Accessibility-Related Publication Distribution in HCI Based on a Meta-Analysis

Mark Colley  
mark.colley@uni-ulm.de  
Institute of Media Informatics, Ulm University  
Ulm, Germany

Taras Kränzle  
taras.kraenzle@alumni.uni-ulm.de  
Institute of Media Informatics, Ulm University  
Ulm, Germany

Enrico Rukzio  
enrico.rukzio@uni-ulm.de  
Institute of Media Informatics, Ulm University  
Ulm, Germany

ABSTRACT

Accessibility research aims to aid humans that experience minor or major disabilities and conditions. However, researchers might have limited exposure to certain disabilities, therefore, focus on those prevalent in their own lives. This work presents a script-based meta-analysis on addressed populations in accessibility research published on top Human-Computer Interaction (HCI) venues (3617 full papers). We categorize the publications regarding the involved people and their disabilities. We found that work on vision disability makes up for almost one third (27.85%) of the work published in general HCI. In light of these findings, we present possible conference- and funding-related explanatory approaches and argue that disability research could more reflect the prevalence of disabilities in the world.

CCS CONCEPTS

• General and reference  →  Surveys and overviews.

KEYWORDS

Accessibility; overview; survey.

ACM Reference Format:

1 INTRODUCTION

According to The International Classification of Functioning, Disability and Health (ICF) [316], there are multiple categories of impairments that can be broadly sub-categorized into four types [129]: Mobility and Physical Impairments, Vision, Hearing, and Cognitive or Learning Disabilities. For example, there are around 1.3 billion people with some, 217 million people with a severe vision impairment, and 36 million people are blind [42, 247]; 466 million people have disabling hearing loss [248], and 50 million people have some form of dementia [249]. The distribution of disability in the USA in 2015, as an example, is 3.6% for hearing, 2.3% for vision, and 4.8% for cognitive impairments [176]. Accessibility plays an important role in various conferences today both for attendance as well as a research topic [1]. The International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS) [90] specifically addresses research on people with disabilities. Additionally, there are special subcommittees at CHI for mental health, learning and accessibility & ageing [245]. However, a quantitative approach to how strongly each disability is addressed in the field of Human-Computer Interaction (HCI) is missing. HCI refers to a "multidisciplinary field of study focusing on the design of computer technology and, in particular, the interaction between humans (the users) and computers" [98]. This is important to shape future research. Complex problems require formulating the challenges first [134]. Knowing the distribution of papers on disabilities can facilitate discussions about potential underrepresentation, especially important research topics, and distribution of research funds. Thus, the first important step is to report relevant data as, for example, done by Microsoft [185].

This work first provides an overview of the addressed population from 2009 — 2019 that were published at ASSETS to define relevant fields and terminology. Based on this overview, a script-based categorization approach was defined. For this, we first categorized the papers of ASSETS manually and then compared this to the script-based paper categorization approach. Such a script-based approach enabled us to categorize many more publications than would be possible manually. After an iterative improvement process, we used this script to search further venues in the accessibility community (W4A [6], TACCESS [243]) and the broader HCI domain (CHI [244], UIST [293], AutoUI [149], and MUM [216]) in the time frame of 2016 – 2019 were included in the quantitative categorization process. We found that conferences vary greatly in the addressed populations such as visual disabilities for ASSETS, TACCESS, and W4A or cognitive and learning disabilities at CHI. Visual impairments were studied and addressed the most (27.85%).

Contribution statement: This work contributes to the body of accessibility-related knowledge by providing a meta-analysis on addressed disabilities in the field of HCI. For this, an R script-based approach was defined, compared to a manual classification of the ASSETS papers from 2009 – 2019 and then applied to other relevant venues. This work shows different focus points of the conferences and reveals the misalignment of the work on disabilities with the prevalence of disabilities in the general population.

2 METHOD

In this paper, we provide a meta-analysis on the addressed disabilities at major HCI venues. First, we describe our manual and script-based process. We reviewed the main proceedings of ASSETS
from 2009 — 2019 (316 publications). We collected all 316 ASSETS papers and formalized our process to classify and review the papers. We chose ASSETS as a premier venue for accessibility-related publications and the time frame to gather a large enough sample.

2.1 Manual Categorization
All ASSETS papers were categorized manually by two authors. For this, the two authors first read the title, the abstract, and the keywords. If a categorization was not possible based on these sections, the entire full paper was read. Both authors categorized all papers. Disagreements were resolved via discussions. One author already has knowledge about and worked with people with vision impairments. Additionally, a medical physician was available for queries regarding nomenclature and disabilities. First, we categorized publications into four types [129]: Mobility and Physical Impairments, Vision, Hearing, and Cognitive or Learning Disabilities. We found that some papers could not be categorized as more than one disability or older adults were addressed. Therefore, we added the categories General Accessibility and Elderly. We explicitly do not see being of older age to be a disability. Additionally, we divided psychological disorders from cognitive disabilities as these seem to become more prevalent [246]. This is also in line with the distinction of the UN Convention on the Rights of Persons with Disabilities (UNCRPD) between mental and intellectual disabilities [121]. As we will discuss in section 5, there is an inevitable conceptual overlap between these disabilities [208], however, not in our sub-category terminology. The results can be seen in Table 1.

2.2 Script-based Categorization
After the manual categorization, the publications were analyzed via a R script using the pdfsearch [187] (see Appendix C; available under https://github.com/M-Colley/accessibility-distribution-meta). We defined relevant terms of the addressed disability (e.g., “visual impairment”). Additionally, we asked a medical physician for relevant keywords. As authors tend to use different vocabulary, we used the first ten synonyms with a rating greater than 200 (highest possible, crowd-sourced) from https://www.powerthesaurus.org as an “excellent resource of English words and their synonyms” [281, p. 4] for the search.

