

# Sharing Practices for Datasets Related to Accessibility and Aging

Rie Kamikubo  
College of Information Studies  
University of Maryland, College Park  
rkamikub@umd.edu

Utkarsh Dwivedi  
College of Information Studies  
University of Maryland, College Park  
udwivedi@umd.edu

Hernisa Kacorri  
College of Information Studies  
University of Maryland, College Park  
hernisa@umd.edu

## ABSTRACT

Datasets sourced from people with disabilities and older adults play an important role in innovation, benchmarking, and mitigating bias for both assistive and inclusive AI-infused applications. However, they are scarce. We conduct a systematic review of 137 accessibility datasets manually located across different disciplines over the last 35 years. Our analysis highlights how researchers navigate tensions between benefits and risks in data collection and sharing. We uncover patterns in data collection purpose, terminology, sample size, data types, and data sharing practices across communities of focus. We conclude by critically reflecting on challenges and opportunities related to locating and sharing accessibility datasets calling for technical, legal, and institutional privacy frameworks that are more attuned to concerns from these communities.

## CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI); Accessibility**; • **Security and privacy** → Human and societal aspects of security and privacy.

## KEYWORDS

dataset; sharing practices; disability; repository; machine learning.

### ACM Reference Format:

Rie Kamikubo, Utkarsh Dwivedi, and Hernisa Kacorri. 2021. Sharing Practices for Datasets Related to Accessibility and Aging. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21)*, October 18–22, 2021, Virtual Event, USA. ACM, New York, NY, USA, 16 pages. <https://doi.org/10.1145/3441852.3471208>

## 1 INTRODUCTION

A wide range of academic disciplines and industries source datasets from people with disabilities and older adults – each with its own data collection method as well as reporting and sharing practices. A recurring challenge they face is navigating the tension between making their data accessible and restricting access to protect those represented in the data [78, 148, 176]. This tension, typically faced by many social data stewards [112], is not unique to accessibility. However, there are benefit and risk traits unique to many accessibility datasets that can make the challenge more pressing, especially when they can be used in AI-infused applications.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
ASSETS '21, October 18–22, 2021, Virtual Event, USA

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-8306-6/21/10...\$15.00  
<https://doi.org/10.1145/3441852.3471208>

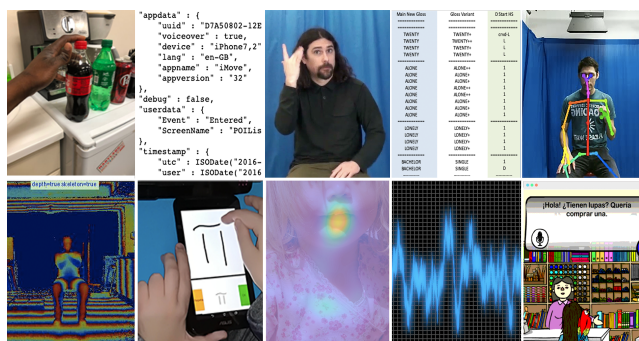


Figure 1: Examples of accessibility datasets such as photos taken by blind users [106], assistive app logs of users with visual impairments [92], sign language videos [85], gloss annotations [130], motion captured signs [80], depth data from older adults' activities [108], stroke gestures by people with motor impairments [166], eye-tracking data from autistic children [54], voice recordings of people with speech impairments [36], and a speech corpus of people with intellectual disabilities [142].

**Benefits.** Curation and sharing of accessibility datasets is crucial for innovation, benchmarking, bias mitigation, and understanding of real-world scenarios, where assistive and general purpose AI-infused applications are deployed. For example, innovation in sign language processing requires videos sourced from the Deaf community [30, 31, 111]. Benchmarking visual question answering can benefit from real-world visual questions from blind people [77]. Bias mitigation could utilize features explicitly representing disability [161]. Understanding the use of mobility applications among blind people in the real world requires longitudinal-data from the intended users [92]. More importantly, representation of people with disabilities and older adults in AI training datasets can contribute to more inclusive AI-infused applications [75, 137, 161, 174].

**Challenges.** Despite their critical role, large datasets directly sourced from people with disabilities and older adults are scarce [30, 90, 128], either publicly available or not [91]. This is partly due to smaller populations [150]. Other factors relate to high variability within a given disability or age group; data often being constrained to very specific tasks, applications, or scenarios making it difficult to aggregate; and annotations often requiring domain knowledge making it difficult to elicit through crowdsourcing tasks. For instance, creating annotated video datasets for sign language synthesis requires linguistic background and sign language fluency [89, 130]. However, increasing representation of these populations yet can amplify ethical and privacy concerns derived from potential data abuse and misuse or re-identification risk [11, 75, 78, 160].

In this paper, we explore how data stewards across different disciplines and over time have balanced risks and benefits when collecting, reporting, and sharing accessibility data. Specifically, we conduct a systematic review of 137 accessibility datasets from 1984 to 2020. As shown in Figure 1, we focus on annotated data resources generated from communities of interest in accessibility and aging that could be used to train or evaluate machine learning models. Datasets are manually located over the period of two years and coded in terms of practices for reporting the communities represented; sample size, data type, and how they relate to the purpose and tasks of the data collection; and sharing practices.

Motivated by an increasing interest in openness, transparency, and accountability in AI, the contributions of this work are: (i) improving the transparency of norms for collecting and sharing data related to accessibility and aging and (ii) calling the research accessibility community to action for guidelines and frameworks. Despite the critical role that data play to replicability, innovation, and a more inclusive future of AI, determining the answer to the questions, “what data can I share and how?” is yet not straightforward.

## 2 RELATED WORK

We discuss and contrast our work with prior work on disability data and data sharing practices in related fields.

### 2.1 Systematic Review of Disability Data

Prior efforts looking at cross-disability datasets have mainly focused on demographics, disability related diagnoses, causes of injury, interventions, outcomes, and costs (e.g. BMS National Database [48]). We also find systematic analyses of accessibility datasets and publications focusing in websites [158], physical environments [52], and data collection and study methods [27, 37, 115, 120]. Complimentary in nature to these efforts, this work focuses on a systematic review of datasets and related publications that source data from people with disabilities and older adults (e.g. photos taken by blind users, video recordings of Deaf signers, and sensor data from gait movements of people with Parkinson’s disease). What is important about these data is that they can be used in AI-infused applications and thus, require a delicate balancing act between benefits and risks for sharing. There are prior attempts reviewing such datasets. However, they tend to be more narrow in focus typically restricted to a population and task e.g. sign language videos of Deaf signers [129, 163], photos taken from people who are blind or low vision [106], and presence of older adults in publicly available face datasets [137].

### 2.2 Data Sharing in Health and Accessibility

Datasets and data sharing have often served to attract, nurture, and challenge people in computing and technology. Many fields, including the health community, have seized this opportunity to promote AI in their area [51, 61, 169]. We observe a similar trend in accessibility, with data challenges in the computer vision, speech, and broader machine learning communities (e.g. Kaggle [86]) designed around tasks and datasets from people with visual impairments (e.g. [77]) and Parkinson’s (e.g. [95, 149]), respectively. Even though

accessibility and health data can share similarities (e.g. smaller populations and sensitive information) risks may differ (e.g. accessing one’s mouse movements or clicks could be easier than having access to one’s blood, saliva, or urea samples). More so, the health community has a longer history of discussions and frameworks around data ethics [176]. For example, venues may require researchers to submit a data sharing plan when registering clinical trials and well-defined data sharing statements along their manuscripts [156]. Such conversations around data artifacts are still ongoing in many computing communities (e.g. [66, 70, 88]). The timing is right for the accessibility community to be involved in these discussions. Within the broader human-computer interaction community, we see that prior work looking at 509 papers on wellness, accessibility, and aging published at ACM CHI 2010–2018, found that only 3 made their data publicly available [11]. While not directly comparable, this number is quite low when compared to another work surveying CHI authors from the same period [168]; researchers found that out of 373 reporting or generating any type of data, 80 shared raw data. Reasons for not sharing included data sensitivity, participant consent, and re-identification risks. We believe that this difference can be also explained by the increased privacy risks for accessibility data that are amplified by the risk of disability disclosure.

## 3 METHOD

To understand the current state of accessibility datasets and reflect on differences across communities of focus and disciplines involved in the data collection in the past, we perform a two-year long iterative process for dataset search, code, and review, which started in November 2017 and ended in March 2020. The scholars involved had varying levels of familiarity with accessibility and AI. This review includes a qualitative and descriptive analysis of the identified datasets spanning over 35 years (1984–2020, N=137 datasets).

### 3.1 Identifying Accessibility Datasets

Dataset search is a difficult task with some researchers arguing that it is a field on its own [38]. Other factors making our search even more challenging are: datasets may not be publicly available; there is a vast terminology for communities of focus in accessibility and aging; and data collection tasks and purpose can highly vary.

*3.1.1 Criteria for Inclusion.* Datasets had to satisfy the following: **Include data sourced from people with disabilities.** Many datasets may include people with disabilities without having explicit information about disability as about 15% of the world’s population live with some form of disability [175]. Given that there is no way for us to verify it, those datasets are not included in our collection. We include only datasets, where people with disability are explicitly mentioned to be contributing data. Terminology related to disability can be challenging. When in doubt, we consulted terminology in the IDEA Act [41, 43], the ADA act [162], and accessibility research. We found that even when appropriate terms were explicitly mentioned they were ambiguous or misleading. Here are a few examples related to data from the Deaf community. Sometimes, data stewards refer to those contributing data as “participants” and in other parts mention that they “talked to the Deaf community” without explicitly stating whether those contributing data were

deaf/Deaf. The term “signer” is often used in the linguistic community to indicate native or fluent signers. However, there are many sign language datasets where the signers are not actually deaf/Deaf or hard-of-hearing people even though the term “signer” is used *e.g.* to refer to interpreters [138]. Also, it is not uncommon for those new to the field to collect data on their own or with other hearing people just by mimicking signs and still use the term “signer” without any other information [96, 105]. Even when terms like “fluent” or “expert” are used, often the data collection process is missing information about the people contributing to the data thus preventing us from discovering if the experts or those fluent in a certain sign language were deaf/Deaf or hard of hearing versus sign language learners or hearing interpreters [30]. When we are not certain, datasets are not included.

**Include data that can be used in AI-infused applications.** We focus on datasets that include images, text, sensor data, and logs generated or reflective of people with disabilities. We focus on AI-infused applications and technologies for end users. Datasets that used sophisticated technology such as fMRI with people with dementia [121] or invasive technologies such as brain electrodes for enabling motor control, do not meet this criteria. While these technologies are promising, they are applicable to clinical settings. They would be out of reach for most of the general public and only be of interest to cliques within their respective research communities.

**3.1.2 Search Strategy.** We employed a multilayer strategy for identifying relevant datasets: (i) *open search* *e.g.*, using specific keywords on search engines, (ii) *focused searches in repositories and publication venues*, and (iii) *focused searches on authors* that we found to have shared or provided contact information for sharing datasets in the past. This was not a linear process. We would go back and forth.

For the **open search**, we used terms that are associated with the communities of focus such as *disability, disabled, accessibility, impairment, visual, disease, speech, hearing, cognitive, mobility, vision, autism, behavioral, developmental, disorder, learning, cleft, dyslexia, dysarthria, Parkinson’s disease, speech pathology, phonetic articulation, sign language, cleft lip, Aspergers syndrome, low vision, depression, and bipolar*. These were complemented with terminology used in the IDEA [41, 43] and The Americans with Disabilities [162] Act along with the terms such as *dataset, database, repository, data, collection, eye tracking, kinect, and glove*.

For the **focused search in repositories and publication venues**, we started with known machine learning repositories such as Kaggle [86], UCI [18], and VisualData [62]. We found that less than 0.15% of available datasets in Kaggle were somehow related to disability and aging; the majority of them focused on health-care rather than assistive or inclusive technologies. Only 4 of them met our criteria. Similar patterns were found in UCI and VisualData with 3 and 2 datasets meeting our criteria, respectfully. We also searched through digital libraries in scientific societies such as ACM [2], IEEE [6], LREC [8], ISCA [7], ACL [1], and CVF [4] as well as open science efforts such as Zenodo [59]. Many of these venues do not typically provide a database or a search engine for the available datasets<sup>1</sup>. When possible, we used the keywords from our open search or an HTML Regex addon, Find Plus [144], to find

matches on lists of publications or datasets. As a last resort, we manually scanned publications, *e.g.*, by zooming in on a specific venue such as ASSETS 2008-2018. In the midst of our research in 2018, we saw many efforts from the industry to help with the discovery and sharing of datasets such as the Registry of Open Data on AWS [3], Microsoft Research Open Data [9], and the Google Search [5], though, they resulted in very few new datasets in our collection.

Last, for the **focused search on authors**, we used a seed list of authors resulting from the previous searches and started growing it organically. As we discovered more datasets, we identified common authors, principal investigators, and funded projects which we kept adding to our list of authors and keywords.

## 3.2 Coding and Analyzing Accessibility Datasets

We explore how data stewards across different disciplines and over time have balanced risks and benefits when collecting, reporting, and sharing accessibility data by looking at current and historical trends through an exploratory analysis. Deciding-what-to-explore-next is one of three key challenges in exploratory data analysis [104]; in our analysis the question formulation were affected by (a) the feasible shape and structure of our codes, (b) the intended audience such as researchers and policy makers, and (c) our domain knowledge. For example, we wanted to explore data size in terms of number of data points per data contributor (a study participant in most datasets). However, accessibility datasets are so diverse in terms of datatype, granularity, and annotation, that having information on the number of data points per participant is not feasible and often not meaningful for a comparison. Even when looking at prior surveys on datasets within a community (*e.g.*, sign language data), we see comparisons at a per participant level being avoided as the size of datasets could be measured either in sentences, individual signs, duration of continuous signing, or richness of linguistic annotations.

**3.2.1 Manual Coding.** We extract the following when available:

**About:** We note the name, year of release, links to the paper and/or dataset, DOI, data stewards (*e.g.*, authors), and contact emails.

