A comprehensive evaluation of life sciences data resources reveals significant accessibility barriers

Sehi L'Yi[1], Harrison G. Zhang[2], Andrew P. Mar[1,3], Thomas C. Smits[1], Lawrence Weru[1], Sofía Rojas[1], Alexander Lex[4], Nils Gehlenborg[1]

[1]Harvard Medical School, MA, USA

[2]Stanford University, CA, USA

[3]University of California, Berkeley, CA, USA

[4]University of Utah, UT, USA

Abstract

Individuals with disabilities participate notably less in the scientific workforce. While the reasons for this discrepancy are multi-faceted, accessibility of knowledge is likely a factor. In the life sciences, digital resources play an important role in gaining new knowledge and conducting data-driven research. However, there is little data on how accessible essential life sciences resources are for people with disabilities. Our work is the first to comprehensively evaluate the accessibility of life sciences resources. We collected large-scale accessibility data about two essential resources: data portals (n=3,112) and journal websites (n=5,099). Our analysis shows that many life sciences resources contain severe accessibility issues (74.8% of data portals and 69.1% of journal websites) and are significantly less accessible than US government websites, which we used as a baseline. Focusing on visual impairment, we further evaluated three data portals in-depth with a blind user, unveiling the practical impact of the identified accessibility issues on common tasks (53.3% success rate), such as data discovery tasks. We believe that our data and analysis results bring new insights into how the scientific community can address critical accessibility barriers and increase awareness of accessibility, leading to more inclusive life sciences research and education. Our analysis results are publicly available at http://inscidar.org/.

Introduction

People with disabilities, such as vision, cognitive, and physical disabilities, encounter barriers in scientific research and education[1]. Researchers and scientific organizations highlighted the importance of addressing this problem and proposed strategies to address the current problems. For example, Swenor and Meeks[1] highlighted the need for a multifaceted approach, proposing 12 best practices. The Advisory Committee to the Director Working Group on Diversity (ACD WGD) of the National Institutes of Health (NIH) released a report[2] in 2022, highlighting the importance of data collection related to disabilities to diversify the life sciences workforce. The National Human Genome Research Institute (NHGRI) of NIH announced a long-term action agenda for the next decade[3] to increase diversity in the genomics workforce, which similarly emphasized the lack of data regarding people with disabilities in the workforce.

There is a discrepancy in the population of individuals with disabilities in the scientific workforce and the US adult population[4]. This indicates potential barriers for individuals with disabilities to join the workforce. There are many factors that are potentially associated with the current barriers, yet digital accessibility is considered one of the important factors[5]. In the life sciences workforce, digital resources, such as data portals and journal websites, play an important role in gaining new knowledge and conducting data-driven research[6]. However, there is little data on how accessible essential life sciences resources are for people with disabilities. For example, 45.2%[7] of people with visual impairments rely on screen reader assistive technologies, such as NVDA[8] and JAWS[9], to identify and understand contents displayed on the screen. However, multiple studies[10–15] found that existing digital resources, such as educational websites and PDF files, largely fail to support screen readers, making resources inaccessible to people with visual impairments. There are several studies that evaluated the accessibility of existing resources, such as university and government websites[10–14] or alternative text ("alt text") in PDF files[15]. However, such data is largely lacking for the life sciences. To include people with a broad range of disabilities in the life sciences workforce, it is vital to understand and address the digital accessibility issues of existing resources.

Our work is the first of its kind to comprehensively evaluate the accessibility of life sciences resources for people with disabilities, including vision, cognitive, and physical disabilities. We present a large-scale dataset that captures the real-world accessibility of life sciences resources—data portals (n=3,112) and journal websites (n=5,099)—using a computational accessibility testing tool. We also conduct a statistical comparison of our data on life sciences resources with US government websites (n=852); websites that need to meet strict legal requirements for accessibility (e.g., Section 508 of the Rehabilitation Act)[16]. Focusing on visual impairments, we also collected accessibility data of select data portals through manual evaluation with a screen reader user who has no residual vision, unveiling the practical impact of the identified accessibility issues on an individual trying to accomplish common tasks on data portals, such as data discovery tasks.