With preliminary criteria, we first analyzed the 316 publications of ASSETS. In an iterative process, we added relevant keywords. When no match was found, the paper was again analyzed independently by two authors and relevant search terms were added to the script. Therefore, the search term lists vary in length. Disagreements were resolved through discussions. In the case of unclear affiliations to one of the categories, the authors also asked a medical physician as these are examined in detecting and working with people with disabilities.

For ASSETS, the classification based on the R script for ASSETS was then compared to our manual categorization. Our script categorized 306/316 (≈ 96.84%) ASSETS paper. We found that 59/77 table cells (one cell per year and addressed disability; e.g., Vision Disability for 2009 had 9 references in Table 1) were populated the same. As a distance function measuring the deviation from the manual categorization, we defined the absolute difference for each cell. This showed that there is a difference of 22 (max 2 per cell, see Appendix B Table 3) publications (e.g., the automatically generated table has 7 references for Vision Disability for 2009, thus the difference is 2). Ten of these can be attributed to the non-classified publications and five can be attributed to non-searchable publications. Nevertheless, we were able to correctly classify 93% of all publications at ASSETS. This leads us to conclude that our approach, while not perfect, provides a reliable picture of accessibility-related publications at HCI venues. The relevant search terms can be found in Appendix A.

Afterward, we repeated this with the additional publications. For this, we followed the PRISMA [217] process. PRISMA refers to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses. The papers had to be published within the time frame of 2016 – 2019. We chose this different time frame because (1) we found no changes over the distribution at ASSETS over the 11 years and (2) CHI allows bulk downloads for this time frame (as of January 2021). They also had to be published in a journal or on a conference with an HCI focus: W4A, TACCESS, CHI, UIST, AutoUI, and MUM. We downloaded all papers from the ACM DL.

First, we repeated the procedure with the papers of W4A and TACCESS, as it is clear that all publications address accessibility. Again, relevant search terms from titles were added manually. After another iteration, with our script, of the 494 publication in ASSETS, W4A, and TACCESS, we could not categorize 15 or 3.23% of the publications. Of these, 5 were deemed maybe-accessibility-related and 10 accessibility-unrelated (see criteria below).
For these 494 publications, on average, the category with the most occurrences had $M=88.41$ ($SD=69.07$) occurrences (e.g., terms from one category occurred $\approx 88$ times). The first quartile was 41. The category with the second most occurrences had $M=14.02$ ($SD=17.00$). The first quartile was 4. Overall, publications had $M=112.99$ ($SD=79.53$) accessibility-related keywords, the first quartile being 59. Based on these numbers, we considered a publication as assigned to a category when a search term was found in the title or the keywords or otherwise if (1) more than 23 references (0.1 quantile of most common keyword category) to one category were found and (2) either the ratio to the second most common keywords was above 2.0 or no more than 14 references (mean of second most common keywords) to another category were found in the entire full paper. 

References refers to the number of appearances of one specific term (e.g., dyslexia) in the entire full paper. As all works include some reference to accessibility, we defined that if at least 20 references (0.05-quantile of the overall keyword occurrences for the accessibility-related venues) to search terms are found, this paper was categorized as maybe accessibility-related. With this threshold, only 10/494 $\approx 2.02\%$ publications of the accessibility-related venues ASSETS, TACCESS, and W4A were wrongly categorized as accessibility-unrelated. 271 were categorized by title, 46 by keywords. This approach also guards against “accessibility-as-icing”, as singular sentences indicating a potential relevance to people with disabilities will not be categorized as relevant.

As we aimed to gain a broader picture of the HCI community, we included the full papers of the conferences CHI, UIST, AutoUI, and MUM in the R-based analysis. We chose CHI and UIST as large top tier HCI venues and AutoUI and MUM as surrogates for smaller, more focused conferences. With the updated criteria, we checked all full papers. Taken together, we categorized 3617 full papers. For every conference, a data table was automatically generated (see supplementary material). Following the PRISMA method [217], we only included the papers outlined above, excluded no records, and performed no qualitative synthesis.

Usage of `pdfsearch` [187] in full text: keyword_search(pdf_file, keyword=keywords, path=TRUE, remove_hyphen=TRUE, ignore_case=TRUE, surround_lines=1, split_pdf=TRUE) where: keywords are the search terms per disability, remove_hyphen is set to TRUE to combine hyphenated words, ignore_case is TRUE to disregard capitalization, and split_pdf is TRUE as this is “most useful with multicolurn pdf files” [187, p. 6].

3 FINDINGS

We report descriptive statistics for all papers and conference-related findings. We found that 1062 of the 3617 (29.36%) full papers to address a disability. Additionally, we found that 400 full papers were potentially accessibility-related. 577 of 1062 (54.33%) were classified by either the title (445 or 41.90%) or the keywords (132 or 12.43%).

ASSETS: We predominantly found vision-related papers (128/316 or 40.51%; see Figure 2a). While we approve of this effort and do not want this to cease, we are worried that other disabilities are not included sufficiently. Hearing and cognitive disabilities were addressed the second most. All other disabilities each represent $\approx 10\%$ or less of the papers.

AutoUI: AutoUI did not address disabilities to a great extent. With our script, we found that one publication addressed cognitive disability and two addressed older adults (2.2% of the 136 analyzed papers).

CHI: 515/2514 (20.49%) papers addressed disability. These works mainly focused on older adults (128, 5.09%), psychological (94, 3.74%), or cognitive (86, 3.42%) disabilities (see Figure 2b). This accounts for the more frequent occurrence of mental health problems [246]. Visual disabilities were the fourth most common disability addressed (84). Other disabilities were only addressed to a minor degree; hearing disability, as the fifth most common, was only addressed in 4.47% of all the accessibility-related papers.