**Contributors:** We extract terms originally used to describe the individuals contributing data.

**Communities of Focus:** We assign a dataset to one of more communities of focus based on the data contributors (see Section 3.2.2).

**Collectors:** We note whether the dataset was collected by an educational institution, industry, or both.

**Access:** We scan the papers and/or online records to see whether there is a link for direct download or an explicitly mentioned point of contact for accessing the dataset as well as associated licensing.

**Ethical Board Clearance:** We mark whether ethical board clearance is mentioned in the datasets or their associated materials.

**Data Size:** We use the number of contributors as a proxy for data size. Control groups when available are coded separately.

**Data Type:** We make a note of the technology used to collect the data and list all formats both for the data and annotations if any.

**Summary:** We write a few sentence-long summary about the people, strategy, and purpose of the data collection.

<sup>1</sup>ACM Digital Library offers a dataset search only for dataset artifacts; does not include datasets that authors may link to in their publications.

**3.2.2 Communities of Focus.** Rigid categorization of disability is difficult [27, 133] and perhaps a questionable task as conditions are vast and fluid [174]. To aid the analysis and presentation, we annotated the datasets in our collection across 10 groups, with many communities falling under more than one. In the Results section, we augment discussions by analyzing the language used to describe the people within these communities (terms are indicated in quotes) and contrast them to current guidelines [79, 134].

**Autism.** This group includes datasets sourced from autistic individuals, both children and adults (e.g., [54] and [58]). The umbrella classification of *Autism Spectrum Disorder (ASD)* refers to disorders of brain development affecting social interaction or verbal and nonverbal communication [74]. It is a group on its own because the community has received attention as a specific subset of accessibility research [115, 155] and in the IDEA Act [42].

**Cognitive.** This group captures a much broader category of datasets (e.g., [12] and [69]) sourced from individuals with cognitive disorders commonly leading to different types of impairments such as Parkinson’s disease, which can cause speech or mobility impairments, or Aphasia, which refers to a language impairment. We often cross-listed these datasets in multiple groups listed below. The group also considers other medical conditions that can lead to cognitive impairments such as bipolar disorder or functional memory deficit.

**Developmental.** This group broadly captures datasets from people with any physical and/or mental disability (e.g., [45] and [64]) that began before the age of 22 [79]. Though many people with a developmental disability do not have an intellectual disability [134], the community is being noted as Intellectual or Developmental Disability (IDD) [115]. We group datasets that particularly mention and include individuals with IDD.

**Health.** This group refers to datasets sourced from people with a diverse demographic typically collected in research related to healthcare as well as health and wellbeing management (e.g., [108] and [53]), though, we see intersections with other sub-disciplines of HCI, particularly aging and rehabilitation [11, 115]. Thus, this group includes older adults or people undergoing rehabilitation, as well as those with specific health concerns that do not belong to other groups such as cardiac diseases.

**Hearing.** This group follows a combination of the IDEA Act category including datasets that involve people who are deaf/Deaf and hard of hearing. In this group, those who are deaf/Deaf often contribute as signers (e.g., [34] and [85]).

**Language.** This group includes datasets from people with language disorders such as aphasia or impairments represented by low verbal IQs that affect a person’s ability to communicate. We group datasets that particularly focus on limited abilities to use and express language (e.g., [50] and [171]).

**Learning.** This group includes datasets from people with conditions that are neurologically-based and that can lead to difficulty in learning and using related skills [134], such as reading or writing. We particularly group datasets that mention and include people with dyslexia, dysgraphia, or dyscalculia (e.g., [143] and [114]).

**Mobility.** This group includes datasets from people with mobility and motor/dexterity impairments (e.g., [166] and [17]). Although the word Mobility is used here for grouping, we generally include

datasets that target people with limited physical functioning of one or more limbs, such as in walking or moving hands or fingers.

**Speech.** This group covers datasets sourced from people with limited or impaired speech patterns, such as found in dysphonia, dysarthria, stuttering, and conditions of cleft lip and cleft palate. Since speech impairments can be caused by Parkinson’s disease this group is cross-listed with other groups. Examples of data are [167] and [124].

**Vision.** This group includes dataset sourced from people who are blind or have low vision (e.g., [65] and [77]), which are commonly noted as the *visually impaired* or those with *vision loss* [115].

### 3.3 Reflections in Our Method and Limitations

While we strived to be methodical in dataset search, it involved a lot of detective work. Thus, the dataset collection could be biased. The overseeing faculty, who has been working in the area of accessibility for more than 12 years, leveraged domain knowledge to complement the search by pointing to publication venues, data repositories, datasets, and data stewards. The initial search was performed by two Master’s students in Information Management who were not familiar with accessibility and were guided through weekly meetings with the overseeing faculty. Then two HCI graduate students (Master’s and PhD level) continued expanding the collection and started coding. They calculated inter rater agreement. Though this was found to be high, there were many errors in the coding. This is partially explained by the fact that the task was inherently messy and challenging. We see similar tasks, even when limited in scope within the field of accessibility and publication venues such as ACM, characterized as “challenging and effortful” [115].

We found disability-related terminology to be a dual challenge. First, it was difficult to come up with an extensive list of past and current terms than can be used as keywords for finding datasets as data stewards may have used depreciated terminology or more fine-grained terms established within a field to refer to specific sub populations. Second, it was difficult to map terms used in datasets and/or associated publications to the communities of focus. Partially, this is explained by the fact that in the broader research community, terminology used to describe the people contributing data is often confusing, as discussed above. However, we believe that the lack of extensive accessibility experience may have also contributed. The team would meet often to discuss challenging cases and resolve disagreements. A fifth student, pursuing a doctoral program, was added to the team; the student had a few years of experience in accessibility. The student had a detailed pass on the annotations, resolved any conflicts with the first PhD student, confirmed difficult cases with the faculty, and helped co-lead the analysis efforts.

## 4 RESULTS

We characterize the current status of accessibility datasets (1984-2020) in terms of the communities of focus, current distribution of the datasets across communities and how that changed over the years, data collection purpose, and language used to describe the people who contributed data. We report trends across data size in terms of the number of people involved referring to both those

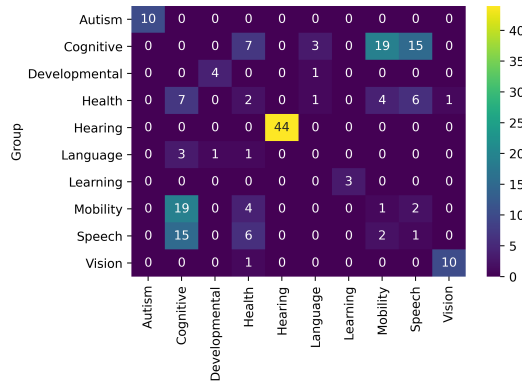


Figure 2: Distribution of dataset count across communities.

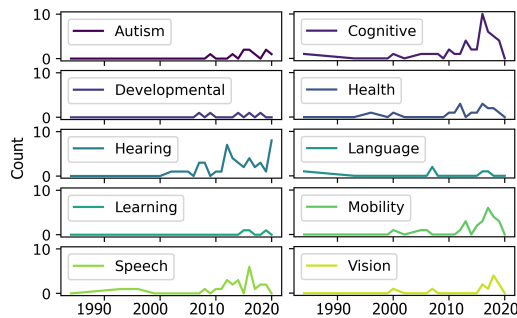


Figure 3: Dataset count over the years across communities.

representing communities of focus and those serving as proxies or control [115]. We explore how data types relate to communities and the purpose of the data collection. Last, we identify common data sharing practices and report on mentions of research clearance.

## 4.1 Communities of Focus

In Figure 2, we see that at a high level the representation across communities is not equally distributed with datasets sourced from the Deaf community and broader Cognitive group dominating; followed by datasets related to communities within the Mobility and Speech groups, which are often cross listed with Cognitive. Surprisingly, the disproportionate attention that communities within the Vision group have received in accessibility research [115] is not reflected in datasets, though we see an uptick in more recent years (Figure 3).

**4.1.1 Autism.** Datasets under this group (a total of 10) starting around 2010, include data both from children (5) and adults (5). Data from children focus mostly on early detection through eye-tracking measurements [35], vocalization composition [177], or eye-contact detection in dyadic interactions [140]. Only one relates to analyzing behaviors of participants interacting with technology such as robots [109]. Data from adults focus on increasing accessibility, such as browsing [58] and text readability [178]. When describing the people contributing data, we find that terms such

as “*person with autism*” or “*person with ASD*” were the most common (7), in contrast to recommended language for this community, either identity-first (e.g., autistic person) or person-first (e.g., person on the autism spectrum) [134]. Datasets often resulted from studies assessing differences across autistic and non-autistic participants. Non-autistic individuals were described as “*neurotypical*” or “*non-ASD/non-autistic*” by accessibility and health informatics researchers [35, 58] and as “*healthy controls*” or “*typically developing*” by others [54, 177]. There were few occurrences of misnomers such as “*high functioning*” which open up the risk of inaccurate labels to describe autistic individuals [16]. Surprisingly, they were present both in broader computing and accessibility publications [54, 57].

**4.1.2 Cognitive.** Datasets under this group (a total of 44) also start around 2010 with a few occurrences as early as 1985. They encompass a wide range of communities typically including people with Parkinson’s (PD) [135], Dementia [118], Dysarthria [67], Alzheimer’s [98], Aphasia [117], Huntington’s (HD) [126], Amyotrophic Lateral Sclerosis (ALS) [81], Episodic Memory Impairment [107], Ataxia [87], or Bipolar disorder [69]. While the majority of datasets focus on Parkinson’s disease (29) we often see a combination (e.g., PD and HD [126]; ALS, PD and HD [81]) with a shared goal of identifying “*neurodegenerative*” signs underlying these conditions, commonly in speech or mobility. This is why in Figure 2 we often cross list datasets in this group with Mobility (19) and/or Speech (15). Guidelines recommend for cognitive disabilities to be defined clearly [79]. We see this being a common trend (42), except when referring to “*people with mild cognitive impairment*,” a diagnostic category in which its formal definition and measures are the field of inquiry [139]. Given that communities in this group have been actively studied in health and clinical domains for a long time, we see more organized data sharing efforts such as TalkBank [116] released in early 2000s followed by AphasiaBank [117].

**4.1.3 Developmental.** Datasets under this group (a total of 5) shyly start appearing before 2010 and include data from people with a specific diagnosis related to developmental disabilities (e.g., Down syndrome) [45, 64] and those without a known diagnosis whose disabilities are described as moderate or mild intellectual disabilities [13, 45, 63]. The majority include data from children (3). Developmental disability represents a broader category of often lifelong disability that can be intellectual and/or physical [79]. The term “*intellectual disabilities*,” often noted as (ID), was found in studies that referred to participants’ limitations in cognitive functioning and adaptive skills (4), such as communication [13] or reading [63]. The datasets associated with studies in ID are typically used in machine learning models to find diagnostic predictors [13, 63]. The term “*developmental disorders*” was found in a study focusing on the identification of what authors call “*ultra-rare developmental diseases*” using facial phenotypes from photographs [64]. One of the datasets includes adolescents with language impairments and low non-verbal IQ [171]; this dataset is thus cross-listed in Figure 2 with Language and Developmental. In contrast to other groups, we do not see a growth over the years and intent of data sharing (see Section 4.4).

**4.1.4 Health.** Datasets under this group (a total of 17) with peaks in the 90s and this past decade, include data from people undergoing

stroke rehabilitation [53], people with depression [69], retinopathy disorders [84], levodopa-induced dyskinesia [110], dysarthric speech after neck cancer surgery [40], or suspected dementia [21]. Datasets are often cross-listed with Cognitive (7), Speech (6), and Mobility (4). The overall space of Health is vast, but within the scope of HCI, studies have been centered around the impact of technologies on the practices and experiences of health professionals and patients [26]. Similarly, datasets in this group relate to applications in healthcare settings, incorporating telehealth mechanisms or automatic patient screening (e.g., predicting Parkinson’s disease progression with smartphone data [17]). Thus, their increase in the last decade could be explained by these research directions. In contrast, in the 90s the emphasis was on clinical studies analyzing dysarthric speech [123] or retinal images [84]. Datasets from research in aging are also prevalent in this group, where terms “*older adults*” and the “*elderly*,” which the latter is rather deprecated in the community [113], are used. The purpose of the data collection often relates to understanding age-related factors, such as motor movements [108, 127]. Although these datasets represent a highly diverse demographic in terms of aging [102], we did not observe a consistent age being used as a threshold.

**4.1.5 Hearing.** Datasets under this group (a total of 44) starting in year 2000, include data from people who are deaf/Deaf or hard of hearing, typically consisting of sign language videos and gloss annotations. Even though this is the most represented group in terms of the number of datasets, researchers in related fields still call for more datasets: larger, more representative, and public [30]. This group is also the most diverse in terms of the research communities involved in the data collection process including computer vision [32], linguistics and natural language processing [55, 130], as well as accessibility and human-computer interaction [85]. However, as discussed in Bragg *et al.* [30], data collection still occurs in separate disciplinary silos. We find that the most common sign languages elicited were American (12), German (8), and French (6); the rest were Polish, Greek, Chinese, Finnish, British, Bangla, Turkish, Czech, Auslan (Australian sign language), Libras (Brazilian sign language), Arabic, Flemish, Spanish, Korean, Russian, and Swiss-German Sign Language. Very few datasets targeted multiple sign languages [44, 122, 159] to establish machine learning benchmarks. When referring to signers contributing data, terms like “*Native*”, “*Expert*”, and “*Fluent*” were often used.