Results

Overall digital accessibility of life sciences websites

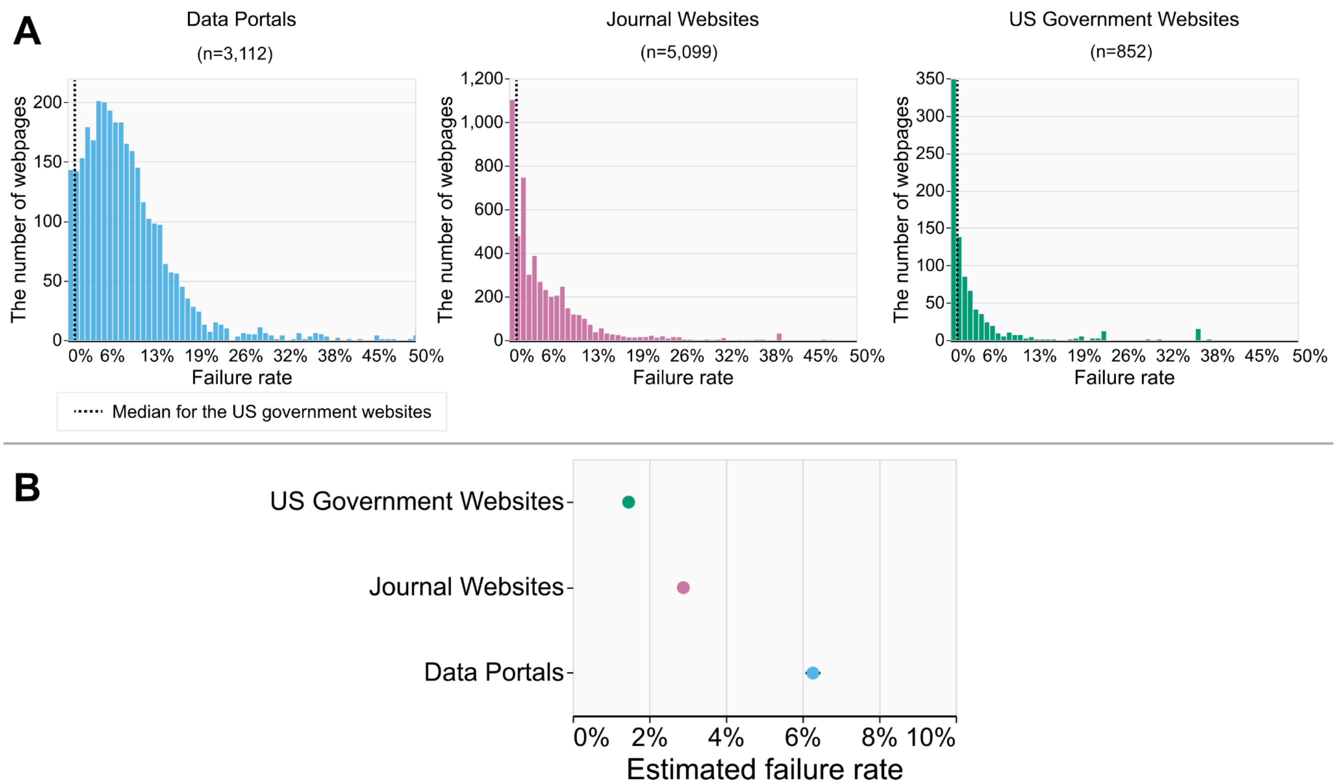

Fig. 1: The failure rates of data portals, journal websites, and US government websites. (A) The distribution of failure rates of individual web pages is shown in histograms. The y-axis represents the number of web pages with the corresponding failure rates (x-axis). The x-axis is truncated to 50% failure rates to make the distributions readable in the charts. There were 46 web pages with failure rates higher than 50% (32 data portals, 9 journals, and 5 US government websites). The median failure rate of the US government websites (1.1%) is drawn as dotted vertical lines for a standard reference. (B) The estimated failure rates of three groups of websites are shown in dot plots. The error bars represent the 95% confidence intervals. The failure rate is a metric that measures the accessibility issues of a website while considering its size. For example, a failure rate of 50% indicates that of all possible accessibility issues, a site actually fails on 50% of the issues (refer to the Methods section).

Our results show that life sciences websites largely fail to meet accessibility standards. The distribution of the failure rates of individual web pages (Fig. 1) shows that data portals are skewed the most toward the higher failure rates compared to the other two resources. Our statistical analysis results reveal that the failure rates of life sciences websites are significantly higher than those of US government websites. Government websites showed the lowest estimated failure rates (1.5%), followed by journal websites (2.9%). Data portals showed the highest estimated failure rates (6.3%), more than twice the rate of journal websites and four times the rate of government websites. Notably, almost all data portals (93.6%) showed higher failure rates than US government websites' median failure rates (dotted lines in Fig. 1). These include widely used data portals, including cBioPortal[17] (19.3%), KEGG[19] (3.7%), ENCODE[18] (2%).

Overall, life sciences websites significantly fail to meet accessibility standards, which leads to critical barriers for people with disabilities in accessing and using these essential resources for life sciences education and research.