MUM: We found that two papers addressed vision, four cognitive, one psychological, five mobility disability and five addressed older adults (in total, 11.49% of the 148 analyzed papers).

TACCESS: We categorized 70/70 papers. Vision-related papers were most dominant (31/70, 44.29%). Older adults, the second most common addressed population, were only addressed in 10/70 (14.29%) of the categorized papers.

UIST: 49/325 (15.08%) addressed disability. Mobility (21, 6.46%) and visual (15, 4.62%) were most dominantly addressed.

W4A: 102/108 (94.4%) papers were categorized. W4A, with addressing website accessibility, focuses mainly on general accessibility (43, 39.81%; this includes accessibility evaluations) and visual impairments (33, 30.56%).

In Table 2, the occurrences per disability addressed are summed up. Most work focuses on vision impairments while the other disabilities are addressed to an equal degree. Hearing-impairment related work is scarcer. When the distribution of disabilities is taken into account, we argue that this distribution could look different and, therefore, while these findings are not worrisome, action should be taken to better address underrepresented disabilities.

4 EXPLANATORY APPROACHES, IMPLICATIONS, AND RECOMMENDATIONS

Almost a third of the publications were focused on vision impairment and another third on the combination of cognitive and psychological disabilities. In this section, we propose explanations for the analysis’ findings.

4.1 Possible Conference-Related Explanations

Driving is a highly visual and often demanding task which is difficult for people with impairments [250]. This could be one reason for the lack of accessibility-related work on AutoUI. However, there are approaches, e.g., to aid people with vision impairments to steer [300]. We also believe that with the advance of autonomous driving, the target group of vehicle-related research will change towards work-related and accessibility-related research (e.g., see [77]).

ASSETS, TACCESS, and UIST seem to focus highly on visual impairments. One interpretation of the dominance of visual-impairment related papers might be that a lot of visual feedback is apparent in today’s society, e.g., billboards, television, and the worldwide web. We predominantly perceive information through a display in 2020, for example via smartphone, tablet, laptop, PC, or screen walls. Therefore, researchers could attribute a high relevance to these topics. With an increased amount of visual
Table 1: Distribution of publications based on the addressed population at ASSETS from 2009 — 2019.

<table>
<thead>
<tr>
<th>Year</th>
<th>Vision Disability</th>
<th>Hearing Disability</th>
<th>Brain, Cognitive or Learning Disabilities</th>
<th>Psychological Disorders</th>
<th>Mobility and Physical Impairments</th>
<th>General Accessibility</th>
<th>Elderly</th>
<th>Overall Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>9</td>
<td>5</td>
<td>13</td>
<td>17</td>
<td>97</td>
<td>119</td>
<td>164</td>
<td>278</td>
</tr>
</tbody>
</table>

Table 2: Distribution of overall papers based on addressed population.

<table>
<thead>
<tr>
<th>Vision Disability</th>
<th>Hearing Disability</th>
<th>Brain, Cognitive or Learning Disabilities</th>
<th>Psychological Disorders</th>
<th>Mobility and Physical Impairments</th>
<th>General Accessibility</th>
<th>Elderly</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>291 (27.83%)</td>
<td>81 (7.75%)</td>
<td>145 (13.88%)</td>
<td>131 (12.54%)</td>
<td>121 (11.58%)</td>
<td>103 (9.86%)</td>
<td>173 (16.56%)</td>
<td>1045</td>
</tr>
</tbody>
</table>

With vision impairments. Therefore, the focus on this population is explainable.

Finally, there are also difficulties regarding recruiting participants of certain disabilities and running usability testing, focus groups and interviews with younger/child populations.

4.2 Bias in Funding Agencies

Funding of research projects is a potential source for the disproportionate distribution of papers. Currently, there are 25 EU-funded research projects on technologies for accessibility [78]. Three address visual impairments, two hearing impairments, one deaf-blindness, five motor impairments, nine cognitive impairments, and five general accessible communication technology. In Canada, currently announced projects are made of seven projects addressing mobility impairment, eight projects addressing general disabilities, three visual, and two cognitive disabilities [238]. Microsoft's AI for Accessibility program [212] has no special focus an a disability. In their featured examples, all categories of disabilities are addressed. While this is not a full view of funding agencies around the world, it seems that the bias towards visual and cognitive and psychological disabilities is, at least today, not induced by funding agencies.

4.3 Publication

Based on our script-based approach and the difficulties we experienced, we propose some enhancements for the publication process such as including the target group and the used methodology in the
Keywords. This could also be done automatically, as done by other venues such as IEEE Xplore Digital Library. Additionally, we recommend a more rigorous approach to the enforcement of metadata validity. This is relevant for several tasks such as (our) script-based approaches but also, for example, bibliometrics. We also stress the need to provide papers in an accessible and searchable format. This is also necessary to ensure that screen readers can function properly.