**4.1.6 Language.** Datasets under this group (a total of 5) appearing early in 1985 and then much later in 2005, include people with aphasia [50, 98] or language impairments who have intact or lowered non verbal IQ [171]. Aphasia is a disorder of linguistic processing related to specific brain regions [46]. Hence, datasets related to Aphasia are cross-listed with Cognitive, though, not always restricted to people exclusively with aphasia. All datasets in this group involve adults [50, 151] except one, which focuses on understanding the impact of language impairments on narrative skills of children as they reach adolescence [171]. Most, collected by researchers in Speech and Hearing Science or Aphasiology, aimed at identifying patterns of linguistic decline [98, 117, 151] with the intent to share data (e.g. through AphasiaBank) only for specific research purposes. In contrast, datasets collected among accessibility researchers for purposes of developing augmentative and

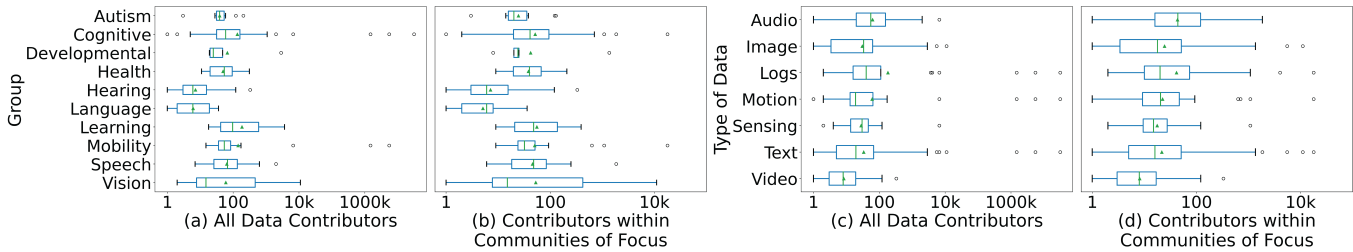
alternative communication devices [15] did not indicate any intent for data sharing.

**4.1.7 Learning.** Datasets under this group (a total of 3) started appearing in 2015 and include data from children and adults with dyslexia; with the target population being often children from 7 to 17 years old [141]. This is perhaps not a surprise as dyslexia is the most common neurobehavioral disorder related to children’s learning ability [152]. Datasets typically aim to support screening of dyslexia [143] and sometimes involved eye tracking to build predictive models [114, 143]. Data collection approaches vary significantly from gamified exercises eliciting data from a large pool of dyslexia-diagnosed children [141] to studies involving university students despite the work’s motivation on children with dyslexia [114].

**4.1.8 Mobility.** Datasets under this group (a total of 27) appearing around year 2000 and quickly growing, include hand or gait movement from people with motor impairments, mostly (21) found in studies related to Parkinson’s disease [164, 173] or dyskinesia [110]. Dyskinesia is commonly seen in Parkinson’s disease patients after prolonged treatment with levodopa, often noted as levodopa-induced dyskinesia (LID), causing involuntary and uncontrollable movements [23]. Often studies include people with not only PD or LID but also various motor impairments, such as cerebral palsy [166], spinal cord injury [132], or spinal muscular atrophy [166]. Datasets vary in the sensing modalities, ranging from mouse cursor movements [173], wearable sensors [157], or stroke-gesture input on touchscreens [166], to stride measurements from insole sensitive resistors [81, 164] or vision-based pose estimation [110]. We observe that over the last decade datasets in this group are often shared.

**4.1.9 Speech.** Datasets under this group (a total of 24) appearing since the 90s, include mainly people with dysarthria (11) and Parkinson’s disease (12) with some overlap between the two. The other datasets included samples of impaired speech due to neck cancer [40], cleft lip [167], and other causes of pathological speech like Amyotrophic Lateral Sclerosis (ALS) [145]. Given that communities in this group have been actively studied for a long time, we see more organized data sharing efforts, common in the speech and language processing research community, such as creating dysarthric speech databases for speech recognition. Dysarthria is a motor speech disorder due to neurological disease or injury, and those with dysarthria experience difficulty in articulating words [47]. Many have build benchmark datasets for dysarthria since early 90s [49, 100, 123] with others following in more recent years [39, 145]. Dysarthric speakers contributing their data often have cerebral palsy [49] and some have Parkinson’s disease [67] or ALS [145]. However, the speech and language community still calls for more datasets with more appropriate sizes for machine learning tasks [68].

**4.1.10 Vision.** Datasets under this group (a total of 11) typically include people who are blind [14, 24] (6) or have low vision [165]. When datasets are collected with real-world assistive applications, where disability status as well as visual acuity or age of onset are not known, the umbrella term, “*people with visual impairments*” or “*visually impaired*” is used to describe those contributing data (e.g. [92]). Datasets are typically collected in the context of accessibility such as navigation [65], object recognition [154], and accessibility of



**Figure 4: Sample size across communities and data types: all contributors (a&c) vs. those within the communities of focus only (b&d).**

web or touchscreen interfaces [24, 165]. There is one exception, where the context is clinical, focusing on screening of Proliferative Diabetic Retinopathy based on retina images [84]. Although datasets in this group involve communities dominating accessibility research [115], they are typically not shared. This might seem surprising. However, we suspect that it merely mirrors local data sharing standards in the accessibility community, where dataset contributions are rare [11, 115]. Perhaps, this could be explained by the awareness that many accessibility researchers have towards potential risks but also the lack of guidelines and frameworks for ethical data sharing [128]. Surprisingly, we did not find any datasets sourced from deaf-blind people, even though in the United States only, this population is estimated to be as high as 2.4 million [83].

## 4.2 Sample Size

Figures 4a and 4b show the number of people ( $N$ ) contributing data across groups; Figure 4a includes those who served as a control and Figure 4b does not. Figures 4c and 4d also show the sample size with and without the control respectively, this time across data types. Distributions are visualized as boxplots with a 1.5 interquartile range (IQR) on a logarithmic scale. For comparison, means are denoted with a triangle. We see that  $N$  is highly variable across and within groups and that those serving as control are contributing to outliers (above 1000K). Looking at Figure 4b, where data are sourced only from those within the communities of focus, we find that overall reported size ranged from 1 to 17,843 (median=20, IQR=51-8). Only few (2) datasets do not provide information on the number of data contributors. Within the 58 datasets that included control groups (e.g., participants with and without dysarthric speech [124]), the reported size ranged from 1 to 31,321,070 (median=31, IQR=68-34).

In this section, we focus our analysis on sample sizes that refer only to the number of participants within the communities of focus (Figure 4b). We observe  $N=3$  being the most common sample size (8 datasets);  $N=1$  and  $N=10$  were second most common (each across 7 datasets); and  $N=2$  being next in line (6 datasets). The majority of datasets (115) have  $N \leq 100$  and  $N \geq 1000$  for only a few (7).

Across all the communities of focus, the Speech group had the largest median value of 48 (85-18). Often studies in this group involved clinical trials for data collection that were performed through collaboration with the local hospital and medical lab, and/or the federation associated with the target disability group. In the extensive project by [67], in partnership with multiple institutions including

a hospital and a team of doctors for over 30 years, they produced two corpora having the largest population size of dysarthric speech data, with a total of 1,850 participants diagnosed with various neurological disorders. In comparison, Whitaker database of dysarthric speech due to cerebral palsy [49] had the smallest sample size (6) and was created by two research teams from different universities.

Although the Language group also collected speech samples like the Speech group, their median sample size was one of the smallest among our communities of focus. The median was 6 (8-2), the same value as the Hearing group of 6 (16-3). This could be due to a lack of clinical collaboration. The largest sample size in the Language group was 36 in [151], being the only dataset (out of 5) that was created in collaboration with a medical institution specialized in the target population (i.e., Aphasia). It also could be due to the complexity of speech tasks associated with the speech recordings. Studies in the Speech group often include a series of short tasks to collect vowel or phonation sounds following a limited set of acoustic measures (e.g., [67], [36]), whereas studies in the Language group include story telling and conversational tasks to assess verbal and non-verbal communication skills that are more complex to analyze (e.g., [151], [171]). The population pool size could also impact the availability of participants; for example, [15] in the Language group collected data from aphasic participants but [67] in the Speech group collected dysarthric speech data from participants with various neurological diagnoses, such as Parkinson’s, Amyotrophic Lateral Sclerosis, and cerebellar diseases.

**4.2.1  $N > 1000$ .** When looking at datasets sourced from more than 1000 participants (a total of 7), we observe that typically a remote data collection in the real world is involved, such as tracking device usage or asking people to upload data regularly using an app. Examples include the mPower study [29] that involved 1087 Parkinson’s disease diagnosed individuals to record and register daily 4 activities via a mobile application: tapping, memory test, walking and voice samples of vowel pronunciations. Similarly, the iMove and Vizwiz datasets were collected with assistive applications deployed in the real-world with 4055 and 11045 participants, respectively [25, 92]. The largest sample size was 17,843, where mouse and keyboard interactions with a search engine were collected and users were classified as having Parkinson’s disease based on proxy information from web search queries (e.g., “I have Parkinson’s”) [180].

**4.2.2  $N \leq 3$ .** Many datasets (a total of 21) had a small sample size, with data generated from no more than 3 people. The majority of

them (15) involved sign language recordings (e.g. [94, 170]). Unlike the large datasets above, where membership to a community of focus was often inferred on proxy information (e.g., use of a screen reader to access an assistive technology app or relevant search terms), these small datasets typically included rich data from series of tasks in controlled settings, where researchers had detailed information about those contributing to the data. For example, for sign language datasets this typically involves lab recordings in a fixed multi-camera setup with additional apparatus used to capture and process data. Given the difficult and costly process of these data collection methods, where people are explicitly asked to complete a series of tasks, it is not surprising to see the same people contributing to studies performed by the same research team (e.g., [136] and [97]). Depending on the amount of linguistic analysis required, such as gloss and non-manual annotations, these datasets could be constrained by the smaller pool of individuals who beyond sign language fluency also have computational linguistic training.

### 4.3 Data Types

We describe the different types of data generated across the communities of focus. Figure 5 shows the distribution of data formats across the communities and Figure 6 shows how the distribution has evolved over the years. We reviewed all datasets and coded the data formats as Audio, Image, Logs, Motion, Sensing, Text, and Video. A dataset often has multiple formats e.g., the raw data and annotations.

**4.3.1 Audio.** This format (in 39 datasets) is typically found in the Speech (24) and Cognitive (23) groups, which often overlap as they collect audio samples targeting dysarthric speech [135]. We also see audio data in the Health group (10) such as speech before and after a neck cancer surgery [40] or audio data from heart sound signals [119]. Speech tasks range from scripted vowels and sentences [67] to story narratives [171], picture descriptions [117], and conversations in pairs [98], which are produced not only by people with speech impairments but also those with language impairments or intellectual disabilities. The recordings are typically done in an isolated booths, especially in studies that collected vowel or phonation sounds [13, 67]. More recently, audio is collected through smartphones and some times in a real-world context, e.g., Parkinson’s detection through body sound collected through smartphones [181]. Audio recordings of speech included English (16), Czech (4), Spanish or Colombian Spanish (4), German (3), French (3), Dutch (2), Korean (1), Italian (1), Taiwanese (1), and Turkish (1).

**4.3.2 Text.** This is the most common format (in 94 datasets) as text is often used to supplement and annotate other types of data. For example, gloss labels are used to annotate sign language videos [56] and phonemic transcriptions are used for spoken utterances of people with cognitive decline [22]. Thus, it is not a surprise to see this format present in the Hearing (43), Cognitive (26), and Speech (20) groups. Only a few (2) datasets were exclusively text based. They involved Reddit posts related to living with a mental illness (e.g., [72]) and corpora related to text readability (e.g. [63]). Text is also found along side formats such images and logs. For example, VizWiz [77] includes crowdsourced text responses from

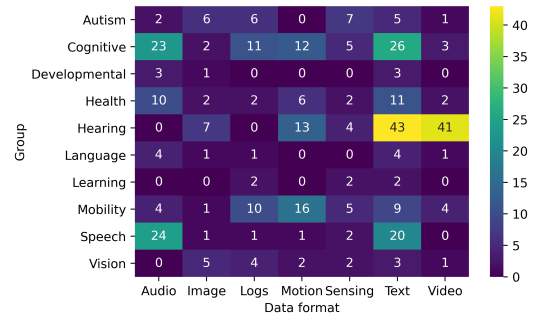


Figure 5: Distribution of data types across communities.

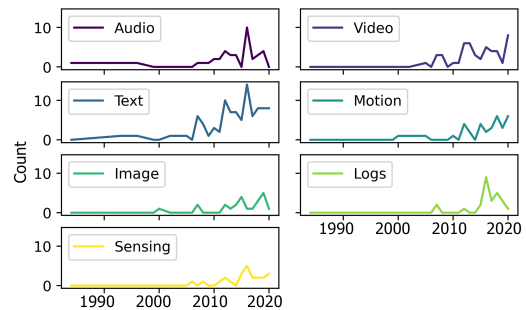


Figure 6: Dataset count over the years across data types.

sighted people to visual questions asked by blind people. Another dataset captures cursor tremors along web search queries [173].

**4.3.3 Image.** This format (in 23 datasets) typically includes photos taken by blind individuals, snapshots of signing, and people’s faces during interactions. It is most commonly found in the Hearing (7), Autism (6), and Vision (5) groups. While sign language data are typically in a video format [131], images are also used e.g., to capture and analyze individual hand shapes in fingerspelling [153]; sometimes in combination with depth images [170]. Images are also collected in behavioral studies of autistic children often including their face during child-adult [140] and child-robot [109] interactions. Photos by people with visual impairments are typically taken from an egocentric viewpoint through smartphone cameras and they tend to include objects to support computer vision technology [106, 154]. Only one dataset include third-person viewpoint images from a fundus (retinal) camera used for locating blood vessels [84].