Accessibility issues

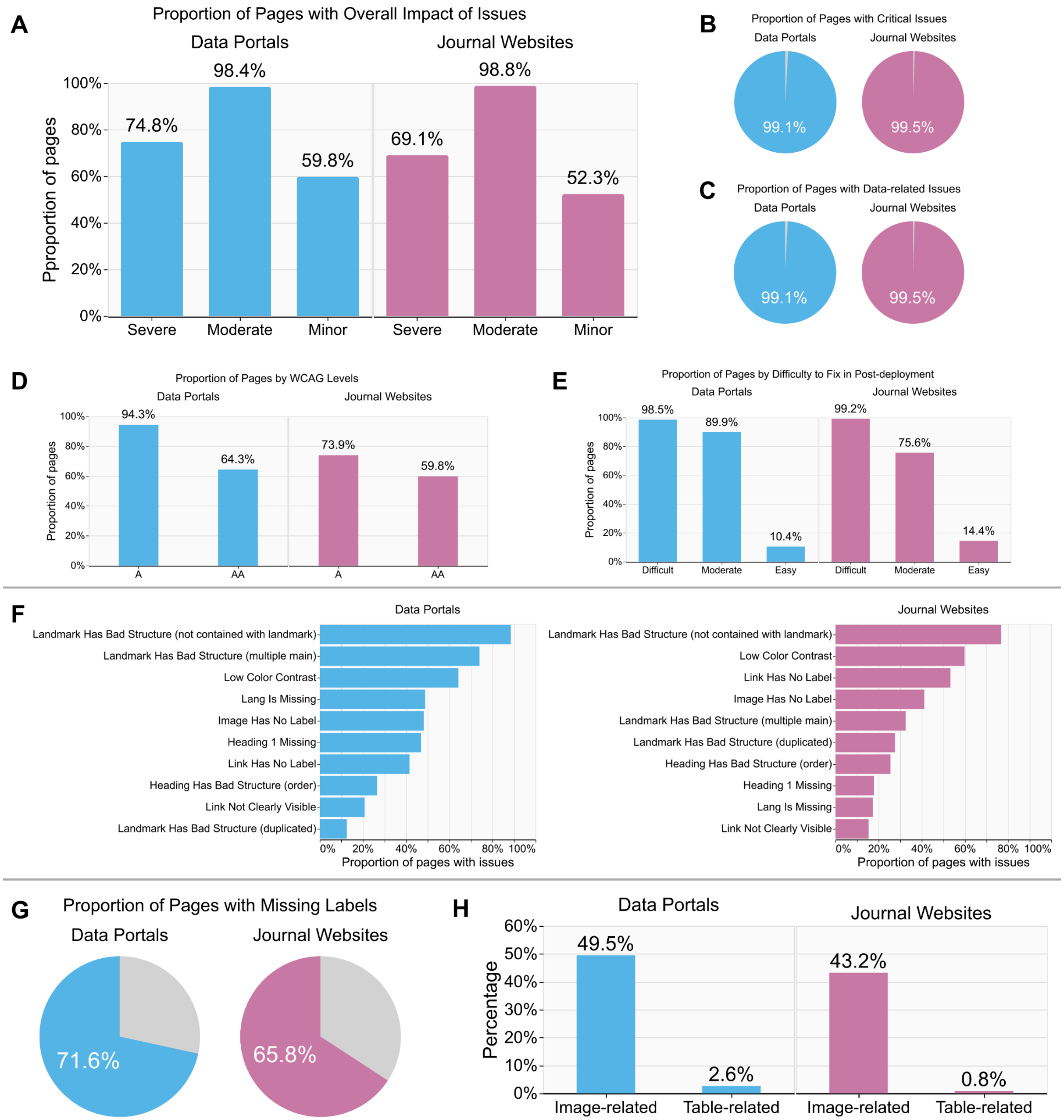

Fig. 2: Accessibility issues found on life sciences websites: Proportion of pages with issues in terms of their (A) overall impact, (B) criticality, i.e., whether the issue is likely to block users performing tasks entirely, (C) whether they are related to the perception of and interaction with data (e.g., issues in tables), (D) WCAG levels[20], and (E) difficulty to fix in post-deployment (e.g., using browser extensions to work around accessibility issues in already deployed websites); (F) Top 10 most frequently observed

accessibility issues. The proportion of pages with (G) missing labels and (H) image-related and table-related issues.

We further analyzed accessibility issues after categorizing them based on several criteria (refer to the Methods section for details). Overall, more than half of the life sciences websites—74.8% data portals and 69.1% journal websites—contain "severe" accessibility issues (Fig. 2A). The severe accessibility issues are identified in our analysis (refer to Methods) by testing several criteria, such as their criticality in performing user tasks and difficulty to fix (Table 1). Most importantly, almost all websites contain issues that (1) can critically block users from performing tasks (Fig. B) and (2) cannot be easily fixed after the deployment of the websites (e.g., using a browser extension to fix issues on existing websites). A typical example is the absence of required labels (Fig. 2G), such as missing alternative text (or "alt text") for images and links, which is one of the most frequently identified accessibility issues on data portals and journal websites (Fig. 2F). Other common types of issues are ill-structured web pages (e.g., broken heading structure that makes it difficult to navigate a web page with a keyboard and a screen reader) and low color contrast (e.g., links that are hard to distinguish from regular text for low-vision users).

Since life sciences data and information are frequently represented in images and tables (e.g., dataset tables in cBioPortal[17]), we explored how many issues relate to images and tables on web pages (Fig. 2H). Our data shows that nearly half of the data portals (49.5%) and one-third of the journal websites (43.2%) had image-related accessibility issues. With regards to tables: 2.6% of data portals and 0.8% of journal websites had accessibility issues, a much lower rate than images. Common table-related issues are missing labels for table headers (e.g., column names), which make it difficult for users with visual impairments to navigate and perceive a large data table efficiently. Since navigating data tables is considered one of the most important tasks for data portal end users[21], even a few accessibility issues on tables can prevent people with visual impairments from accessing essential information.