5 THOUGHTS ON AND LIMITATIONS OF OUR APPROACH & FUTURE WORK

Categorization is shaped by culture and worldview and, itself shapes these [43]. It can also "suggest bias or reflect negative, disparaging, or patronizing attitudes toward individuals or groups of individuals" [129, p. 62]. With our suggested approach, potentially harmful considered terms (e.g., elderly, see [129]) were used for classification. Additionally, including a medical physician could have included a biased view on the topic of disability. Hanson et al. [129] state that language alters and we strongly support this change towards unharmful terminology, however, for our purpose, these search terms were necessary. Also, some overlap between the categorization is inevitable. We assured that most search terms can only be put into one category, however, for example, vision disability is often addressed for older adults, resulting in an overlap. The creation of the search terms, in general, was difficult and demanded an extensive use of reading at least titles and keywords of the paper. Still, it cannot be eliminated that some keywords were overlooked.
Our R-based categorization of work potentially could have misclassified some works due to unaccounted or overlapping search terms, not accounting for abbreviations etc. Additionally, the difference in length for search terms per disability could have biased the results. While the categorization via title or keywords, performed well, it had difficulties when general terms were used in the keywords alongside more specific ones. For example, keywords that first have assistive technology and then autism would be categorized under General Accessibility while, if the keywords would be reversed it would be categorized under Psychological Disorder. This became clear when the script-based approach was compared to the manual categorization for ASSETS papers.

Classification solely based on the title, abstract, and keywords was difficult. While our script searched the entire paper, adding such relevant information in these parts would significantly ease the reading process. Also, we found that numerous files have incorrect metadata. This limited script-based process. For example, the TACCESS papers’ title is TACCESS-Volumese-Number of Article, e.g. TACCESS1201-03. Keywords are also missing in the TACCESS publications’ metadata. Additionally, numerous PDF files were not text-searchable. For ASSETS, these were five publications. While this resembles only 1.6%, this could still have biased our results.

6 CONCLUSION

In conclusion, we have shown that people with disabilities are addressed in numerous conferences in HCI. Some of the venues seem to address specific disabilities disproportionately. ASSETS and TACCESS, for example, seem to mainly address visual impairments. To uncover this finding, we employed a script-based methodology. We proposed some explanations such that visual disabilities are more graspable for researchers and practitioners. Our work is, to the best of our knowledge, the first approach to quantitatively reveal the addressed populations with disabilities in HCI.

ACKNOWLEDGMENTS

This work was conducted within the project ’SituWare’ funded by the Federal Ministry for Economic Affairs and Energy (BMWi). Mark Colley was supported by the Startup Funding B of Ulm University.

REFERENCES


Our R-based categorization of work potentially could have misclassified some works due to unaccounted or overlapping search terms, not accounting for abbreviations etc. Additionally, the difference in length for search terms per disability could have biased the results. While the categorization via title or keywords, performed well, it had difficulties when general terms were used in the keywords alongside more specific ones. For example, keywords that first have assistive technology and then autism would be categorized under General Accessibility while, if the keywords would be reversed it would be categorized under Psychological Disorder. This became clear when the script-based approach was compared to the manual categorization for ASSETS papers.

Classification solely based on the title, abstract, and keywords was difficult. While our script searched the entire paper, adding such relevant information in these parts would significantly ease the reading process. Also, we found that numerous files have incorrect metadata. This limited script-based process. For example, the TACCESS papers’ title is TACCESS-Volumese-Number of Article, e.g. TACCESS1201-03. Keywords are also missing in the TACCESS publications’ metadata. Additionally, numerous PDF files were not text-searchable. For ASSETS, these were five publications. While this resembles only 1.6%, this could still have biased our results.

6 CONCLUSION

In conclusion, we have shown that people with disabilities are addressed in numerous conferences in HCI. Some of the venues seem to address specific disabilities disproportionately. ASSETS and TACCESS, for example, seem to mainly address visual impairments. To uncover this finding, we employed a script-based methodology. We proposed some explanations such that visual disabilities are more graspable for researchers and practitioners. Our work is, to the best of our knowledge, the first approach to quantitatively reveal the addressed populations with disabilities in HCI.

ACKNOWLEDGMENTS

This work was conducted within the project ’SituWare’ funded by the Federal Ministry for Economic Affairs and Energy (BMWi). Mark Colley was supported by the Startup Funding B of Ulm University.

REFERENCES


Our R-based categorization of work potentially could have misclassified some works due to unaccounted or overlapping search terms, not accounting for abbreviations etc. Additionally, the difference in length for search terms per disability could have biased the results. While the categorization via title or keywords, performed well, it had difficulties when general terms were used in the keywords alongside more specific ones. For example, keywords that first have assistive technology and then autism would be categorized under General Accessibility while, if the keywords would be reversed it would be categorized under Psychological Disorder. This became clear when the script-based approach was compared to the manual categorization for ASSETS papers.

Classification solely based on the title, abstract, and keywords was difficult. While our script searched the entire paper, adding such relevant information in these parts would significantly ease the reading process. Also, we found that numerous files have incorrect metadata. This limited script-based process. For example, the TACCESS papers’ title is TACCESS-Volumese-Number of Article, e.g. TACCESS1201-03. Keywords are also missing in the TACCESS publications’ metadata. Additionally, numerous PDF files were not text-searchable. For ASSETS, these were five publications. While this resembles only 1.6%, this could still have biased our results.

6 CONCLUSION

In conclusion, we have shown that people with disabilities are addressed in numerous conferences in HCI. Some of the venues seem to address specific disabilities disproportionately. ASSETS and TACCESS, for example, seem to mainly address visual impairments. To uncover this finding, we employed a script-based methodology. We proposed some explanations such that visual disabilities are more graspable for researchers and practitioners. Our work is, to the best of our knowledge, the first approach to quantitatively reveal the addressed populations with disabilities in HCI.

ACKNOWLEDGMENTS

This work was conducted within the project ’SituWare’ funded by the Federal Ministry for Economic Affairs and Energy (BMWi). Mark Colley was supported by the Startup Funding B of Ulm University.