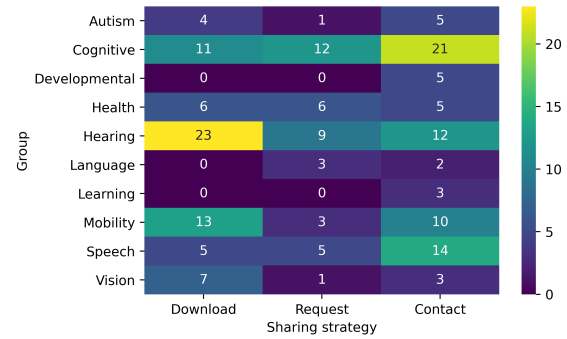
**4.3.4 Sensing.** This format (in 23 datasets) involves data collected by sensing instruments that are not easily captured by the motion category such as eye-tracking measurements [57] or heart sound [119] and EEG [132] signals measured using a medical device. It is typically present in datasets from autism studies (7) as they often employ eye tracking with autistic children and adults. Some of these studies try to observe differences in gaze fixations between autistic and non-autistic adults when reading [179]. Other assess traits by tracking gaze patterns of autistic children when viewing images [54] or autistic adults when searching information in the

web [57]. Few incorporate multimodal sensing; *e.g.* Rehg *et al.* [140] include wrist-worn sensors recording electrodermal activity and accelerometry. Otherwise, the default is using a stationary eye tracker attached to a display with engineered visual stimuli (*e.g.*, Tobii T120 in [54]). However, many comment on the limitations of this approach, especially when involving autistic children [54]. Thus, some attempt to obtain eye contact estimation through egocentric images of the child’s face from the experimenter’s viewpoint [140].

**4.3.5 Video.** This format (in 48 datasets) captures videos of people signing [28], talking [45], walking [14], describing pictures [50], or interacting with technology (*e.g.*, rehabilitation robots [53]). The majority (41) of these datasets fall in the Hearing group. Typically, they capture sign language through multiple cameras placed at a fixed location (*e.g.* [130]). 3D Kinect videos are also common (14). These datasets started appearing after 2010, the year Kinect was launched. Looking at the dataset distributions over time in Figures 3 and 6, we see that datasets in the Hearing group tend to reflect the availability of research instruments and computational processes enabling fine-grained analysis of body movements and facial expressions transitioning from video and motion capture to videos only. We see a similar trend in Mobility, where datasets typically involve motion sensors but now look at pose estimation in videos (*e.g.*, [110]).

**4.3.6 Motion.** This format (in 35 datasets) mostly covers data using motion sensors typically embedded in devices such as smartphones to measure hand and walking movements of individuals with mobility impairments affected by neurological disorders such as Parkinson’s disease [20] or spinal cord injury [166]. Thus, the majority (15) of the datasets fall under the Mobility group. Different approaches are used to collect such data over the last decade; some opt for smartphone-mounted hand gloves to quantify tremor symptoms [103] others use touchscreen tablets to collect stroke-gesture input [166]. Initially researchers relied on customized motion sensors such as accelerometers and gyroscopes attached in footwear [101] or shoes with force-sensitive resistors measuring foot contact [82]. This format is also present in some (14) datasets in the Hearing group, where motion capture gloves and other equipment is used to measure hand and finger positions [94, 111]. We see some (5) datasets also in the Health group capturing pen- and mouse-based interactions of older adults [127] and daily activity movements of people with unipolar or bipolar disorders through wearable sensors [69].

**4.3.7 Logs.** This format (in 26 datasets) typically captures user interactions in mobile applications [92, 147], keyboards [71], and search engines [172, 173]. We see this format in the Cognitive (11) and Mobility (9) groups with datasets being often cross-listed. The datasets typically include people with Parkinson’s. For example, a keystroke logging app, Tappy, was installed on participants’ personal computers at home to record key press and release timings for early detection of Parkinson’s [12]. Logs are also commonly found in the Autism community (6) *e.g.*, subjective measures on text complexity by autistic people [60], which were often combined with eye tracking measurements for text readability [178, 179].



**Figure 7: Distribution of sharing strategies employed across communities.**

## 4.4 Data Sharing Practices

Figure 7 shows how datasets from each community are distributed across strategies that dataset creators employed for sharing or not sharing data. Out of the 137 datasets that we analyzed, only 52 can be downloaded directly (*e.g.*, through a webpage from the dataset creators [122]) and 27 are available upon request (*e.g.*, through registration of name, institution, and purpose of the use of data [108]). The remaining 58 do not include any sharing intent or information; however, we still link to them as we have observed a lack of standardized process for documenting datasets that require guessing or further inquires to understand their intent of sharing. More often, authors choose to provide links or request information for the datasets in some footnote on manuscripts or supplementary materials (*e.g.*, [72]) without specifying data access policy, making it challenging to discover their motivation for sharing data resources.

**4.4.1 Publicly Available Accessibility Datasets.** Datasets in this category can be directly downloaded from personal and project-specific websites (26); repositories like Kaggle (4), UCI Machine Learning Repository (3), and PhysioNet (5); OrtoLang (3); Zenodo (3); Synapse.org (1); and Open Science Foundation (1). This strategy was most commonly found across datasets from the Hearing group (23), mostly related to sign language videos and gloss annotations. The majority of them were shared by computational linguists and computer vision researchers, often having their personal or project site for documentation but the information provided varied in its quality and detail. Dataset creators from [130] documented an overview of participants contributing data, data types and size, and annotation of datasets, as well as a detailed explanation on how the data was collected, processed, and analyzed. Providing such documentation did not seem to be a standard practice; often the creators provide a short description of the dataset on their personal/project webpage and cite their related publications for more detail (*e.g.*, [97]). The strategy of direct download was also common among datasets sourced from people who are blind or have low vision (7). They were similarly hosted on personal/project websites sharing their photos (*e.g.*, [76]), touchscreen gestures (*e.g.*, [165]), and walking patterns (*e.g.*, [65]).

We also see the Mobility (13) and Cognitive (11) groups contributing publicly available datasets *e.g.*, providing keystroke logs

collected from users with and without Parkinson’s disease [12], where 8 of these datasets are cross-listed in these two groups. This community seems to leverage public data repositories more often than personal websites to host their datasets, specifically Kaggle (4) and PhysioNet (4), documenting data purpose and description, along with instructions on how to use the data properly. In a few occasions, they link additional data resources (e.g., analysis tool [53]). Relevant acknowledgments and citations are often provided on the repository. Unlike Kaggle that does not specifically target disability data, PhysioNet [73] is specific to the community for donating and searching biomedical research data and software. It provides a standardized process for documenting datasets, including data access policy and license, and discovery of related publications (e.g., [81]).

**4.4.2 Datasets Shared Upon Request.** Datasets in this category can be accessed only upon request through specific procedures. The most common practice we observe is to have a dedicated dataset webpage with a note to contact one of the authors (typically the Project Investigator) given an email address without any further details on eligibility or process. Another practice is to describe the license agreement, the requirements to obtain the data, as well as the types of data that would be shared. This information was either included on the project webpage or included on a dedicated section of the publication where the data were introduced, named *Distribution*. For example, in the BosphorusSign dataset [34] this section reads: “*The collected corpus will be available to download for academic purposes upon filling a license agreement available from the BosphorusSign website. The provided data will include...*”

Contributing datasets in this category come from the Cognitive (12) and Hearing (9) groups, similar to the representation of these communities in publicly-available datasets though less in number. The contrasting pattern found in datasets shared upon request is the appearance of sharing efforts from the Health (6), Speech (5), and Language (3) groups. These datasets were often hosted on TalkBank [116], a data sharing platform for research associated with human communication. It provides separate repositories for different research areas including DementiaBank [21] and AphasiaBank [117]. Datasets studying communication of patients with Primary Progressive Aphasia (e.g., [50]) provide audio and video data of the discourse on AphasiaBank. To access data, the TalkBank system requires membership registration, and clinical data such as those on AphasiaBank are restricted to faculty members or require permission from those who are already a member of the system.

**4.4.3 Non-Shared Datasets.** In our collection, we see that this non-sharing strategy is not unique to a specific community; it is prevalent across data from different user groups including Vision, Hearing, Cognitive, Speech, and Mobility, as well as Autism. More so, we observe that all of our datasets from the Developmental and Learning groups follow this strategy, which encompass the community of what is called “invisible disabilities,” disabilities that are less apparent to others and perhaps more sensitive for disclosure. We also observe that children are often involved in these unshared datasets (e.g., [13, 109]), where parents have to agree to share the videos, images, or audio files that capture their behaviors. Given that the majority of human-computer interaction researchers that work with these populations do not share data [11], the lack of sharing strategies is not a surprise. Potential factors for not sharing

include data sensitivity, participant consent, and re-identification risks [168] as there are increased privacy concerns for accessibility data with the risk of disability disclosure.

**4.4.4 Datasets Cleared from Ethical Boards.** Compliance with human subjects research standards and requirements can play an important role in researchers’ data sharing practices [19, 125]. For instance, one cannot share data when the consent form was silent about data sharing or promised that the data would not be shared outside of the research team. Communicating the term “data” to participants can also be tricky within the space of the consent form. We review how many of the papers and resources associated with the datasets in our collection claim that the research has been cleared from institutional review boards (IRBs) or institutions imposing policies and restrictions on the collection and use of research data.

Out of 137 datasets, 49 reported to be approved by ethical boards, with the following breakdown: Download (22), Request (7), and Contact (20). Datasets in the Mobility (19) and Cognitive (25) groups, which contribute relatively more publicly available datasets than other, have also higher reporting rate of being cleared from ethical boards. Within these two groups, datasets with mentions of ethical clearance that are downloadable (13) constitute of more logs (7) and motion (7) data than respective non-shared datasets (13) – here, text (6) and audio (5) are found, often as a combination of speech and transcript data (e.g., [146]). This reflects the recurring concerns of contributing audio data associated with the risk of speaker identification as mentioned in one of the TalkBank user guidelines [10].

Interestingly, only a few (3) datasets in the Hearing group, which contributed most of the publicly available datasets in our collection, mentioned ethical clearance in their research work. This could be due to those contributing data not being participants but members of the research team or consultants. We yet note that non-reporting of ethical clearance in publications does not necessarily mean the work was not cleared from ethical boards. NSF-funded projects are required to obtain IRB approval before issuance of an award. The process of projects requesting to be hosted on TalkBank or PhysioNet also involves screening of IRB permission for data sharing. In support, these repositories provide informed consent and data confidentiality guidelines that match the ways they operate (i.e., data restricted to authorized members or open to anyone).

## 5 DISCUSSION

We were able to collect 137 datasets that represent populations from different communities in accessibility and aging. Across and within these communities of focus, we observed varying data collection, reporting, and sharing practices that changed over the last three decades with the adoption of technologies and growth of interest, participation, and openness in certain research domains. For example, active contributions from the Hearing group resulted from having a community of researchers from different disciplines often working towards sign language datasets for recognition and translation technologies. Such collaboration was also seen in communities performing clinical studies to collect data for early detection of cognitive or speech impairments. The advances in research instruments, such as Kinect camera for in-depth analysis of body movements or smart devices to collect longitudinal data, have also

widened the possibilities for researchers to pursue specific problems in that research space. Along with these research trends, certain communities have already established open data platforms (e.g., TalkBank, PhysioNet) to facilitate data sharing and reuse. However, the majority of datasets are still hosted or announced on personal or project-specific websites. On the other hand, communities that are less represented in data collection and sharing (e.g., Autism, Developmental, Learning) highlight the challenges associated with sharing based on parental consent and child assent as well as involuntary and inaccurate disclosure of visible and invisible disability.

**Locating Accessibility Datasets.** Our “detective” work highlights that accessibility datasets are difficult to locate and require domain and community knowledge, which can hinder accessible and inclusive AI innovations. This is partially due to inconsistent terminology. The lack of a common language between repositories and publications makes it difficult to find connections and to identify secondary data-reuse cases [99]. Such issues are more prominent for accessibility datasets, since rigid categorization of disability is difficult [27, 133] and perhaps a questionable task as conditions are vast and fluid [174]. There is additional ambiguity in how people describe those contributing data. For example, it is difficult to tell whether “subjects with cognitive decline” relate to those “with early Dementia” or whether “signers” actually indicate Deaf/deaf or hard-of-hearing data contributors. Datasets lack consistent descriptions and require manual screening and, on few occasions guessing or further inquiries, highlighting the importance of a standardized process for documenting datasets (e.g. [70]). More often, links or request information for the datasets are buried in some footnote or a specific section on manuscripts, making it challenging to discover. Even the Google Dataset search [33] released on September 2018, would leave us wanting for more. To enable broader dataset discovery and transparency, we launched a parallel thread to this work; one involving a data surfacing repository called IncluSet [91]. Of course, assuming that everyone will know about one more repository would defeat the purpose. Thus, we implemented the Google Schema (react-schemaorg), allowing IncluSet to surface accessibility datasets to broader search engines. In IncluSet, researchers don’t have to share their data just point to them; anyone coming across an accessibility dataset can point to it and have our team review it.

**Balancing Risks and Benefits.** In this work we highlight the benefits of creating and sharing accessibility datasets. However, there are many privacy and ethical concerns associated with such practices as people who have distinct data patterns may be more susceptible to data abuse and misuse [11, 75, 78, 160]. Naturally, the risk of deduction [11] can increase when reporting data on smaller populations. For instance, blind participants’ age, gender, visual acuity, onset, and mobility aids, typically reported in navigation studies [93], combined with researchers’ location, may reveal their identities to those living in the area. Given that some disability communities can be really small, the effectiveness of privacy-preserving techniques can also be affected [128], calling for novel approaches (e.g. privacy-enhancing distortions on sign language datasets [31]). Even when re-identification is not a risk, consent and disclosure can be, as disability status is sensitive. We believe that by sharing accessibility datasets we can attract, nurture, and challenge data scientists and technologists to include people with disabilities and older adults at the forefront of AI innovations. However,

the same datasets that are collected to mitigate bias against people with disabilities or to support them through novel AI-infused assistive tech, can be used against them by “detecting” their disabilities. This can happen even when disclosure is not voluntary, posing further discrimination risks e.g., for one’s healthcare and employment [174]. Thus, we call for better sharing practices as well as technical, legal, and institutional privacy frameworks that are more attuned to concerns from these communities e.g., risks of inaccurate or non-consenting disclosure of a disability. We hope that researchers utilize our data and insights to reflect and discuss with others the future of data sharing and ownership in accessibility research and other field that are often connected through interdisciplinary efforts. We note that our review is not a call to include underrepresented communities, that we aim to benefit, in models that follow rigid categorization that can pose risks for non voluntary disability disclosure. On the contrary, we are hoping it will help us better understand sharing practices and potential concerns that can feed into the conversations to follow.