Comparison by hosting institutions

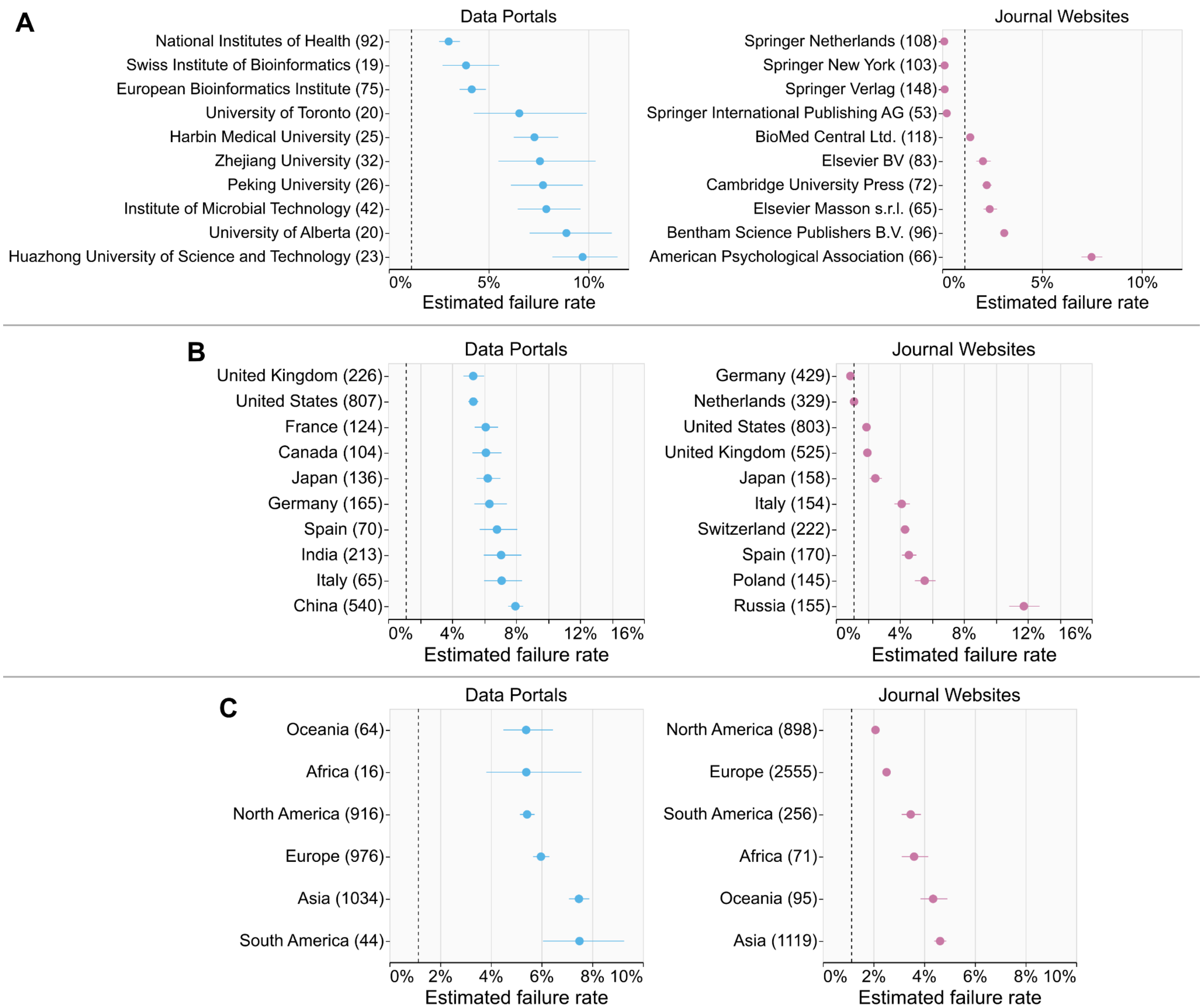

Fig. 3: The estimated failure rates of data portals and journal websites grouped by their (A) hosting institutions and publishers, respectively, (B) countries, and (C) continents. The error bars represent the 95% confidence intervals. The numeric values inside brackets in the y-axis represent the sample size n of corresponding groups. The US government websites' median failure rate (1.1%) is drawn as vertical dashed lines for reference. Only the top 10 hosting institutions (out of 141 for data portals and 121 journals) and the top 10 counties (out of 35 for data portals and 66 for journals) with the largest sample sizes (n) are shown.

We further analyzed the collected data at the aggregated level using several key categories. We first see whether hosting institutions of life sciences websites have affected accessibility (Fig. 3A). Our data shows that some institutions's pages vary a lot with respect to failure rates, while pages hosted by others have more consistent failure rates (error bars in Fig. 3A). Interestingly, data portals hosted by universities generally showed higher failure rates with higher variability (as seen with wider confidence intervals in Figure 3A), while data portals hosted by (inter)national institutions (i.e., the three rows on the top), such as European Bioinformatics Institute (EMBL-EBI) and National Institutes of Health (NIH), showed significantly lower and less variable failure rates.

Comparison by countries and continents

Despite the notable differences between life sciences data portals and journal websites in terms of their contents and use cases, failure rates by geographic region were overall consistent between the two groups (Fig. 3B–C). In both groups, China and India showed significantly higher failure rates than the United States and the United Kingdom (Fig. 3B). Similarly, at the continent level, Asia showed significantly higher failure rates compared to Europe, Oceania, and North America (Fig. 3C). This trend aligns with a recent large-scale accessibility study in general websites[24], where websites in English showed better accessibility compared to websites in other languages, such as Mandarin. Notably, none of the countries and continents showed better accessibility results than the US government websites (dotted lines in Fig. 3B–C), except for journal websites by Germany-based publishers (e.g., *Biological Chemistry*, ISSN: 1437-4315, by Walter de Gruyter). National disability policies are considered to play an essential role in digital accessibility, and our data shows a consistent pattern. In a cross-country accessibility study[25], several Asian countries, including China, are considered to have "weak" disability policies. In contrast, European countries, including Germany and the United Kingdom, are considered to have "strong" policies. In 2002, Germany enacted the ordinance on barrier-free information technology (BITV), which extends the Web Content Accessibility Guidelines (WCAG). This mandates that all public websites in Germany meet this accessibility standard. In line with this, journal websites from Germany showed significantly lower failure rates than other countries.