REFERENCES


Our R-based categorization of work potentially could have misclassified some works due to unaccounted or overlapping search terms, not accounting for abbreviations etc. Additionally, the difference in length for search terms per disability could have biased the results. While the categorization via title or keywords, performed well, it had difficulties when general terms were used in the keywords alongside more specific ones. For example, keywords that first have assistive technology and then autism would be categorized under General Accessibility while, if the keywords would be reversed it would be categorized under Psychological Disorder. This became clear when the script-based approach was compared to the manual categorization for ASSETS papers.

Classification solely based on the title, abstract, and keywords was difficult. While our script searched the entire paper, adding such relevant information in these parts would significantly ease the reading process. Also, we found that numerous files have incorrect metadata. This limited script-based process. For example, the TACCESS papers’ title is TACCESS-Volumese-Number of Article, e.g. TACCESS1201-03. Keywords are also missing in the TACCESS publications’ metadata. Additionally, numerous PDF files were not text-searchable. For ASSETS, these were five publications. While this resembles only 1.6%, this could still have biased our results.

6 CONCLUSION

In conclusion, we have shown that people with disabilities are addressed in numerous conferences in HCI. Some of the venues seem to address specific disabilities disproportionately. ASSETS and TACCESS, for example, seem to mainly address visual impairments. To uncover this finding, we employed a script-based methodology. We proposed some explanations such that visual disabilities are more graspable for researchers and practitioners. Our work is, to the best of our knowledge, the first approach to quantitatively reveal the addressed populations with disabilities in HCI.

ACKNOWLEDGMENTS

This work was conducted within the project ’SituWare’ funded by the Federal Ministry for Economic Affairs and Energy (BMWi). Mark Colley was supported by the Startup Funding B of Ulm University.

REFERENCES


A SEARCH TERMS


Table 3: Distances between manual and automatic categorization for ASSETS.

<table>
<thead>
<tr>
<th>Year</th>
<th>Vision</th>
<th>Hearing</th>
<th>Brain, Cognitive, or Learning</th>
<th>Psychological</th>
<th>Mobility and Physical</th>
<th>General</th>
<th>Elderly</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2011</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2012</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2013</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2014</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2015</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2016</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2017</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2018</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2019</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Combined</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

B METHOD

C R SCRIPT

```r
library(pacman)
pacman::p_load(paste0s, readxl, xlsx, pdfsearch,
    tabulizer, formattable, htmltools, webshot, DT,
    Rfast, ggplot2, broom)

# outsource all definitions for better overview
source("PATH_TO_DEFINITIONS.R", echo = FALSE)

number_of_problems <- 0
number_of_not_categorized <- 0
potentially_related_work <- 0
unrelated_work <- 0
categorized_by_title <- 0
categorized_by_keywords <- 0
df_occurrences <- data.frame(  
  Conference = character(),  
  Vision = integer(),  
  Hearing = integer(),  
  Cognitive = integer(),  
  Psychological = integer(),  
  Mobility = integer(),  
  General = integer(),  
  Elderly = integer(),  
  stringsAsFactors = FALSE)

df_occurrences_un_categorized <- data.frame(  
  Conference = character(),  
  Vision = integer(),  
  Hearing = integer(),  
  Cognitive = integer(),  
  Psychological = integer(),  
  Mobility = integer(),  
  General = integer(),  
  Elderly = integer(),  
  stringsAsFactors = FALSE)
```
df_accessibility_related_keywords_found <- data.frame(Conference = character(),
number = integer(),
stringsAsFactors = FALSE)

# Set directory to get all relevant pdfs, do this for all years
setwd("PATH_FOR_FILES_TOP_FOLDER")

# get all conference directories
collection_directories <- list.files(".", recursive = FALSE)
sink("output_pdf_search_newest.txt")
sink(stdout(), type = "message")

for (d in 1:length(collection_directories)) {
  # go into x'th conference
  setwd(collection_directories[d])
  df_conference <- data.frame(Year = character(),
    "Vision Disability" = integer(),
    "Hearing Disability" = integer(),
    "Cognitive Disability" = integer(),
    "Psychological Disability" = integer(),
    "Mobility Disability" = integer(),
    "General Disability" = integer(),
    "Elderly" = integer(),
    stringsAsFactors = FALSE)

  # get all years
  year_directories <- list.files(".", recursive = FALSE)

  for (y in 1:length(year_directories)) {
    vision_related_publications <- 0
cognitive_related_publications <- 0
psychological_related_publications <- 0
mobility_related_publications <- 0
general_related_publications <- 0
elderly_related_publications <- 0
  
  # go into x'th year

  setwd(year_directories[y])
  year <- as.integer(substring(year_directories[y], 3))
  files <- list.files(pattern = "\pdf")

  for (i in 1:length(files)) {
    # message all the names of the files
    # message(i, ": " , substr(files[i], 1, nchar(files[i]) - 4))

    # pdf_file <- system.file("pdf", files[i], package = "pdfsearch")
    # The pdftools function for extracting text is pdf_text
    pdf_file <- files[i]

    result <- read_Metadata(pdf_file = pdf_file)
title_available <- TRUE

    # Title was extracted
    # Attention: result if gone right is an array of relevant information including author, title
    # if gone wrong, then it is simply ""
    # Metadata can be available but title can still be NULL!
    # Use "" || to avoid error: "$ operator is invalid for atomic vectors"
    if (length(result) == 1 || is.null(result$title)) {
      message("Title of PDF NOT AVAILABLE")
title_available <- FALSE
    } else {
      message("Title of PDF is:", result$title)
    }