## 6 CONCLUSION

Datasets directly sourced from underrepresented communities such as people with disabilities and older adults can contribute to more inclusive AI applications as well as innovative assistive technologies. However, they are scarce. In this paper, we reflect on the data collection and sharing practices for accessibility datasets across a vast number of disciplines for the past 35 years. While not an exhaustive search, as dataset search is inherently a challenging task, our analysis has implications for our fields’ position in current practices, and where we should go from here. Specifically, we contribute a deep understanding of the current status of accessibility datasets (1984-2020) in terms of their distribution across communities represented, data collection purpose, and language used to describe the people who contributed with data. We report trends across data size in terms of the number of people involved referring to both those representing communities of focus and those serving as proxies or control. We explore how data types relate to communities and the purpose of the data collection. More importantly, we identify common data sharing practices and report on clearance from ethical boards.

## ACKNOWLEDGMENTS

We thank Sravya Amancherla, Mayanka Jha, Riya Chanduka, and Amnah Mahmood for their contributions in collecting and coding the datasets reported in this paper. We also thank our anonymous reviewers for further strengthening this paper. This work is supported by National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR), ACL, HHS (#90REGE0008).

## REFERENCES

- [1] 2021. ACL: Association for Computational Linguistics. ACL Data and Code Repository. [https://aclweb.org/aclwiki/ACL\\_Data\\_and\\_Code\\_Repository](https://aclweb.org/aclwiki/ACL_Data_and_Code_Repository).
- [2] 2021. ACM: Association for Computing Machinery. <https://www.acm.org/>.
- [3] 2021. Amazon. Registry of Open Data on AWS. <https://registry.opendata.aws/>.
- [4] 2021. CVF: Computer Vision Foundation. <https://www.thecvf.com/>.
- [5] 2021. Google Search. <https://www.google.com/search/howsearchworks/>.
- [6] 2021. IEEE: Institute of Electrical and Electronics Engineers. <https://www.ieee.org/>.
- [7] 2021. ISCA: International Speech Communication Association. <https://www.isca-speech.org/iscaweb/index.php>.

- [8] 2021. LREC: Language Resources and Evaluation. <http://www.lrec-conf.org/>.
- [9] 2021. Microsoft. Microsoft Research Open Data. <https://msropendata.com/>.
- [10] 2021. TalkBank IRB Approval. <https://talkbank.org/share/irb/>.
- [11] Jacob Abbott, Haley MacLeod, Novia Nurain, Gustave Ekobe, and Sameer Patil. 2019. Local Standards for Anonymization Practices in Health, Wellness, Accessibility, and Aging Research at CHI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery (ACM), 1–14. <https://doi.org/10.1145/3290605.3300692>
- [12] Warwick R Adams. 2017. High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing. *PLOS ONE* 12, 11 (11 2017), 1–20. <https://doi.org/10.1371/journal.pone.0188226>
- [13] Gaurav Aggarwal and Latika Singh. 2018. Evaluation of Supervised Learning Algorithms Based on Speech Features as Predictors of the Diagnosis of Mild to Moderate Intellectual Disability. *3D Research* 9, 4 (2018), 55. <https://doi.org/10.1007/s13319-018-0207-6>
- [14] Dragan Ahmetovic, Uran Oh, Sergio Mascetti, and Chieko Asakawa. 2018. Turn Right: Analysis of Rotation Errors in Turn-by-Turn Navigation for Individuals with Visual Impairments. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '18)*. Association for Computing Machinery (ACM), 333–339. <https://doi.org/10.1145/3234695.3236363>
- [15] Meghan Allen, Joanna McGrenere, and Barbara Purves. 2007. The Design and Field Evaluation of PhotoTalk: A Digital Image Communication Application for People with Aphasia. In *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '07)*. Association for Computing Machinery (ACM), 187–194. <https://doi.org/10.1145/1296843.1296876>
- [16] Gail A Alvares, Keely Bebbington, Dominique Cleary, Kiah Evans, Emma J Glasson, Murray T Maybery, Sarah Pillar, Mirko Uljarević, Kandice Varcin, John Wray, et al. 2020. The misnomer of 'high functioning autism': Intelligence is an imprecise predictor of functional abilities at diagnosis. *Autism* 24, 1 (2020), 221–232. <https://doi.org/10.1177/1362361319852831>
- [17] Pramod Anantharam, Krishnaprasad Thirunarayan, Vahid Taslimi, and Amit P Sheth. 2013. Predicting parkinson's disease progression with smartphone data. (2013). <https://corescholar.libraries.wright.edu/knoesis/569/>
- [18] Kevin Bache and Moshe Lichman. 2013. UCI Machine Learning Repository. <http://archive.ics.uci.edu/ml>
- [19] Mark Barnes, Judith Carrithers, and Jeremy Sugarman. 2020. Ethical and Practical Concerns about IRB Restrictions on the Use of Research Data. *Ethics & Human Research* 42, 6 (2020), 29–34. <https://doi.org/10.1002/eahr.500072>
- [20] O. Bazgir, J. Frounchi, S. A. H. Habibi, L. Palma, and P. Pierleoni. 2015. A neural network system for diagnosis and assessment of tremor in parkinson disease patients. In *Proceedings of the 22nd Iranian Conference on Biomedical Engineering (ICBME)*, 1–5. <https://doi.org/10.1109/ICBME.2015.7404105>
- [21] James T Becker, François Boiler, Oscar L Lopez, Judith Saxton, and Karen L McGonigle. 1994. The natural history of Alzheimer's disease: description of study cohort and accuracy of diagnosis. *Archives of neurology* 51, 6 (1994), 585–594. <https://doi.org/10.1001/archneur.1994.00540180063015>
- [22] Daniela Beltrami, Laura Calzà, Gloria Gagliardi, Enrico Ghidoni, Norina Marcello, Rema Rossini Favretti, and Fabio Tamburini. 2016. Automatic identification of mild cognitive impairment through the analysis of Italian spontaneous speech productions. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC '16)*, 2086–2093. <https://aclanthology.org/L16-1331>
- [23] Erwan Bezard, Jonathan M Brotchie, and Christian E Gross. 2001. Pathophysiology of levodopa-induced dyskinesia: potential for new therapies. *Nature Reviews Neuroscience* 2, 8 (2001), 577–588. <https://doi.org/10.1038/35086062>
- [24] Jeffrey P. Bigham, Anna C. Cavender, Jeremy T. Brudvik, Jacob O. Wobbrock, and Richard E Ladner. 2007. WebinSitu: A Comparative Analysis of Blind and Sighted Browsing Behavior. In *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '07)*. Association for Computing Machinery (ACM), 51–58. <https://doi.org/10.1145/1296843.1296854>
- [25] Jeffrey P. Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samuel White, and Tom Yeh. 2010. VizWiz: Nearly Real-Time Answers to Visual Questions. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. Association for Computing Machinery (ACM), 333–342. <https://doi.org/10.1145/1866029.1866080>
- [26] Ann Blandford. 2019. HCI for health and wellbeing: Challenges and opportunities. *International Journal of Human-Computer Studies* 131 (2019), 41–51. <https://doi.org/10.1016/j.ijhcs.2019.06.007>
- [27] Brianna Blaser and Richard E Ladner. 2020. Why is Data on Disability so Hard to Collect and Understand?. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Vol. 1. IEEE, 1–8. [https://www.washington.edu/doit/sites/default/files/atoms/files/RESPECT\\_2020\\_DisabilityData.pdf](https://www.washington.edu/doit/sites/default/files/atoms/files/RESPECT_2020_DisabilityData.pdf)
- [28] Julian Bleicken, Thomas Hanke, Uta Salden, and Sven Wagner. 2016. Using a Language Technology Infrastructure for German in order to Anonymize German Sign Language Corpus Data. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC '16)*. European Language Resources Association (ELRA), 3303–3306. <https://aclanthology.org/L16-1526>
- [29] Brian M. Bot, Christine Suver, Elias Chaibub Neto, Michael Kellen, Arno Klein, Christopher Bare, Megan Doerr, Abhishek Prapat, John Wilbanks, E. Ray Dorsey, Stephen H. Friend, and Andrew D Trister. 2016. The mPower study, Parkinson disease mobile data collected using ResearchKit. *Scientific Data* 3 (March 2016), 160011. <https://doi.org/10.1038/sdata.2016.11>
- [30] Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreault, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler, and Meredith Ringel Morris. 2019. Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. Association for Computing Machinery (ACM), 16–31. <https://doi.org/10.1145/3308561.3353774>
- [31] Danielle Bragg, Oscar Koller, Naomi Caselli, and William Thies. 2020. Exploring Collection of Sign Language Datasets: Privacy, Participation, and Model Performance. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20)*. Association for Computing Machinery (ACM), Article 33, 14 pages. <https://doi.org/10.1145/3373625.3417024>
- [32] Hannah Bull, Annelies Braffort, and Michèle Gouffès. 2020. MEDI-API-SKEL - A 2D-Skeleton Video Database of French Sign Language With Aligned French Subtitles. In *Proceedings of the 12th Language Resources and Evaluation Conference (LREC '20)*, 6063–6068. <https://aclanthology.org/2020.lrec-1.743>
- [33] Matthew Burgess and Natasha Noy. 2018. Building Google Dataset Search and Fostering an Open Data Ecosystem. (Sept. 2018). <https://ai.googleblog.com/2018/09/building-google-dataset-search-and.html>
- [34] Necati Cihan Camgöz, Ahmet Alp Kandroğlu, Serpil Karabükü, Meltem Keleşir, Ayşe Sumru Özsoy, and Lale Akarun. 2016. BosphorusSign: A Turkish Sign Language Recognition Corpus in Health and Finance Domains. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC '16)*. European Language Resources Association (ELRA), 1383–1388. <https://www.aclweb.org/anthology/L16-1220>
- [35] Romuald Carette, Mahmoud Elbattah, Federica Cilia, Gilles Dequen, Jean-Luc Guerin, and Jérôme Bosche. 2019. Learning to Predict Autism Spectrum Disorder based on the Visual Patterns of Eye-tracking Scanpaths. In *Proceedings of the 12th International Conference on Health Informatics*, 103–112. <https://doi.org/10.5220/0007402601030112>
- [36] Ugo Cesari, Giuseppe De Pietro, Elio Marciano, Ciro Niri, Giovanna Sannino, and Laura Verde. 2018. A new database of healthy and pathological voices. *Computers & Electrical Engineering* 68 (May 2018), 310–321. <https://doi.org/10.1016/j.compeleceng.2018.04.008>
- [37] Mary Chamie. 1989. Survey design strategies for the study of disability. *World health statistics quarterly* 42, 3 (1989), 122–140. <https://apps.who.int/iris/handle/10665/48609>
- [38] Adriane Chapman, Elena Simperl, Laura Koesten, George Konstantinidis, Luis-Daniel Ibáñez, Emilia Kacprzak, and Paul Groth. 2020. Dataset search: a survey. *The VLDB Journal* 29, 1 (2020), 251–272. <https://doi.org/10.1007/s00778-019-00564-x>
- [39] Dae-Lim Choi, Bong-Wan Kim, Yeon-Whoo Kim, Yong-Ju Lee, Yongnam Um, and Minhwa Chung. 2012. Dysarthric Speech Database for Development of QoLT Software Technology. In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC '12)*. European Language Resources Association (ELRA), Istanbul, Turkey, 3378–3381. [http://www.lrec-conf.org/proceedings/lrec2012/pdf/388\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2012/pdf/388_Paper.pdf)
- [40] R.P. Clapham, L. van der Molen, R.J.J.H. van Son, M. van den Brekel, and F.J.M. Hilgers. 2012. NKI-CCRT Corpus - Speech Intelligibility Before and After Advanced Head and Neck Cancer Treated with Concomitant Chemoradiotherapy. In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC '12)*. European Language Resources Association (ELRA), Istanbul, Turkey, 3350–3355. [http://www.lrec-conf.org/proceedings/lrec2012/pdf/230\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2012/pdf/230_Paper.pdf)
- [41] US Congress. 2017. Individuals with Disabilities Education Act. <https://sites.ed.gov/idea/statuteregulations/>
- [42] US Congress. 2017. Sec. 300.8 (c) (1) (i) - Individuals with Disabilities Education Act. <https://sites.ed.gov/idea/reg/b/a/300.8/c/1/i>
- [43] US Congress. 2018. Sec. 300.8 Child with a disability - Individuals with Disabilities Education Act. <https://sites.ed.gov/idea/reg/b/a/300.8>
- [44] HM Cooper, Eng-Jon Ong, Nicolas Pugeault, and Richard Bowden. 2012. Sign language recognition using sub-units. *Journal of Machine Learning Research* 13 (2012), 2205–2231. [https://doi.org/10.1007/978-3-319-57021-1\\_3](https://doi.org/10.1007/978-3-319-57021-1_3)
- [45] Mario Corrales-Astorgano, David Escudero-Mancebo, Yurena Gutiérrez-González, Valle Flores-Lucas, César González-Ferreras, and Valentín Cardeñoso-Payo. 2016. On the Use of a Serious Game for Recording a Speech Corpus of People with Intellectual Disabilities. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC '16)*. European Language Resources Association (ELRA), 2094–2099. <https://aclanthology.org/L16-1332>
- [46] Antonio R Damasio. 1992. Aphasia. *New England Journal of Medicine* 326, 8 (1992), 531–539. <https://doi.org/10.1056/NEJM199202203260806>