Manual evaluation results

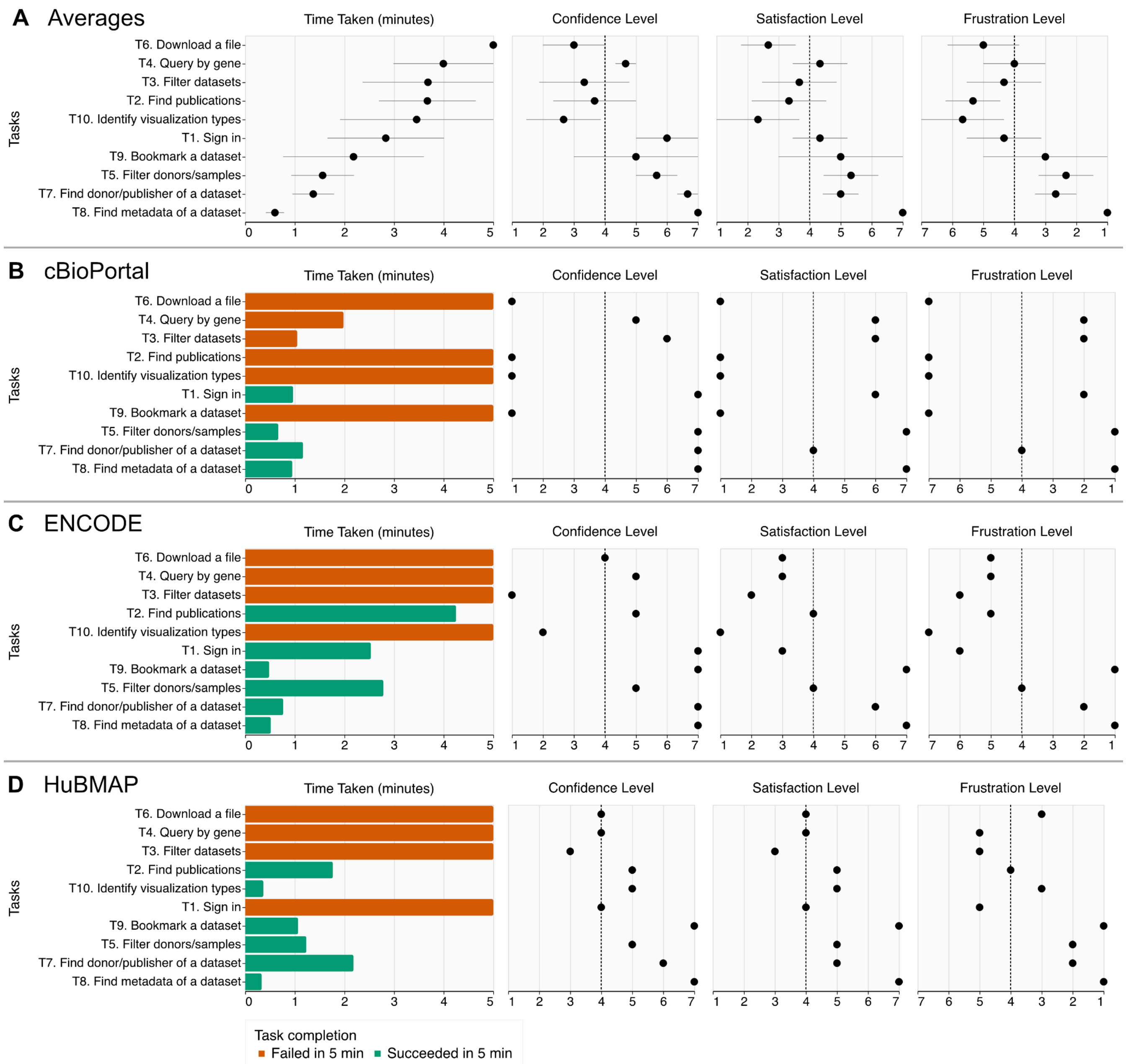

Fig. 4: Results of manual evaluation of three data portals[17,18,26] with a blind screen reader user. The user performed ten typical tasks on each data portal and gave subjective scores about their confidence, satisfaction, and frustration levels for each task. (A) The three results are aggregated with mean completion time and subjective responses, where error bars represent standard deviation. (B–D) Individual results of three data portals are shown: (B) cBioPortal[17], (C) ENCODE Data Portal[18], and (D) HuBMAP Data Portal[26]. HuBMAP Data Portal is developed and maintained by the research group led by the senior author of this paper. Note that the purpose of the manual evaluation is not to compare data portals but instead to understand how accessibility issues identified by computational approaches influence the actual use cases.

Our manual evaluation of selected important data portals with a blind user complements the computational evaluation results, showing the potential influence of the accessibility issues we identified in real-world use cases. The study participant performed a total of 30 tasks in three data portals[17,18,26], which had failure rates of 19.3% (cBioPortal), 2% (ENCODE), and 8.8% (HuBMAP) in our computational evaluation. Refer to the Methods section for the manual evaluation protocol. The participant successfully performed only about half of the tasks (16 out of 30, 53.3%), spending 2.8 minutes out of a maximum of 5 minutes per task on average (Fig. 4). The participant, on average, felt that they were somewhat confident (4.8 out of 7), somewhat satisfied (4.3), and somewhat not frustrated (3.76) with the data portals for each task, where their subjective responses vary across tasks and data portals (Fig. 4A—B). Overall, the participant was able to find metadata of a given dataset or study in all data portals (T5, T7, T8). However, the participant consistently failed to find and download specific datasets (T3, T4, T6).

Through this study, we found evidence of how different types of accessibility issues prevent users from performing typical data portal tasks. For example, the most common and critical accessibility issues hindering the completion of the tasks were related to buttons, links, and images without labels (e.g., alt-texts). Developers usually have to provide such labels when implementing data portals. Sometimes, data portals use icons for buttons without alt text, making it extremely difficult for the user to identify buttons for downloading files (T6) or bookmarking datasets (T9). Sometimes, users can still understand the function of the miss-labeled components through trial and error. However, the complexity of pages in data portals made it very difficult for the participant to guess the functionalities of buttons with missing labels.

