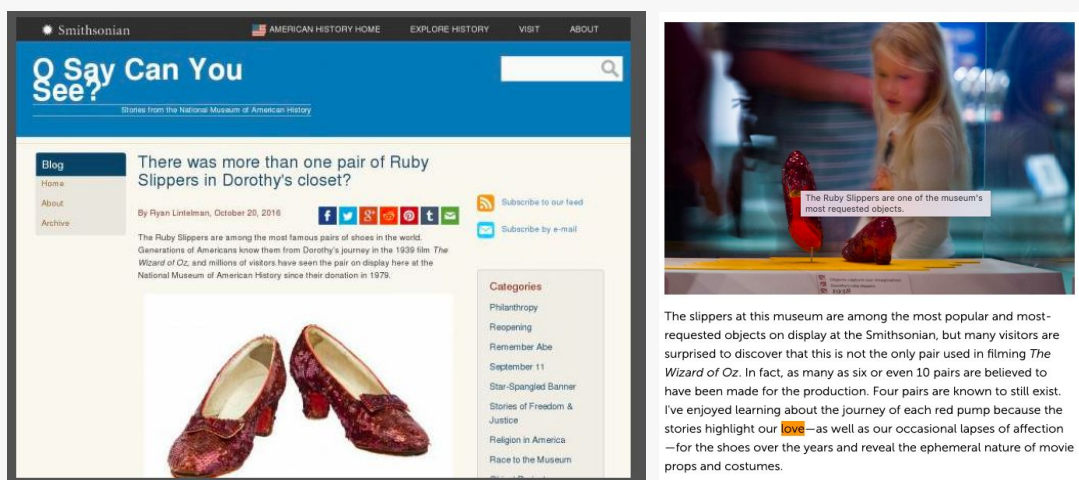

Fig. 37: Screenshot of Article Type Item From Smithsonian Learning Lab (see Resources Appendix).

A further recommendation that will require more research from user surveys is not just to meet the needs of the Lab's current largest user audiences worldwide, but the ones that seem to be having the hardest time finding successful results. The languages with the most users and the languages with the highest pageview time after search are **French and Russian** and thus are highly recommended for translation inclusion. Other languages in smaller user groups but with long amounts of time after search to take note of are **Polish, Arabic, Indonesian, and Dutch**. These are ordered in analysis from most time spent after a search, which can be interpreted in different ways, such as having the session open longer for research or struggling with translations. Each motivation behind time after the search tells a different story of user behavior. Either way, it shows use-case studies of people who speak these languages using the Lab.

6. Recommendations

Choosing which languages to prioritize in translation will come down to Smithsonian Learning Lab's personal goals in maintaining the current audience or reaching out to audiences that are currently lacking in support. Translating the website can also be a good tool for **Search Engine Optimization** by improving the user experience. The Lab may want to look into Google Website Translator, a more accurate version of the Google Translate AI, which has now been offered as a free widget for institutions since the start of the COVID-19 pandemic in 2020 (see Resources appendix).

Since this is Artificial Intelligence, it is important to create a **disclaimer** during the search tool redesign that tells audiences using a different language that there are possible errors in translation from the original English entry. When human resources are available, having curation oversight on the translations will be helpful. Once the description is added in translation, the site search tool will be able to search objects when that language is typed into the search.

Furthermore, Google searches will lead users from other languages to your site with this move for search engine optimization. This is a lot of work to do multiple translations but is well worth it since it meets Smithsonian Learning Lab's mission of the promotion of education across disciplines, creating educational networks that are long-lasting, and sharing knowledge and creations with a global community. Adding the language refinement in the search will be necessary to keep searches relevant to the user's query.

7. Conclusion

In analyzing the Smithsonian Learning Lab website through the lens of the search function, we took a two-pronged approach to define a successful search. Our analysis was modeled after two ideas of a successful search based on user motivations: exploratory or results-driven. Additionally, we segmented our analysis further by splitting our findings and recommendations into two categories: Search Behavior and Search Results. By segmenting our analysis in this manner, we aimed to provide the most in-depth findings and recommendations possible.

The project team brought multiple findings and recommendations into consideration for Smithsonian Learning Lab to optimize the tool use. In optimizing the ability of visitors to utilize the tool effectively, whether they are results-driven or exploratory, our recommendations align with user needs to achieve successful searches. In accordance with our findings, we put forth the following suggestions:

- Simplify the search process
- Feature popular collections
- Feature resources for educators
- Make search refinement easier
- Optimize results page layout
- Connect collection keywords
- Encourage explorations of connected resources
- Call for Smithsonian-created collections
- Enable searching for museums
- Improve language and accessibility

In conducting our analysis, we faced limitations related to a lack of data on true user motivations, end-use of content, and a lack of accurate Google Analytics goals currently in place for the Smithsonian Learning Lab.