- [47] Frederic L Darley, Arnold E Aronson, and Joe R Brown. 1969. Differential diagnostic patterns of dysarthria. *Journal of speech and hearing research* 12, 2 (1969), 246–269. <https://doi.org/10.1044/jshr.1202.246>
- [48] Burn Model System National Data and Statistical Center. 1994. Burn Model System: Advancing recovery through knowledge. <http://burndata.washington.edu/>
- [49] JR Deller Jr, MS Liu, LJ Ferrier, and P Robichaud. 1993. The Whitaker database of dysarthric (cerebral palsy) speech. *The Journal of the Acoustical Society of America* 93, 6 (1993), 3516–3518. <https://doi.org/10.1121/1.405684>
- [50] Roxanne DePaul. 2016. DementiaBank English PPA Corpus. <https://doi.org/10.21415/TsZH5T>
- [51] Carol C. Diamond, Farzad Mostashari, and Clay Shirky. 2009. Collecting And Sharing Data For Population Health: A New Paradigm. *Health Affairs* 28, 2 (2009), 454–466. <https://doi.org/10.1377/hlthaff.28.2.454>
- [52] Chaohai Ding, Mike Wald, and Gary Wills. 2014. A Survey of Open Accessibility Data. In *Proceedings of the 11th Web for All Conference (W4A '14)*. Association for Computing Machinery (ACM), Article 37, 4 pages. <https://doi.org/10.1145/2596695.2596708>
- [53] Elham Dolatabadi, Ying Xuan Zhi, Bing Ye, Marge Coahran, Giorgia Lupinacci, Alex Mihailidis, Rosalie Wang, and Babak Taati. 2017. The Toronto Rehab Stroke Pose Dataset to Detect Compensation during Stroke Rehabilitation Therapy. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '17)*. Association for Computing Machinery (ACM), 375–381. <https://doi.org/10.1145/3154862.3154925>
- [54] Huiyu Duan, Guangtao Zhai, Xiongkuo Min, Zhaohui Che, Yi Fang, Xiaokang Yang, Jesús Gutiérrez, and Patrick Le Callet. 2019. A Dataset of Eye Movements for the Children with Autism Spectrum Disorder. In *Proceedings of the 10th ACM Multimedia Systems Conference (Amherst, Massachusetts) (MMSys '19)*. Association for Computing Machinery (ACM), New York, NY, USA, 255–260. <https://doi.org/10.1145/3304109.3325818>
- [55] Sarah Ebling, Necati Cihan Camgöz, Penny Boyes Braem, Katja Tissi, Sandra Sidler-Miserez, Stephanie Stoll, Simon Hadfield, Tobias Haug, Richard Bowden, Sandrine Tornay, Marzieh Razavi, and Mathew Magimai-Doss. 2018. SMILE Swiss German Sign Language Dataset. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC '18)*. European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/L18-1666>
- [56] Eleni Efthimiou and Stavroula-Evita Fotinea. 2007. GSLC: Creation and Annotation of a Greek Sign Language Corpus for HCI. In *Proceedings of the 4th International Conference on Universal Access in Human Computer Interaction: Coping with Diversity (UAHCI'07)*. Springer-Verlag, 657–666. [https://doi.org/10.1007/978-3-540-73279-2\\_73](https://doi.org/10.1007/978-3-540-73279-2_73)
- [57] Sukru Eraslan, Victoria Yaneva, Yeliz Yesilada, and Simon Harper. 2019. Web users with autism: eye tracking evidence for differences. *Behaviour & Information Technology* 38, 7 (2019), 678–700. <https://doi.org/10.1080/0144929X.2018.1551933>
- [58] Sukru Eraslan, Yeliz Yesilada, Victoria Yaneva, and Simon Harper. 2020. Autism Detection Based on Eye Movement Sequences on the Web: A Scanpath Trend Analysis Approach. In *Proceedings of the 17th International Web for All Conference (W4A '20)*. Association for Computing Machinery (ACM), Article 11, 10 pages. <https://doi.org/10.1145/3371300.3383340>
- [59] European Organization For Nuclear Research and OpenAIRE. 2013. Zenodo. <https://doi.org/10.25495/7GXK-RD71>
- [60] Richard Evans, Victoria Yaneva, and Irina Temnikova. 2016. Predicting reading difficulty for readers with autism spectrum disorder. In *Proceedings of 10th International Conference on Language Resources and Evaluation (LREC '16)*. European Language Resources Association. <https://wlv.openrepository.com/handle/2436/622447>
- [61] Benedikt Fecher, Sascha Friesike, and Marcel Hebing. 2015. What Drives Academic Data Sharing? *PLOS One* 10, 2 (02 2015), 1–25. <https://doi.org/10.1371/journal.pone.0118053>
- [62] Jie Feng. 2018. VisualData - Search Engine for Computer Vision Datasets. <https://www.visualdata.io/discovery>
- [63] Lijun Feng. 2009. Automatic Readability Assessment for People with Intellectual Disabilities. *SIGACCESS - Accessible Computing* 93 (Jan. 2009), 84–91. <https://doi.org/10.1145/1531930.1531940>
- [64] Quentin Ferry, Julia Steinberg, Caleb Webber, David R FitzPatrick, Chris P Ponting, Andrew Zisserman, and Christoffer Nellaker. 2014. Diagnostically relevant facial gestalt information from ordinary photos. *Elife* 3 (2014), e02020. <https://doi.org/10.7554/eLife.02020>
- [65] German H. Flores and Roberto Manduchi. 2016. WeAllWalk: An Annotated Data Set of Inertial Sensor Time Series from Blind Walkers. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (Reno, Nevada, USA) (ASSETS '16)*. Association for Computing Machinery, New York, NY, USA, 141–150. <https://doi.org/10.1145/2982142.2982179>
- [66] Association for Computing Machinery (ACM). 2020. Artifact Review and Badging Version 1.1. <https://www.acm.org/publications/policies/artifact-review-and-badging-current>.
- [67] Cécile Fougerson, Lise Crevier-Buchman, Corinne Fredouille, Alain Ghio, Christine Meunier, Claude Chevré-Muller, Nicolas Audibert, Jean-François Bonastre, Antonia Colazo-Simon, Céline Delooue, Danielle Duez, Cédric Gendrot, Thierry Legou, Nathalie Lévêque, Claire Pillot-Loiseau, Serge Pinto, Gilles Pouchoulin, Danièle Robert, Jacqueline Vaissière, François Viallet, and Coralie Vincent. 2010. Developing an acoustic-phonetic characterization of dysarthric speech in French. In *Proceedings of the 7th International Conference on Language Resources, Technologies and Evaluation (LREC '10, Vol. 1)*. European Language Resources Association (ELRA) (Ed.), 2831–2838. <https://hal.archives-ouvertes.fr/hal-00528520>
- [68] Kathleen C. Fraser, Nicklas Linz, Hali Lindsay, and Alexandra König. 2019. The importance of sharing patient-generated clinical speech and language data. In *Proceedings of the 6th Workshop on Computational Linguistics and Clinical Psychology*. Association for Computational Linguistics, 55–61. <https://doi.org/10.18653/v1/W19-3007>
- [69] Enrique García-Ceja, Michael Riegler, Petter Jakobsen, Jim Tørresen, Tine Nordgreen, Ketil J. Oedegaard, and Ole Bernt Fasmer. 2018. Depresjon: A Motor Activity Database of Depression Episodes in Unipolar and Bipolar Patients. In *Proceedings of the 9th ACM Multimedia Systems Conference (MMSys '18)*. Association for Computing Machinery (ACM), 472–477. <https://doi.org/10.1145/3204949.3208125>
- [70] Timmit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna M. Wallach, Hal Daumé III, and Kate Crawford. 2018. Datasheets for Datasets. *CoRR abs/1803.09010* (2018). [arXiv:1803.09010](http://arxiv.org/abs/1803.09010)
- [71] Luca Giancardo, Alvaro Sanchez-Ferro, Teresa Arroyo-Gallego, Ian Butterworth, Carlos S Mendoza, Paloma Montero, Michele Matarazzo, José A Obeso, Martha L Gray, and R San José Estépar. 2016. Computer keyboard interaction as an indicator of early Parkinson's disease. *Scientific reports* 6 (2016), 34468. <https://doi.org/10.1038/srep34468>
- [72] George Gkotsis, Anika Oellrich, Sumithra Velupillai, Maria Liakata, Tim J. P. Hubbard, Richard J. B. Dobson, and Rina Dutta. 2017. Characterisation of mental health conditions in social media using Informed Deep Learning. *Scientific Reports* 7 (March 2017), 45141. <https://doi.org/10.1038/srep45141>
- [73] Ary L Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H. Eugene Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101, 23 (2000), e215–e220. <https://doi.org/10.1161/01.cir.101.23.e215>
- [74] Roy Grant and Molly Nozyce. 2013. Proposed changes to the American Psychiatric Association diagnostic criteria for autism spectrum disorder: Implications for young children and their families. *Maternal and child health journal* 17, 4 (2013), 586–592. <https://doi.org/10.1007/s10995-013-1250-9>
- [75] Anhong Guo, Ece Kamar, Jennifer Wortman Vaughan, Hanna Wallach, and Meredith Ringel Morris. 2020. Toward Fairness in AI for People with Disabilities: A Research Roadmap. *SIGACCESS - Accessible Computing* 125, Article 2 (March 2020), 1 pages. <https://doi.org/10.1145/3386296.3386298>
- [76] D. Gurari, Q. Li, C. Lin, Y. Zhao, A. Guo, A. Stangl, and J. P. Bigham. 2019. VizWiz-Priv: A Dataset for Recognizing the Presence and Purpose of Private Visual Information in Images Taken by Blind People. In *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR '19)*. 939–948. <https://doi.org/10.1109/CVPR.2019.00103>
- [77] Danna Gurari, Qing Li, Abigale J. Stangl, Anhong Guo, Chi Lin, Kristen Grauman, Jiebo Luo, and Jeffrey P Bigham. 2018. VizWiz Grand Challenge: Answering Visual Questions from Blind People. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (Jun 2018). <https://doi.org/10.1109/cvpr.2018.00380>
- [78] Foad Hamidi, Kellie Ponerer, Aaron Massey, and Amy Hurst. 2018. Who Should Have Access to My Pointing Data? Privacy Tradeoffs of Adaptive Assistive Technologies. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '18)*. Association for Computing Machinery, 203–216. <https://doi.org/10.1145/3234695.3239331>
- [79] Vicki L Hanson, Anna Cavender, and Shari Trewin. 2015. Writing about accessibility. *Interactions* 22, 6 (2015), 62–65. <https://interactions.acm.org/archive/november-december-2015/writing-about-accessibility>
- [80] Saad Hassan, Larwan Berke, Elahe Vahdani, Longlong Jing, Yingli Tian, and Matt Huenerfauth. 2020. An Isolated-Signing RGBD Dataset of 100 American Sign Language Signs Produced by Fluent ASL Signers. In *Proceedings of the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives*. European Language Resources Association (ELRA), 89–94. <https://www.aclweb.org/anthology/2020.signlang-1.14>
- [81] Jeffrey M Hausdorff, Apinya Lertratanakul, Merit E Cudkovicz, Amie L Peterson, David Kaliton, and Ary L Goldberger. 2000. Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis. *Journal of applied physiology* (2000). <https://doi.org/10.1152/jappl.2000.88.6.2045>
- [82] Jeffrey M Hausdorff, Susan L Mitchell, Renee Firtion, Chung-Kang Peng, Merit E Cudkovicz, Jeanne Y Wei, and Ary L Goldberger. 1997. Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. *Journal of applied physiology* 82, 1 (1997), 262–269. <https://doi.org/10.1152/jappl.1997.82.1.262>