We also observed several barriers the participant encountered that the computational evaluation could not identify. For example, the participant had to apply several filters (e.g., assay types or donor ages) using checkboxes, search fields, and even sometimes sliders to search for specific datasets (T3, T4). The participant was able to use these components correctly for some data portals. However, the user found it difficult to grasp all the filters applied in the data portals. This often made the participant apply the wrong filters, leading to inaccurate search results. Another example is that the participant consistently found it confusing to use the auto-complete features of keyword search since the user did not perceive suggested keywords and did not understand that the participant had to select one of the suggested keyboards to proceed with the search. These examples show how specific tasks in data portals could be challenging to perform even when accessibility standards are met.

Discussion

Benefits of centralized development for accessibility

In our evaluation results, the centralized development and maintenance of data resources at national or international institutions may have resulted in more consistent and better accessibility of their life sciences websites[22]. For instance, EMBL-EBI has implemented and maintains a website construction framework—Visual Framework[23]—for their life sciences websites. This framework incorporates accessibility guidelines and best practices as built-in features, which have been iteratively improved over time and are consistently used across websites that the institution maintains. Since learning and implementing accessibility standards is considered notably challenging and time-consuming for developers, using a sharable framework with accessibility standards built in seems to be largely beneficial in supporting better and consistent accessibility.

Implications to real-world accessibility

As our manual evaluation showed, our computational accessibility results underestimate the real-world accessibility issues for actual users. This means that even though our computational results showed a low failure rate (e.g., 1%) for certain life sciences (data) resources, it does not mean that they are overall (e.g., 99%) accessible to users with disabilities. For example, failures to implement accessibility standards on a few critical components of a web page (e.g., a figure containing important life sciences concepts or a data table containing data essential for data-driven research) can still make the web page entirely useless for users with a disability. In our manual evaluation, many tasks that our participant could not complete successfully, such as downloading files and bookmarking datasets, were caused by a single or a handful of inaccessible components on the website. Therefore, we conclude that the real-world accessibility of life sciences data portals and journal websites is much worse than our computational testing results indicate. Hence, automatic approaches can only show an incomplete picture, and conducting user studies with people with disabilities is required to assess real-world accessibility better.

Going beyond existing accessibility standards

Our manual evaluation showed that merely implementing accessibility standards does not guarantee that users can successfully perform typical tasks in data portals. For example, even though two of the data portals in the manual evaluation implemented accessibility standards for interactive components well (e.g., providing readable labels to buttons, checkboxes, and sliders), the participant found it difficult to interact with large datasets (e.g., applying proper filters to select datasets of interest). We think that making life sciences data accessible to people with disabilities requires more than simply following general accessibility standards; it requires more studies and specialized solutions in the life sciences field. For example, given that each data portal can have different use cases in mind, it is vital to include people with disabilities in the design process (i.e., co-design with people with disability) to accurately support accessibility in different contexts.

Limitations

A limitation of our work is that we evaluated only the landing pages of life sciences resources in our computational accessibility testing. In future work, we will focus on reliably identifying common types of subpages (search page, browse page, data page, article page, etc.) of resources at scale. However, given the vital role of landing pages (i.e., people need to first access the landing pages to reach other subpages), we believe that our analysis of home pages provides essential insights into the extent to which current life sciences resources are accessible. Moreover, we mitigate the impact of this limitation on our overall recommendations by conducting manual evaluations with a screen reader user by testing multiple meaningful pages per website (e.g., data page, detail page, and documentation).

Future work

In the future, we will expand the accessibility evaluation of life sciences resources in several regards. First, we will repeat the automatic evaluation regularly in the future (e.g., quarterly or annually) to see temporal trends and improvements in accessibility in the community. The streamlined computational approaches we built in this study will allow us to perform continued evaluations. We will also expand the number and diversity of the targeted life sciences resources. Specifically, we plan to include additional scientific tools for using, exploring, and analyzing data, such as visualization tools (e.g., genome browsers). In our manual evaluation, we tried to see whether a screen reader user can identify the essential information of visualizations embedded in data portals (i.e., visualization types). However, many different accessibility aspects of visualization need to be tested to ensure its usefulness to actual users. To test the accessibility of visualization, we will adopt accessibility standards developed for data visualizations, such as Chartability[27].

Methods

Computational accessibility evaluation

Following best practices in accessibility research[28], we used both computational and manual approaches to test the accessibility of existing life sciences resources. The computational approach enabled us to collect data about accessibility at scale, which we complemented with a manual evaluation with an actual screen reader user who is totally blind to assess real-world accessibility issues and put our large-scale data into a practical context.

To identify data portals and journal websites for the evaluation, we collected lists of websites from public repositories, i.e., Database Commons[29] for data portals and Scientific Journal Rankings (SJR)[30] for journal websites. Of all the data portals (n=6,378) and journal websites (n=27,955), we excluded those that were not considered life sciences resources (e.g., journals in physics) (n=17,774) and that we had issues connecting to (e.g., websites that are no longer available) (n=8,438).

As a result, we selected a total of 3,112 data portals and 5,009 journal websites for our evaluation. In addition to life sciences resources, we collected 852 US government websites as the baseline for our comparative analysis[31], the websites that need to meet strict legal requirements for accessibility (e.g., Section 508 of the Rehabilitation Act)[16]. To test the accessibility of the collected websites, we used the Axe accessibility testing tool[32], which showed reliable and comprehensive evaluation results compared to other tools in previous studies[33]. The Axe supports testing various items that are compliant with Section 508[34] and WCAG 2.1[20] (Web Content Accessibility Guidelines). Using Axe, we collected the accessibility states of web pages, examining whether web pages violate different accessibility issues, such as low contrast ratio, missing alternative text, and empty contents. For improved interpretability of the issues defined by Axe[32] for our study result analysis, we assigned more informative categories to individual issues (Table. 1). Refer to the supplemental data for the full table.