    # we search in the title and if any keyword was found, we assume that this is the disability studied
    # we once don't look at this to get information on the distribution!
    if (title_available) {
      # message("Looking in TITLE")
pdf_title <- result$title
  }
result_disability_vision <- keyword_search(pdf_title, keyword = keyword_disability_vision, path = FALSE, ignore_case = TRUE, split_pdf = FALSE)
result_disability_hearing <- keyword_search(pdf_title, keyword = keyword_disability_hearing, path = FALSE, ignore_case = TRUE, split_pdf = FALSE)
result_disability_cognitive <- keyword_search(pdf_title, keyword = keyword_disability_cognitive, path = FALSE, ignore_case = TRUE, split_pdf = FALSE)
result_disability_psychological <- keyword_search(pdf_title, keyword = keyword_disability_psychological, path = FALSE, ignore_case = TRUE, split_pdf = FALSE)
result_disability_mobility <- keyword_search(pdf_title, keyword = keyword_disability_mobility, path = FALSE, ignore_case = TRUE, split_pdf = FALSE)
result_disability_older <- keyword_search(pdf_title, keyword = keyword_disability_older, path = FALSE, ignore_case = TRUE, split_pdf = FALSE)
result_disability_general <- keyword_search(pdf_title, keyword = keyword_disability_general, path = FALSE, ignore_case = TRUE, split_pdf = FALSE)

if (dim(result_disability_vision)[1] > 0) {
  # a term was found in TITLE!
  message("Vision disability term was found in TITLE. Number: ", dim(result_disability_vision)[1])
  result_vision_related_publications <- vision_related_publications + 1
categorized_by_title <- categorized_by_title + 1
  # next to avoid double counting!
  # but only next when there really is one, otherwise we look in entire document
  next
}
if (dim(result_disability_hearing)[1] > 0) {
  # a term was found in TITLE!
  message("Hearing disability term was found in TITLE. Number: ", dim(result_disability_hearing)[1])
  result_hearing_related_publications <- hearing_related_publications + 1
categorized_by_title <- categorized_by_title + 1
  next
}
if (dim(result_disability_cognitive)[1] > 0) {
  # a term was found in TITLE!
  message("Cognitive disability term was found in TITLE. Number: ", dim(result_disability_cognitive)[1])
  result_cognitive_related_publications <- cognitive_related_publications + 1
categorized_by_title <- categorized_by_title + 1
  next
}
if (dim(result_disability_psychological)[1] > 0) {
  # a term was found in TITLE!
  message("Psychological disability term was found in TITLE. Number: ", dim(result_disability_psychological)[1])
  result_psychological_related_publications <- psychological_related_publications + 1
categorized_by_title <- categorized_by_title + 1
  next
}
if (dim(result_disability_mobility)[1] > 0) {
  # a term was found in TITLE!
  message("Mobility disability term was found in TITLE. Number: ", dim(result_disability_mobility)[1])
  result_mobility_related_publications <- mobility_related_publications + 1
categorized_by_title <- categorized_by_title + 1
  next
}
if (dim(result_disability_older)[1] > 0) {
  # a term was found in TITLE!
  message("Elderly disability term was found in TITLE. Number: ", dim(result_disability_older)[1])
  result_elderly_related_publications <- elderly_related_publications + 1
categorized_by_title <- categorized_by_title + 1
  next
}
if (dim(result_disability_general)[1] > 0) {
  # a term was found in TITLE!
  message("General disability term was found in TITLE. Number: ", dim(result_disability_general)[1])
  result_general_related_publications <- general_related_publications + 1
categorized_by_title <- categorized_by_title + 1
  next
}

keywords_available <- TRUE

# keywords_available was extracted
# Attention: result if gone right is an array of relevant information including author, title
# if gone wrong, then it is simply ""
# see https://stat.ethz.ch/R-manual/R-devel/library/base/html/Logic.html
if (length(result) == 1 || is.null(result$keywords)) {
  message("Keywords of PDF NOT AVAILABLE")
  keywords_available <- FALSE
} else {
  message("Keywords are ", result$keywords)
}

# we search in the keywords and if any keyword was found, we assume that this is the disability studied
# we once don’t look at this to get information on the distribution
if (keywords_available) {
  # message("Looking in KEYWORDS")
  pdf_keywords <- result$keywords
  result_disability_vision <- keyword_search(pdf_keywords, keyword = "disability_vision
category = FALSE, ignore_case = TRUE, split_pdf = FALSE)
  result_disability_hearing <- keyword_search(pdf_keywords, keyword = "disability_hearing
category = FALSE, ignore_case = TRUE, split_pdf = FALSE)
  result_disability_cognitive <- keyword_search(pdf_keywords, keyword = "disability_cognitive
category = FALSE, ignore_case = TRUE, split_pdf = FALSE)
  result_disability_psychological <- keyword_search(pdf_keywords, keyword = "disability_psychological
category = FALSE, ignore_case = TRUE, split_pdf = FALSE)
  result_disability_mobility <- keyword_search(pdf_keywords, keyword = "disability_mobility
category = FALSE, ignore_case = TRUE, split_pdf = FALSE)
  result_disability_elder <- keyword_search(pdf_keywords, keyword = "disability_elder
category = FALSE, ignore_case = TRUE, split_pdf = FALSE)
  result_disability_general <- keyword_search(pdf_keywords, keyword = "disability_general
category = FALSE, ignore_case = TRUE, split_pdf = FALSE)