- [83] Helen Keller Services (HKS). 2020. Introduction to Deaf-Blindness. <https://www.helenkeller.org/hknc/lesson/introduction-deaf-blindness>.
- [84] A.D. Hoover, Valentina Kouznetsova, and Michael Goldbaum. 2000. Locating Blood Vessels in Retinal Images by Piecewise Threshold Probing of a Matched Filter Response. *IEEE transactions on medical imaging* 19 (04 2000), 203–10. <https://doi.org/10.1109/42.845178>
- [85] Matt Huenerfauth and Hernisa Kacorri. 2014. Release of experimental stimuli and questions for evaluating facial expressions in animations of American Sign Language. In *Proceedings of the 6th Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel, The 9th International Conference on Language Resources and Evaluation (LREC '14)*. [http://dx.doi.org/10.1007/978-3-642-39188-0\\_55](http://dx.doi.org/10.1007/978-3-642-39188-0_55)
- [86] Kaggle Inc. 2017. Kaggle: Your Machine Learning and Data Science Community. <https://www.kaggle.com>.
- [87] Ronnachai Jaroensri, Amy Zhao, Guha Balakrishnan, Derek Lo, Jeremy D. Schmahmann, Fredo Durand, and John Guttag. 2017. A Video-Based Method for Automatically Rating Ataxia. 68 (18–19 Aug 2017), 204–216. <http://proceedings.mlr.press/v68/jaroensri17a.html>
- [88] Eun Seo Jo and Timnit Gebru. 2020. Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT\* '20)*. Association for Computing Machinery (ACM), 306–316. <https://doi.org/10.1145/3351095.3372829>
- [89] Hernisa Kacorri. 2016. *Data-Driven Synthesis and Evaluation of Syntactic Facial Expressions in American Sign Language Animation*. Ph.D. Dissertation. CUNY Academic Works. [https://academicworks.cuny.edu/gc\\_etds/1375](https://academicworks.cuny.edu/gc_etds/1375)
- [90] Hernisa Kacorri. 2017. Teachable Machines for Accessibility. *SIGACCESS - Accessible Computing* 119, 10–18. <https://doi.org/10.1145/3167902.3167904>
- [91] Hernisa Kacorri, Utkarsh Dwivedi, Sravya Amancherla, Mayanka Jha, and Riya Chanduka. 2020. *IncluSet: A Data Surfacing Repository for Accessibility Datasets*. Association for Computing Machinery (ACM). <https://doi.org/10.1145/3373625.3418026>
- [92] Hernisa Kacorri, Sergio Mascetti, Andrea Gerino, Dragan Ahmetovic, Hironobu Takagi, and Chieko Asakawa. 2016. Supporting Orientation of People with Visual Impairment: Analysis of Large Scale Usage Data. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '16)*. Association for Computing Machinery (ACM), 151–159. <https://doi.org/10.1145/2982142.2982178>
- [93] Hernisa Kacorri, Eshed Ohn-Bar, Kris M. Kitani, and Chieko Asakawa. 2018. Environmental Factors in Indoor Navigation Based on Real-World Trajectories of Blind Users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery (ACM), 1–12. <https://doi.org/10.1145/3173574.3173630>
- [94] Mohammed Waleed Kadous. 2002. *Temporal Classification: Extending the Classification Paradigm to Multivariate Time Series*. Ph.D. Dissertation. AUS.
- [95] Kaggle. 2016. Predicting Parkinson's Disease Progression with Smartphone Data. <https://www.kaggle.com/c/predicting-parkinson-s-disease-progression-with-smartphone-data>.
- [96] Soner Can KALKAN. 2018. Turkish Sign Language (Fingerspelling). <https://www.kaggle.com/feronial/turkish-sign-languagefingerspelling>
- [97] Tomasz Kapuscinski, Mariusz Oszust, Marian Wysocki, and Dawid Warchol. 2015. Recognition of Hand Gestures Observed by Depth Cameras. *International Journal of Advanced Robotic Systems* 12, 4 (2015), 36. <https://doi.org/10.5772/60091>
- [98] Daniel Kempler, Susan Curtiss, and Catherine Jackson. 1987. Syntactic preservation in Alzheimer's disease. *Journal of Speech, Language, and Hearing Research* 30, 3 (1987), 343–350. <https://doi.org/10.1044/jshr.3003.343>
- [99] Nushrat Khan, Catherine J. Pink, and Mike Thelwall. 2020. Identifying Data Sharing and Reuse with Scholix: Potentials and Limitations. *Patterns* 1, 1 (2020), 100007. <https://doi.org/10.1016/j.patter.2020.100007>
- [100] Heejin Kim, Mark Hasegawa-Johnson, Adrienne Perlman, Jon Gunderson, Thomas S Huang, Kenneth Watkin, and Simone Frame. 2008. Dysarthric speech database for universal access research. In *Proceedings of the 9th Annual Conference of the International Speech Communication Association*. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.139.4285>
- [101] Jochen Klucken, Jens Barth, Patrick Kugler, Johannes Schlachetzki, Thore Henze, Franz Marxreiter, Zacharias Kohl, Ralph Steidl, Joachim Hornegger, Bjoern Eskofier, et al. 2013. Unbiased and mobile gait analysis detects motor impairment in Parkinson's disease. *PLoS one* 8, 2 (2013), e56956. <https://doi.org/10.1371/journal.pone.0056956>
- [102] Bran Knowles, Vicki L. Hanson, Yvonne Rogers, Anne Marie Piper, Jenny Waycott, and Nigel Davies. 2019. HCI and Aging: Beyond Accessibility. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19)*. Association for Computing Machinery (ACM), 1–8. <https://doi.org/10.1145/3290607.3290925>
- [103] N. Kostikis, D. Hristu-Varsakelis, M. Arnaoutoglou, and C. Kotsavasiloglou. 2015. A Smartphone-Based Tool for Assessing Parkinsonian Hand Tremor. *IEEE Journal of Biomedical and Health Informatics* 19, 6 (2015), 1835–1842. <https://doi.org/10.1109/JBHI.2015.2471093>
- [104] Heidi Lam. 2008. A framework of interaction costs in information visualization. *IEEE transactions on visualization and computer graphics* 14, 6 (2008), 1149–1156. <https://doi.org/10.1109/TVCG.2008.109>
- [105] Ghazanfar Latif, Nazeeruddin Mohammad, Jaafar Alghazo, Roaa AlKhalaf, and Rawan AlKhalaf. 2019. Arasl: Arabic alphabets sign language dataset. *Data in brief* 23 (2019), 103777. <https://doi.org/10.17632/y7pckrw6z2.1>
- [106] Kyungjun Lee and Hernisa Kacorri. 2019. Hands Holding Clues for Object Recognition in Teachable Machines. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery (ACM), 1–12. <https://doi.org/10.1145/3290605.3300566>
- [107] Matthew L. Lee and Anind K Dey. 2007. Providing Good Memory Cues for People with Episodic Memory Impairment. In *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '07)*. Association for Computing Machinery (ACM), 131–138. <https://doi.org/10.1145/1296843.1296867>
- [108] Daniel Leightley, Moi Hoon Yap, Jessica Coulson, Yoann Barnouin, and Jamie S McPhee. 2015. Benchmarking human motion analysis using Kinect one: An open source dataset. In *Proceedings of the 2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA '15)*. IEEE, 1–7. <https://doi.org/10.1109/APSIPA.2015.7415438>
- [109] Marco Leo, Marco Del Coco, Pierluigi Carcagni, Cosimo Distanto, Massimo Bernava, Giovanni Pioggia, and Giuseppe Palestra. 2015. Automatic Emotion Recognition in Robot-Children Interaction for ASD Treatment. In *Proceedings of the 2015 IEEE International Conference on Computer Vision Workshop (ICCVW '15)*. IEEE Computer Society, 537–545. <https://doi.org/10.1109/ICCVW.2015.76>
- [110] Michael H Li, Tiago A Mestre, Susan H Fox, and Babak Taati. 2018. Vision-based assessment of parkinsonism and levodopa-induced dyskinesia with pose estimation. *Journal of neuroengineering and rehabilitation* 15, 1 (2018), 97. <https://doi.org/10.1186/s12984-018-0446-z>
- [111] Pengfei Lu and Matt Huenerfauth. 2012. CUNY American Sign Language Motion-Capture Corpus: First Release. In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC '12)*. <http://eniac.cs.cuny.edu/matt/pubs/lu-huenerfauth-2012-lrec.pdf>
- [112] Ian Lundberg, Arvind Narayanan, Karen Levy, and Matthew J. Salganik. 2019. Privacy, Ethics, and Data Access: A Case Study of the Fragile Families Challenge. *Socius* 5 (2019). <https://doi.org/10.1177/2378023118813023>
- [113] Nancy E Lundebjerg, Daniel E Trucil, Emily C Hammond, and William B Applegate. 2017. When it comes to older adults, language matters: Journal of the American Geriatrics Society adopts modified American Medical Association style.
- [114] Joakim Lustig. 2016. Identifying dyslectic gaze pattern: Comparison of methods for identifying dyslectic readers based on eye movement patterns. <http://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A955646>
- [115] Kelly Mack, Emma McDonnell, Dhruv Jain, Lucy Lu Wang, Jon E. Froehlich, and Leah Findlater. 2021. What Do We Mean by “Accessibility Research”? A Literature Survey of Accessibility Papers in CHI and ASSETS from 1994 to 2019. , Article 371 (2021), 18 pages. <https://doi.org/10.1145/3411764.3445412>
- [116] Brian MacWhinney, Steven Bird, Christopher Cieri, and Craig Martell. 2004. TalkBank: Building an open unified multimodal database of communicative interaction. In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC '04)*. Evaluations and Language Resources Distribution Agency, 525–528. <http://www.lrec-conf.org/proceedings/lrec2004/pdf/392.pdf>
- [117] Brian MacWhinney, Davida Fromm, Margaret Forbes, and Audrey Holland. 2011. AphasiaBank: Methods for studying discourse. *Aphasiology* 25, 11 (2011), 1286–1307. <https://dx.doi.org/10.1080%2F02687038.2011.589893>
- [118] Kiyoko Makimoto, Miyae Yamakawa, Nobuyuki Ashida, Younhee Kang, and Kyung-Rim Shin. 2006. Japan-Korea Joint Project on Monitoring People with Dementia. (01 2006). [https://www.researchgate.net/publication/237799013\\_Japan-Korea\\_Joint\\_Project\\_on\\_Monitoring\\_People\\_with\\_Dementia](https://www.researchgate.net/publication/237799013_Japan-Korea_Joint_Project_on_Monitoring_People_with_Dementia)
- [119] Manolis Maragoudakis and Euripides Loukis. 2011. MCMC Bayesian Inference for Heart Sounds Screening in Assistive Environments. In *Proceedings of the 4th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '11)*. Association for Computing Machinery (ACM), Article 16, 9 pages. <https://doi.org/10.1145/2141622.2141642>
- [120] Jason Markesich, Stacie Feldman, and William Rafferty. 2021. 2021 Compendium of Disability Data Collection Methods. (January 2021). <https://www.mathematica.org/download-media?MediaItemID={D29116BA-C4BE-4241-B6F9-D3D7F2981A7A}>
- [121] Daniele Mascali, Mauro DiNuzzo, Tommaso Gili, Marta Moraschi, Michela Fratini, Bruno Maraviglia, Laura Serra, Marco Bozzali, and Federico Giove. 2015. Resting-state fMRI in dementia patients. <https://doi.org/10.7910/DVN/29352>
- [122] Silke Matthes, Thomas Hanke, Anja Regen, Jakob Storz, Satu Worseck, Eleni Efthimiou, Athanasia-Lida Dimou, Annelies Braffort, John Glauert, and Eva Safar. 2012. Dicta-Sign—building a multilingual sign language corpus. In *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon (LREC '12)*. <https://www.sign-lang.uni-hamburg.de/lrec/lrec/pubs/12016.pdf>