Table 1. A list of categories we assigned to individual Axe accessibility issues. We reviewed 83 unique accessibility issues observed in our study with Axe[32] and created the categories introduced here for our analysis.

Category	Label Description	Values
Overall Impact	The overall impact of a given issue in the context of life sciences data resources. This label summarizes the four categories below (Criticality, WCAG Level, Difficulty to Fix, and Data Related). For example, if an issue is data-related, critical, difficult to fix, and related to WCAG Level A, the issue is considered to severely impact user tasks.	Severe Moderate Minor
Criticality	This category describes whether the issue can entirely block users from performing their tasks.	Critical Less Critical
WCAG Level	This category represents the minimum WCAG[20] level of conformance that is related to a given issue. An issue with	A AA

	Level A indicates that it is the most essential issue for success criteria. An issue with Level AA in our study represents that it is an important issue for enhancing accessibility but is not part of Level A. Therefore, issues with the WCAG Level are the highest priority.	
Difficulty to Fix in Post-Deployment	This captures how difficult it is to fix a given issue after the deployment ("post-deployment"), such as addressing issues on already deployed data portals using browser extensions.	Difficult Moderate Easy
Data Related	This label describes whether the issue is related to perception or interaction with data (e.g., a table in a wrong structure).	Yes No
Missing-label Related	This category captures whether the issue is related to missing labels (e.g., missing alternative text of images).	Yes No

After getting the accessibility results, we calculated the failure rates (FR)[35] of individual websites as an accessibility measurement, which is one of the most widely used metrics. This enables fair comparison between websites, compared to using the absolute counts of violations, as the failure rate takes the size and complexity of websites into account[35].

$$I_p = \frac{B_p}{P_p} \quad (1)$$

The failure rate of a web page I_p is defined as the ratio between the actual points of accessibility violations B_p and the potential points of violations P_p.

We merged the accessibility evaluation results of individual resources with their contextual information provided from the original repositories (i.e., Database Commons and SJR). This includes geospatial dimensions (e.g., country and city), temporal dimensions (e.g., year founded), and impact scores (e.g., citation counts) of data portals and journals.

We used Python (3.10.13) with Altair[36] (5.1.2) in Jupyter Notebooks[37] to collect, manipulate, and visualize data. The accessibility reports are collected using the 4.9.1 version of Axe-core.

Statistical analysis

We performed statistical analyses on the accessibility states to find significant differences across subgroups: resource types (i.e., data portals, journal websites, and US government websites), countries, continents, and hosting institutions. We performed Dersimonian and Laird's random-effects meta-analyses[38] to summarize the failure rates of the websites within subgroups and used a mixed-effects model to account for potential heterogeneity and differences in underlying distributions of failure rates between different websites. The failure rate for a given journal website, government website, or data portal was counted as a single statistic during meta-analysis, and standard errors were estimated using the total number of checks performed for a given page. We accessed significant differences between subgroups of websites by comparing 95% confidence intervals estimated from the meta-analysis.

Manual accessibility evaluation

To see the potential impact of the identified accessibility issues on actual users, we additionally conducted a manual evaluation using three data portals—cBioPortal[17], HuBMAP Data Portal[26], and

ENCODE Data Portal[18]—that are widely used. A co-author of the paper who is blind (hereafter, "a screen reader user") tested data portals using the JAWS screen reader assistive technology[9]. The first author led the user studies. The manual evaluation was conducted remotely via Zoom. The overall procedure of the manual evaluation was designed based on Pernice and Nielsen[39]. We conducted three study sessions on different days, evaluating a single data portal in each session. The session lasted between 1.5 and 2.5 hours.

In each session, the screen reader user was asked to perform ten tasks and given 5 minutes to complete each task. We designed ten typical user tasks for each data portal based on a task analysis study[21]. Example tasks include finding specific datasets by applying filters (e.g., "Find kidney datasets for donors over the age of 65.") and identifying linked publications (e.g., "Find a list of publications that used data in this data portal. How many peer-reviewed papers are there?"). Refer to the Supplemental Note to find all tasks for the three data portals. After either finishing each task or exceeding the 5-minute time limit, the user was asked to answer three questions on a 7-point Likert scale: (1) How confident are you that you performed the given task accurately?; (2) How satisfying was it to use the website to perform the task?; and (3) How frustrating was it to use the website for a given task? The user was asked to elaborate on why they came up with specific scores. After repeating this process for all ten tasks, we finished the session with an interview to collect their overall impression of the accessibility of the given data portal. Through this manual testing, we collected the time taken and the success of each task, as well as scores for three subjective questions.

Data availability

All data used in our study is available through Figshare (DOI: 10.6084/m9.figshare.26801032.v1). All source code to collect, analyze, and visualize data is publically available on GitHub at https://github.com/inscidar/analysis-notebooks.

Acknowledgements

This study was in part funded by NIH grants R01HG011773 and K99HG013348.

Competing interests

N.G. and A.L. are co-founders and equity owners of Datavisyn. HuBMAP Data Portal is developed and maintained by the research group led by N.G. The remaining authors declare no competing interests.

References

1. Swenor, B. & Meeks, L. M. Disability Inclusion — Moving Beyond Mission Statements. *New England Journal of Medicine* vol. 380 2089–2091 Preprint at https://doi.org/10.1056/nejmp1900348 (2019).

2. Andrisse, S. *et al.* Advisory Committee to the Director Working Group on Diversity membership. Preprint at https://acd.od.nih.gov/documents/presentations/02142021_DiversityReport.pdf (2022).