  # anyone with a keyword?
  if (dim(result_disability_vision)[1] > 0) {
    next to avoid double counting!
    # but only next when there really is one, otherwise we look in entire document
    categorized_by_keywords <- categorized_by_keywords + 1
  }
  if (dim(result_disability_hearing)[1] > 0) {
    message("Hearing disability term was found in KEYWORDS Number: ", dim(result_disability_hearing)[1])
    hearing_related_publications <- hearing_related_publications + 1
categorized_by_keywords <- categorized_by_keywords + 1
  }
  if (dim(result_disability_cognitive)[1] > 0) {
    message("Cognitive disability term was found in KEYWORDS Number: ", dim(result_disability_cognitive)[1])
    cognitive_related_publications <- cognitive_related_publications + 1
categorized_by_keywords <- categorized_by_keywords + 1
  }
  if (dim(result_disability_psychological)[1] > 0) {
    message("Psychological disability term was found in KEYWORDS Number: ", dim(result_disability_psychological)[1])
    psychological_related_publications <- psychological_related_publications + 1
categorized_by_keywords <- categorized_by_keywords + 1
  }
  if (dim(result_disability_mobility)[1] > 0) {
    message("Mobility disability term was found in KEYWORDS Number: ", dim(result_disability_mobility)[1])
    mobility_related_publications <- mobility_related_publications + 1
categorized_by_keywords <- categorized_by_keywords + 1
  }
  if (dim(result_disability_elder)[1] > 0) {
    message("Elderly disability term was found in KEYWORDS Number: ", dim(result_disability_elder)[1])
    elderly_related_publications <- elderly_related_publications + 1
categorized_by_keywords <- categorized_by_keywords + 1
  }
  if (dim(result_disability_general)[1] > 0) {
    message("General disability term was found in KEYWORDS Number: ", dim(result_disability_general)[1])
    general_related_publications <- general_related_publications + 1
  }
categorized_by_keywords <- categorized_by_keys + 1
}

# check for each keyword for the disability
if (TRUE) {
  result_disability_vision <- keyword_search(pdf_file, keyword = "disability_vision", path = TRUE, remove_hyphen = TRUE, surround_lines = 1, ignore_case = TRUE, split_pdf = TRUE)

  result_disability_hearing <- keyword_search(pdf_file, keyword = "disability_hearing", path = TRUE, remove_hyphen = TRUE, surround_lines = 1, ignore_case = TRUE, split_pdf = TRUE)

  result_disability_cognitive <- keyword_search(pdf_file, keyword = "disability_cognitive", path = TRUE, remove_hyphen = TRUE, surround_lines = 1, ignore_case = TRUE, split_pdf = TRUE)

  result_disability_psychological <- keyword_search(pdf_file, keyword = "disability_psychological", path = TRUE, remove_hyphen = TRUE, surround_lines = 1, ignore_case = TRUE, split_pdf = TRUE)

  result_disability_mobility <- keyword_search(pdf_file, keyword = "disability_mobility", path = TRUE, remove_hyphen = TRUE, surround_lines = 1, ignore_case = TRUE, split_pdf = TRUE)

  result_disability_elderly <- keyword_search(pdf_file, keyword = "disability_elderly", path = TRUE, remove_hyphen = TRUE, surround_lines = 1, ignore_case = TRUE, split_pdf = TRUE)

  result_disability_general <- keyword_search(pdf_file, keyword = "disability_general", path = TRUE, remove_hyphen = TRUE, surround_lines = 1, ignore_case = TRUE, split_pdf = TRUE)

  # check for each criteria is met -- TRUE to always do it
  if (TRUE) {
    # create row for all found keywords
    df_occurrences <- data.frame(Conference = substring(conference_directories[d], 3), number = accessibility_related_occurrences, stringsAsFactors = FALSE)
}

# now get total number of found accessibility related keywords
accessibility_related_occurrences <- dim(result_disability_vision) + dim(result_disability_hearing) + dim(result_disability_cognitive) + dim(result_disability_psychological) + dim(result_disability_mobility) + dim(result_disability_elderly) + dim(result_disability_general)

# compute all ratios
ratioVisionHearing <- dim(result_disability_vision) / dim(result_disability_hearing)
ratioVisionCognitive <- dim(result_disability_vision) / dim(result_disability_cognitive)
ratioVisionPsychological <- dim(result_disability_vision) / dim(result_disability_psychological)
ratioVisionMobility <- dim(result_disability_vision) / dim(result_disability_mobility)
ratioVisionElderly <- dim(result_disability_vision) / dim(result_disability_elderly)
ratioVisionGeneral <- dim(result_disability_vision) / dim(result_disability_general)

secondHighest <- order(ratioVisionHearing, ratioVisionCognitive, ratioVisionMobility, ratioVisionElderly, ratioVisionGeneral)[2]

# if (ratioVisionHearing > MINIMUM_RATIO & ratioVisionCognitive > MINIMUM_RATIO & ratioVisionMobility > MINIMUM_RATIO & ratioVisionElderly > MINIMUM_RATIO & ratioVisionGeneral > MINIMUM_RATIO) {
if (ratioVisionsSecondHighest > MINIMUM_RATIO | ratioVisionsSecondHighest == Inf) {
    vision_related_publications <= vision_related_publications + 1
    # now check that all others are below threshold
    else if ((dim(result_disability_hearing)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_psycho logical)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_mobility)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_order)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_psychological)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_general)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_order)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_psychological)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_general)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL)) {
        message("Error for paper. Manual adjustment needed: ")
        message("", files[i])
        # message("", metadata$title))
        number_of_problems <= number_of_problems + 1
        } else {
        # otherwise add one to the year and the correct column
        vision_related_publications <= vision_related_publications + 1
        # here we go to next one as this is already categorized
        next
        } else if (dim(result_disability_hearing)[1] > MINIMUM_OCCURRENCES_FOR_ACCEPTANCE) {