- [123] X. Menendez-Pidal, J.B. Polikoff, S.M. Peters, J.E. Leonzio, and H.T. Bunnell. 1996. The Nemours database of dysarthric speech. In *Proceeding of 4th International Conference on Spoken Language Processing (ICSLP '96)*. 1962–1965 vol.3. <https://doi.org/10.1109/ICSLP.1996.608020>
- [124] Christine Meunier, Cecile Fougeron, Corinne Fredouille, Brigitte Bigi, Lise Crevier-Buchman, Elisabeth Delais-Roussarie, Laurianne Georgeton, Alain Ghio, Imed Laaridh, Thierry Legou, Claire Pillot-Loiseau, and Gilles Pouchoulin. 2016. The TYPALOC Corpus: A Collection of Various Dysarthric Speech Recordings in Read and Spontaneous Styles. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC '16)*. 4658–4665. <https://aclanthology.org/L16-1738>
- [125] Michelle N Meyer. 2018. Practical tips for ethical data sharing. *Advances in methods and practices in psychological science* 1, 1 (2018), 131–144. <https://doi.org/10.1177/2515245917747656>
- [126] Roman Čmejla, Hana Růžicková, Jiří Klempíř, Michal Novotný, Jan Ruzs, and Evžen Růžicka. 2016. Hypernasality associated with basal ganglia dysfunction: evidence from Parkinson's disease and Huntington's disease. *PeerJ* 4 (2016), e2530. <https://dx.doi.org/10.7717/2Fpeerj.2530>
- [127] Karyn Anne Moffatt. 2010. *Addressing age-related pen-based target acquisition difficulties*. Ph.D. Dissertation. University of British Columbia. <http://www.sigaccess.org/2010/01/addressing-age-related-pen-based-target-acquisition-difficulties/>
- [128] Meredith Ringel Morris. 2020. AI and Accessibility. *Commun. ACM* 63, 6, 35–37. <https://doi.org/10.1145/3356727>
- [129] Amit Morayosef and Yoav Goldberg. 2021. Sign Language Processing. <https://sign-language-processing.github.io/>
- [130] Carol Neidle, Ashwin Thangali, and Stan Sclaroff. 2012. Challenges in development of the american sign language lexicon video dataset (aslvld) corpus. In *5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon, LREC*. Citeseer. <https://hdl.handle.net/2144/31899>
- [131] Carol Neidle and Christian Vogler. 2012. A new web interface to facilitate access to corpora: Development of the ASLLRP data access interface (DAI). In *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon, 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon (LREC '12)*. Citeseer. <https://hdl.handle.net/2144/31886>
- [132] Spiros Nikolopoulos, Kostas Georgiadis, Fotis Kalaganis, Georgios Liaros, Ioulietta Lazarou, Katerina Adam, Anastasios Papazoglou-Chalikias, Elisavet Chatziari, Vangelis P Oikonomou, Panagiotis Petrantonakis, Ioannis Kompatsiaris, Chandan Kumar, Raphael Menges, Steffen Staab, Daniel Muller, Korok Sengupta, Sevasti Bostantjopoulou, Zoe Katsarou, Gabi Zeilig, Meir Plotnin, Amihai Gotlieb, Sofia Fountoukidou, Jaap Ham, Dimitrios Athanasiou, Agnes Mariakaki, Dario Comanducci, Eduardo Sabatini, Walter Nistico, and Markus Plank. 2017. The Mamec Project-A Dataset For Multimodal Human-Computer Interaction Using Biosignals And Eye Tracking Information. <https://doi.org/10.5281/ZENODO.834153>
- [133] Institute of Medicine. 1992. *The Second Fifty Years: Promoting Health and Preventing Disability*. The National Academies Press. <https://doi.org/10.17226/1578>
- [134] Research & Training Center on Independent Living. 2020. Guidelines: How to write and report about people with disabilities, 9th edition. <https://rtcil.org/guidelines-9th-edition>
- [135] JR Orozco-Arroyave, F Hönig, JD Arias-Londoño, JF Vargas-Bonilla, K Daqrouq, S Skodda, J Ruzs, and E Nöth. 2016. Automatic detection of Parkinson's disease in running speech spoken in three different languages. *The Journal of the Acoustical Society of America* 139, 1 (2016), 481–500. <https://doi.org/10.1121/1.4939739>
- [136] M. Oszust and M. Wysocki. 2013. Polish sign language words recognition with Kinect. In *Proceedings of 6th International Conference on Human System Interactions (HSI '13)*. 219–226. <https://doi.org/10.1109/HSI.2013.6577826>
- [137] Joon Sung Park, Michael S. Bernstein, Robin N. Brewer, Ece Kamar, and Meredith Ringel Morris. 2021. Understanding the Representation and Representativeness of Age in AI Data Sets. *CoRR* abs/2103.09058 (2021). <https://arxiv.org/abs/2103.09058>
- [138] Daniel Keyzers Philippe Dreu, Thomas Deselaers and Hermann Ney. 2006. RWTH German Fingerspelling Database. <https://www-i6.informatik.rwth-aachen.de/aslr/fingerspelling.php>
- [139] Kelly Quinn, Christina E Miyawaki, Raina Croff, Mia T Vogel, Basia Belza, Anita M Souza, Minhui Liu, Valerie J Edwards, and Daniela B Friedman. 2020. Terms and Measures of Cognitive Health Associated with Dementia and Alzheimer's Disease: A Scoping Review. *Research on aging* 42, 5-6 (2020), 174–185.
- [140] James M Rehg, Agata Rozga, Gregory D Abowd, and Matthew S Goodwin. 2014. Behavioral imaging and autism. *IEEE Pervasive Computing* 13, 2 (2014), 84–87. <http://dx.doi.org/10.1109/MPRV.2014.23>
- [141] Luz Rello, Ricardo Baeza-Yates, Abdullah Ali, Jeffrey P Bigham, and Miquel Serra. 2020. Predicting risk of dyslexia with an online gamified test. *Plos one* 15, 12 (2020), e0241687. <https://doi.org/10.1371/journal.pone.0241687>
- [142] Luz Rello, Ricardo Baeza-Yates, and Joaquim Llisterri. 2014. DysList: An Annotated Resource of Dyslexic Errors. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC '14)*. European Languages Resources Association (ELRA), 1289–1296. [http://www.lrec-conf.org/proceedings/lrec2014/pdf/612\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2014/pdf/612_Paper.pdf)
- [143] Luz Rello and Miguel Ballesteros. 2015. Detecting Readers with Dyslexia Using Machine Learning with Eye Tracking Measures. In *Proceedings of the 12th International Web for All Conference (W4A '15)*. Association for Computing Machinery (ACM), Article 16, 8 pages. <https://doi.org/10.1145/2745555.2746644>
- [144] Brandon Richardson. 2017. find+. <https://addons.mozilla.org/en-US/firefox/addon/brandon1024-find/>
- [145] Frank Rudzicz, Aravind Kumar Namasivayam, and Talya Wolff. 2012. The TORGO database of acoustic and articulatory speech from speakers with dysarthria. *Language Resources and Evaluation* 46, 4 (2012), 523–541. <https://doi.org/10.1007/s10579-011-9145-0>
- [146] Jan Ruzs, Roman Cmejla, Tereza Tykalova, Hana Ruzickova, Jiri Klempir, Veronika Majerova, Jana Picmausova, Jan Roth, and Evzen Ruzicka. 2013. Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. *The Journal of the Acoustical Society of America* 134, 3 (2013), 2171–2181.
- [147] Betül Erdogdu Sakar, M Erdem Isenkul, C Okan Sakar, Ahmet Sertbas, Fikret Gurgun, Sakir Delil, Hulya Apaydin, and Olcay Kursun. 2013. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics* 17, 4 (2013), 828–834. <https://doi.org/10.1109/JBHI.2013.2245674>
- [148] Anand D. Sarwate and Kamalika Chaudhuri. 2013. Signal Processing and Machine Learning with Differential Privacy: Algorithms and Challenges for Continuous Data. *IEEE Signal Processing Magazine* 30, 5 (2013), 86–94. <https://doi.org/10.1109/MSP.2013.2259911>
- [149] Björn Schuller, Stefan Steidl, Anton Batliner, Simone Hantke, Florian Hönig, Juan Rafael Orozco-Arroyave, Elmar Nöth, Yue Zhang, and Felix Weninger. 2015. The Interspeech 2015 computational paralinguistics challenge: nativeness, parkinson's & eating condition. In *Proceedings of the 16th annual conference of the international speech communication association*. 478–482. <https://doi.org/10.1.1.728.2334>
- [150] Andrew Sears and Vicki L Hanson. 2012. Representing Users in Accessibility Research. *ACM Trans. Access. Comput* 4, 2, Article 7 (March 2012), 6 pages. <https://doi.org/10.1145/2141943.2141945>
- [151] Rajani Sebastian, Carol B Thompson, Nae-Yuh Wang, Amy Wright, Aaron Meyer, Rhonda B Friedman, Argye E Hillis, and Donna C Tippet. 2018. Patterns of decline in naming and semantic knowledge in primary progressive aphasia. *Aphasiology* 32, 9 (2018), 1010–1030. <https://doi.org/10.1080/02687038.2018.1490388>
- [152] Sally E Shaywitz. 1998. Dyslexia. *New England Journal of Medicine* 338, 5 (1998), 307–312. <https://www.nejm.org/doi/full/10.1056/nejm199801293380507>
- [153] Samaa M. Shohieb, Hamdy K. Elminir, and A.M. Riad. 2015. SignsWorld Atlas; a benchmark Arabic Sign Language database. *Journal of King Saud University - Computer and Information Sciences* 27, 1 (2015), 68–76. <https://doi.org/10.1016/j.jksuci.2014.03.011>
- [154] Joan Sosa-García and Francesca Odone. 2017. "Hands On" Visual Recognition for Visually Impaired Users. *ACM Transactions on Accessible Computing* 10, 3, Article 8 (Aug. 2017), 30 pages. <https://doi.org/10.1145/3060056>
- [155] Katta Spiel, Christopher Frauenberger, Os Keyes, and Geraldine Fitzpatrick. 2019. Agency of Autistic Children in Technology Research—A Critical Literature Review. *ACM Transactions of Computer Human Interactions* 26, 6, Article 38 (Nov. 2019), 40 pages. <https://doi.org/10.1145/3344919>
- [156] Darren B Taichman, Peush Sahni, Anja Pinborg, Larry Peiperl, Christine Laine, Astrid James, Sung-Tae Hong, Abraham Haileamlak, Laragh Gollgoly, Fiona Godlee, Frank A Frizelle, Fernando Florenzano, Jeffrey M Drazen, Howard Bauchner, Christopher Baethge, and Joyce Backus. 2017. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. *The Lancet* 389, 10086 (June 2017), e12–e14. [https://doi.org/10.1016/S0140-6736\(17\)31282-5](https://doi.org/10.1016/S0140-6736(17)31282-5)
- [157] I. Thomas, J. Westin, M. Alam, F. Bergquist, D. Nyholm, M. Senek, and M. Memedi. 2018. A Treatment-Response Index From Wearable Sensors for Quantifying Parkinson's Disease Motor States. *IEEE Journal of Biomedical and Health Informatics* 22, 5 (2018), 1341–1349. <https://doi.org/10.1109/JBHI.2017.2777926>
- [158] Christian Thomsen and Torben Bach Pedersen. 2006. Building a Web Warehouse for Accessibility Data. In *Proceedings of the 9th ACM International Workshop on Data Warehousing and OLAP (DOLAP '06)*. Association for Computing Machinery (ACM), 43–50. <https://doi.org/10.1145/1183512.1183522>
- [159] Pamela Toman. 2016. Complexity Beyond the Trigram: Identifying Sign Languages from Video Using Neural Networks. (2016). [http://cs231n.stanford.edu/reports/2016/pdfs/207\\_Report.pdf](http://cs231n.stanford.edu/reports/2016/pdfs/207_Report.pdf)
- [160] Jutta Treviranus. 2019. The Value of Being Different. In *Proceedings of the 16th Web For All 2019 Personalization - Personalizing the Web (W4A '19)*. Association for Computing Machinery (ACM), Article 1, 7 pages. <https://doi.org/10.1145/3315002.3332429>

- [161] Shari Trewin, Sara Basson, Michael Muller, Stacy Branham, Jutta Treviranus, Daniel Gruen, Daniel Hebert, Natalia Lyckowski, and Erich Manser. 2019. Considerations for AI Fairness for People with Disabilities. *AI Matters* 5, 3 (Dec. 2019), 40–63. <https://doi.org/10.1145/3362077.3362086>
- [162] Civil Rights Division US Department of Justice. 2021. Information and Technical Assistance on the Americans with Disabilities Act. <https://www.ada.gov/>
- [163] Hamid Vaezi Joze and Oscar Koller. 2019. MS-ASL: A Large-Scale Data Set and Benchmark for Understanding American Sign Language. In *Proceedings of the British Machine Vision Conference (BMVC '19)*. <https://www.microsoft.com/en-us/research/publication/ms-asl-a-large-scale-data-set-and-benchmark-for-understanding-american-sign-language/>
- [164] Juan Camilo Vásquez-Correa, Tomas Arias-Vergara, Juan Rafael Orozco-Arroyave, Björn Eskofier, Jochen Klucken, and Elmar Nöth. 2018. Multimodal assessment of Parkinson's disease: a deep learning approach. *IEEE journal of biomedical and health informatics* 23, 4 (2018), 1618–1630. <https://doi.org/10.1109/jbhi.2018.2866873>
- [165] Radu-Daniel Vatavu, Bogdan-Florin Gheran, and Maria Doina Schipor. 2018. The impact of low vision on touch-gesture articulation on mobile devices. *IEEE Pervasive Computing* 17, 1 (2018), 27–37. <https://doi.org/10.1109/MPRV.2018.011591059>
- [166] Radu-Daniel Vatavu and Ovidiu-Ciprian Ungurean. 2019. Stroke-Gesture Input for People with Motor Impairments: Empirical Results & Research Roadmap. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery (ACM), 1–14. <https://doi.org/10.1145/3290605.3300445>
- [167] C. M. Vikram, Nagaraj Adiga, and S. R. Mahadeva Prasanna. 2019. Detection of Nasalized Voiced Stops in Cleft Palate Speech Using Epoch-Synchronous Features. *IEEE/ACM Trans. Audio, Speech and Lang. Proc* 27, 7 (July 2019), 1189–1200. <https://doi.org/10.1109/TASLP.2019.2913089>
- [168] Chat Wacharamanatham, Lukas Eisenring, Steve Haroz, and Florian Echtler. 2020. Transparency of CHI Research Artifacts: Results of a Self-Reported Survey. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery (ACM), 1–14. <https://doi.org/10.1145/3313831.3376448>
- [169] Mark Walport and Paul Brest. 2011. Sharing research data to improve public health. *The Lancet* 377, 9765 (2019/09/15 2011), 537–539. [https://doi.org/10.1016/S0140-6736\(10\)62234-9](https://doi.org/10.1016/S0140-6736(10)62234-9)
- [170] Dawid Warchol, Tomasz Kapuściński, and Marian Wysocki. 2019. Recognition of fingerspelling sequences in polish sign language using point clouds obtained from depth images. *Sensors* 19, 5 (2019), 1078. <https://doi.org/10.3390/s19051078>
- [171] Danielle Wetherell, Nicola Botting, and Gina Conti-Ramsden. 2007. Narrative skills in adolescents with a history of SLI in relation to non-verbal IQ scores. *Child Language Teaching and Therapy* 23, 1 (2007), 95–113. <https://doi.org/10.1177/0265659007072322>
- [172] Ryen W White, P Murali Doraiswamy, and Eric Horvitz. 2018. Detecting neurodegenerative disorders from web search signals. *NPJ digital medicine* 1, 1 (2018), 1–4. <https://doi.org/10.1038/s41746-018-0016-6>
- [173] Ryen W White and Eric Horvitz. 2019. Population-scale hand tremor analysis via anonymized mouse cursor signals. *NPJ digital medicine* 2, 1 (2019), 1–7. <https://doi.org/10.1038/s41746-019-0171-4>
- [174] Meredith Whittaker, Meryl Alper, Cynthia L Bennett, Sara Hendren, Liz Kazunas, Mara Mills, Meredith Ringel Morris, Joy Rankin, Emily Rogers, Marcel Salas, et al. 2019. Disability, Bias, and AI. *AI Now Institute, November* (2019). <https://wecount.inclusivedesign.ca/uploads/Disability-bias-AI.pdf>
- [175] World Health Organization (WHO). 2020. Disability and health. <https://www.who.int/news-room/fact-sheets/detail/disability-and-health>
- [176] Vicki Xafis, G Owen Schaefer, Markus K Labude, Iain Brassington, Angela Ballantyne, Hannah Yeefen Lim, Wendy Lipworth, Tamra Lysaght, Cameron Stewart, Shirley Sun, et al. 2019. An ethics framework for big data in health and research. *Asian Bioethics Review* 11, 3 (2019), 227–254. <https://doi.org/10.1007/s41649-019-00099-x>
- [177] Dongxin Xu, Jeffrey A. Richards, Jill Gilkerson, Umit Yapanel, Sharmistha Gray, and John Hansen. 2009. Automatic Childhood Autism Detection by Vocalization Decomposition with Phone-like Units. In *Proceedings of the 2nd Workshop on Child, Computer and Interaction (WOCCI '09)*. Association for Computing Machinery (ACM), Article 5, 7 pages. <https://doi.org/10.1145/1640377.1640382>
- [178] Victoria Yaneva, Irina Temnikova, and Ruslan Mitkov. 2015. Accessible Texts for Autism: An Eye-Tracking Study. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS '15)*. Association for Computing Machinery (ACM), 49–57. <https://doi.org/10.1145/2700648.2809852>
- [179] Victoria Yaneva, Irina Temnikova, and Ruslan Mitkov. 2016. A Corpus of Text Data and Gaze Fixations from Autistic and Non-Autistic Adults. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC '16)*. European Language Resources Association (ELRA). <https://aclanthology.org/L16-1077>
- [180] Brit Youngmann, Liron Allerhand, Ora Paltiel, Elad Yom-Tov, and David Arkadir. 2019. A machine learning algorithm successfully screens for Parkinson's in web users. *Annals of clinical and translational neurology* 6, 12 (2019), 2503–2509. <https://doi.org/10.1002/acn3.50945>
- [181] Hanbin Zhang, Chen Song, Aosen Wang, Chenhan Xu, Dongmei Li, and Wenyao Xu. 2019. PDVocal: Towards Privacy-Preserving Parkinson's Disease Detection Using Non-Speech Body Sounds. In *Proceedings of the 25th Annual International Conference on Mobile Computing and Networking (MobiCom '19)*. Association for Computing Machinery, Article 16, 16 pages. <https://doi.org/10.1145/3300061.3300125>