3. Building a Diverse Genomics Workforce: An NHGRI Action Agenda. Preprint at (2021).

4. Swenor, B. K., Munoz, B. & Meeks, L. M. A decade of decline: Grant funding for researchers with disabilities

2008 to 2018. *PLoS One* **15**, e0228686 (2020).

5. Campanile, J. *et al.* Accessibility and disability inclusion among top-funded U.S. Undergraduate Institutions. *PLoS One* **17**, e0277249 (2022).
6. Tenopir, C. *et al.* Data sharing by scientists: practices and perceptions. *PLoS One* **6**, e21101 (2011).
7. WebAIM. Survey of Users with Low Vision #2 Results. Preprint at https://webaim.org/blog/low-vision-survey2-results/ (2018).
8. NVDA: a free, open source, globally accessible screen reader for the blind and vision impaired. *NV Access* https://www.nvaccess.org/.
9. JAWS: the world's most popular screen reader. *Freedom Scientific* https://www.freedomscientific.com/products/software/jaws/.
10. Campoverde-Molina, M., Luján-Mora, S. & García, L. V. Empirical Studies on Web Accessibility of Educational Websites: A Systematic Literature Review. *IEEE Access* **8**, 91676–91700 (2020).
11. Hackett, S., Parmanto, B. & Zeng, X. A retrospective look at website accessibility over time. *Behav. Inf. Technol.* **24**, 407–417 (2005).
12. Ismail, A. & Kuppusamy, K. S. Accessibility of Indian universities' homepages: An exploratory study. *J. King Saud Univ. - Comput. Inf. Sci.* **30**, 268–278 (2018).
13. Kuzma, J., Weisenborn, G., Philippe, T., Gabel, A. & Dolechek, R. Analysis of U.S. Senate Web Sites For Disability Accessibility. *Int. J. Bus. Syst. Res.* **9**, 174–181 (2009).
14. Menzi-Çetin, N., Alemdağ, E., Tüzün, H. & Yıldız, M. Evaluation of a university website's usability for visually impaired students. *Univers. access inf. soc.* **16**, 151–160 (2017).
15. Chintalapati, S., Bragg, J. & Wang, L. L. A Dataset of Alt Texts from HCI Publications: Analyses and Uses Towards Producing More Descriptive Alt Texts of Data Visualizations in Scientific Papers. *arXiv [cs.HC]* (2022).
16. *Section 508 of the Rehabilitation Act*, 29 U.S.C. § 794d. Preprint at (1998).
17. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci.*

Signal. **6**, l1 (2013).

18. Davis, C. A. *et al.* The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.* **46**, D794–D801 (2018).
19. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
20. Web Content Accessibility Guidelines (WCAG) 2.1. https://www.w3.org/TR/WCAG21/.
21. Common Fund Data Ecosystem. CFDE Use Case Library. *CFDE Use Case Library* https://use-cases.nih-cfde.org.
22. Bi, T. *et al.* Accessibility in software practice: A practitioner's perspective. *ACM Trans. Softw. Eng. Methodol.* **31**, 1–26 (2022).
23. European Bioinformatics Institute. Accessibility guidelines of Visual Framework. https://stable.visual-framework.dev/guidance/accessibility/.
24. WebAIM. The WebAIM Million. *WebAIM: Web Accessibility In Mind* https://webaim.org/projects/million/ (2023).
25. Joanne M. Kuzma, Dorothy Yen, Klaus Oestreicher. *Global E-Government Web Accessibility: Empirical Examination EU, Asian African Sites*.
26. HuBMAP Consortium. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature* **574**, 187–192 (2019).
27. Elavsky, Bennett & Moritz. How accessible is my visualization? Evaluating visualization accessibility with Chartability. *Comput. Graph. Forum* (2022) doi:10.1111/cgf14522/v41i3pp057-070.
28. Vigo, M., Brown, J. & Conway, V. Benchmarking web accessibility evaluation tools: measuring the harm of sole reliance on automated tests. in *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility* 1–10 (Association for Computing Machinery, New York, NY, USA, 2013).
29. Database Commons a catalog of worldwide biological databases. *Database Commons* https://ngdc.cncb.ac.cn/databasecommons/browse.
30. SCImago. Scimago Journal & Country Rank [Portal]. *SJR* http://www.scimagojr.com.

31. Cybersecurity and Infrastructure Security Agency. A repository that contains the official, full list of registered domains in the .gov zone. *.gov data* https://github.com/cisagov/dotgov-data.

32. Deque Systems, Inc. axe: digital accessibility testing tools. *axe: digital accessibility testing tools* https://www.deque.com/axe/.

33. Pool, J. R. Accessibility metatesting: Comparing nine testing tools. *arXiv [cs.HC]* (2023).

34. US Government. Section508 of the Rehabilitation Act. *Section508* https://www.section508.gov/.

35. Martins, B. & Duarte, C. Large-scale study of web accessibility metrics. *Univers. Access Inf. Soc.* **23**, 411–434 (2024).

36. VanderPlas, J. *et al.* Altair: Interactive statistical visualizations for python. *J. Open Source Softw.* **3**, 1057 (2018).

37. Loizides, F. & Schmidt, B. *Positioning and Power in Academic Publishing: Players, Agents and Agendas: Proceedings of the 20th International Conference on Electronic Publishing*. (IOS Press, 2016).

38. DerSimonian, R. & Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **7**, 177–188 (1986).

39. Nielsen, Norman Group. *How to Conduct Usability Studies for Accessibility*. (Sams Publishing, Indianapolis, IN, 1999).