Going forward, we recommend the Smithsonian Learning Lab transitions to Google Analytics 4, as Universal Analytics is set to phase out in 2023. We also recommend setting up strategic event triggers in Google Analytics in order to track key user behaviors. Furthermore, we believe conducting a usability study in order to further understand user behavior would be of great benefit to the Lab.

Some of our recommendations will take some serious rework of the website, and would call for a digital strategy to implement such changes. This will shed light on the need for human resources and a UX team for redesigning the search tool to allow for more refinements and translation of assets to multiple languages.

7. Conclusion

Looking forward, if a rework of the search tool is adopted, and the dashboard is used as a point of reference for the digital strategy of the Lab, future capacities could grow to include more data in tracking which users make a collection, the number of sessions (and session duration) vs the number of collections, and how many accounts are created.

By improving the Google Analytics usage within the project team of Smithsonian Learning Lab to reflect search terms, refinements, user behavior, and user ethnographic backgrounds, the Lab team can efficiently understand changes and growth over time as they rework the search tool. Putting this dashboard and other redesign recommendations into practice will ensure that Smithsonian Learning Lab continues to be a pioneering destination for digital resources and collections related to cultural heritage, science, art, museums, and education on a global scale.

8. Appendix

Images

Fig. 38: Site Search Overview of Smithsonian Learning Lab on Google Analytics.

Fig. 39: Table containing most popular search start pages from the Lab website on Google Analytics.

8. Appendix

Resources

"Amelia Earhart." *Smithsonian Learning Lab*, https://learninglab.si.edu/resources/view/131#more-info.

"Amelia Earhart." *Wikidata*, https://www.wikidata.org/wiki/Q3355.

Gu, XinXing. "Google Translate's Website Translator - Available for Non-Commercial Use." *Google*, Google, 2020, https://developers.google.com/search/blog/2020/05/google-translates-website-translator.

"Main Page." *Wikidata*, https://www.wikidata.org/wiki/Wikidata:Main_Page.

Lintelman, Ryan. "There Was More than One Pair of Ruby Slippers in Dorothy's Clo..." Smithsonian Learning Lab, 20 Oct. 2016, https://learninglab.si.edu/q/r/1353020.

"What Is SEO / Search Engine Optimization?" *Search Engine Land*, 18 Apr. 2022, https://searchengineland.com/guide/what-is-seo.

List of Figures

8. Appendix

Figure 9: Filter vs Non-Filter Searches for Museums.

Figure 10: Top 10 Subject Search Filters.

Figure 11: Screenshot of Smithsonian Learning Lab Home Page Where Subject Filters are Linked.

Figure 12: Age Search Filters.

Figure 13: Age Search Filters by Published Collections.

Figure 14: Number of Search Refinements by Search Term.

Figure 15: Percentage of Sessions with Search By Country.

Figure 16: US Users vs. Other Countries.

Figure 17: English vs. Non-English Users.

Figure 18: Number of Users, Time After Search, and Results Pageviews After Search with Search Terms.

Figure 19: Medium of Sessions and Percentage of Sessions with Search.

Figure 20: Behavior Flow for Sessions with Search.

Figure 21: Total Unique Searches and Time After Search with Search Terms.

Figure 22: Shown Results, Relevance, and Percent Relevant on First Page with Search Term.

Figure 23: Results Page of Search Term "space."

Figure 24: Usage of Sorting Filters.

Figure 25: Default Filter Preview vs Collapsed Filter Preview.

Figure 26: Search Results Page for "Japanese Art" Filtered by Images.

Figure 27: Redesign of Smithsonian Learning Lab Home Page.

Figure 28: Mockup of Proposed Featured Collections Module.

8. Appendix

Figure 29: Mockup of Proposed Age Level Module.

Figure 30: Mockup of Search Results Page Redesign.

Figure 31: Mockup of Search Page Collections Tab Redesign.

Figure 32: Mockup of Search Result Interaction.

Figure 33: Mockup of "Did You Mean..?" Example After Search.

Figure 34: Mockup of New Linked Icon on Left-Hand Side of Item Page to Show Related Collections and Other Linked Data.

Figure 35: Screenshot of Current Keyword-Tagging Feature for Collections.

Figure 36: Screenshot of Wikidata Linked Open Data Entry and Smithsonian Learning Lab Item "Amelia Earhart" (See Resources Appendix).

Figure 37: Screenshot of Article Type Item From Smithsonian Learning Lab (see Resources Appendix).

Figure 38: Site Search Overview of Smithsonian Learning Lab on Google Analytics.

Figure 39: Table containing most popular search start pages from the Lab website on Google Analytics.