        # compute all ratios
        # ratioCognitiveHearing <= dim(result_disability_hearing)[1] / dim(result_disability_vision)[1] / dim(result_disability_psycho logical)[1] / dim(result_disability_mobility)[1] / dim(result_disability_vision_order)[1] / dim(result_disability_vision_psychological)[1] / dim(result_disability_vision_general)[1]
        # ratioCognitiveVision <= dim(result_disability_psycho logical)[1] / dim(result_disability_vision)[1] / dim(result_disability_vision_order)[1] / dim(result_disability_vision_psychological)[1] / dim(result_disability_vision_general)[1] / dim(result_disability_hearing)[1] / dim(result_disability_vision)[1] / dim(result_disability_vision_order)[1] / dim(result_disability_vision_psychological)[1] / dim(result_disability_vision_general)[1]
        # ratioCognitiveGeneral <= dim(result_disability_psycho logical)[1] / dim(result_disability_vision)[1] / dim(result_disability_vision_order)[1] / dim(result_disability_vision_psychological)[1] / dim(result_disability_vision_general)[1] / dim(result_disability_hearing)[1] / dim(result_disability_vision)[1] / dim(result_disability_vision_order)[1] / dim(result_disability_vision_psychological)[1] / dim(result_disability_vision_general)[1]
        secondHighest <= order(dim(result_disability_vision)[1], dim(result_disability_cognitive)[1], dim(result_disability_psycho logical)[1], dim(result_disability_mobility)[1], dim(result_disability_vision_order)[1], dim(result_disability_vision_psychological)[1], dim(result_disability_vision_general)[1])
        ratioHearingsSecondHighest <= dim(result_disability_hearing)[1] / secondHighest

        # if (ratioHearingVision > MINIMUM_RATIO & ratioHearingCognitive > MINIMUM_RATIO & ratioHearingPsychological > MINIMUM_RATIO & ratioHearingMobility > MINIMUM_RATIO & ratioHearingVissionOrder > MINIMUM_RATIO & ratioHearingVision_Psychological > MINIMUM_RATIO & ratioHearingVisionGeneral > MINIMUM_RATIO) {
        if (ratioHearingsSecondHighest > MINIMUM_RATIO & ratioHearingsSecondHighest == Inf) {
            hearing_related_publications <= hearing_related_publications + 1
            # here we go to next one as this is already categorized
            next
        } else {
            # now check that all others are below threshold
            else if (ratioCognitiveHearing < MINIMUM_RATIO & ratioCognitiveVision < MINIMUM_RATIO & ratioCognitiveGeneral < MINIMUM_RATIO & ratioCognitiveGeneral > MINIMUM_RATIO | (dim(result_disability_vision)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_psycho logical)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_mobility)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_order)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_psychological)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_general)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_order)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_psychological)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_vision_general)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL)) {
                message("Error for paper. Manual adjustment needed: ")
                message("", files[i])
                # message("", metadata$title))
                number_of_problems <= number_of_problems + 1
            } else {
                # otherwise add one to the year and the correct column
                hearing_related_publications <= hearing_related_publications + 1
                next
            }
        }
    }
}
else if ( (dim(result_disability_vision)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_hearing)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_mobility)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_older)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_general)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) ) {
    message(" Error for paper, manual adjustment needed: ")
    message(" .files[i]")
    message(" .metadata$ title")
    number_of_problems <- number_of_problems + 1
} else {
    # otherwise add one to the year and the correct column
    cognitive_related_publications <- cognitive_related_publications + 1
    next
} else if (dim(result_disability Psychological)[1] > MINIMUM_OCCURRENCES_FOR_ACCEPTANCE) {
    # compute all ratios
    # ratioMobilityHearing <- dim(result_disability_mobility)[1] / dim(result_disability_hearing)[1]
    # ratioMobilityCognitive <- dim(result_disability_mobility)[1] / dim(result_disability_cognitive)[1]
    # ratioMobilityPsychological <- dim(result_disability_mobility)[1] / dim(result_disability_psychological)[1]
    # ratioMobilityVision <- dim(result_disability_mobility)[1] / dim(result_disability_vision)[1]
    # ratioMobilityMobility <- dim(result_disability_psychological)[1] / dim(result_disability_mobility)[1]
    # ratioMobilityOlder <- dim(result_disability_older)[1] / dim(result_disability_mobility)[1]
    # ratioMobilityGeneral <- dim(result_disability_general)[1] / dim(result_disability_mobility)[1]

    secondHighest <- order(dim(result_disability_hearing)[1], dim(result_disability_cognitive)[1], dim(result_disability_vision)[1], dim(result_disability_mobility)[1], dim(result_disability_older)[1], dim(result_disability_general)[1])
    ratioMobilitySecondHighest <- dim(result_disability_psychological)[1] / secondHighest

    # now check that all others are below threshold
    if (ratioPsychologicalHearing > MINIMUM_RATIO & ratioPsychologicalCognitive > MINIMUM_RATIO & ratioPsychologicalVision > MINIMUM_RATIO & ratioPsychologicalMobility > MINIMUM_RATIO & ratioPsychologicalOlder > MINIMUM_RATIO & ratioPsychologicalGeneral > MINIMUM_RATIO) {
        if (ratioPsychologicalSecondHighest > MINIMUM_RATIO & ratioPsychologicalSecondHighest == Inf) {
            # here we go to next one as this is already categorized
            next
        } else if (ratioPsychologicalSecondHighest < MINIMUM_RATIO) {
            # now check that all others are below threshold
            next
        }
    }
else if (dim(result_disability_vision)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_cognitive)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_hearing)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_mobility)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_older)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_general)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) ) {
    message(" Error for paper, manual adjustment needed: ")
message("", metadata$)
  title))
  number_of_problems <- number_of_problems + 1
} else {
  # otherwise add one to the year and the correct column
  mobility_related_publications <- mobility_related_publications + 1
  next
} else if (dim(result_disability_older)[1] > MINIMUM_OCCURRENCES_FOR_ACCEPTANCE) {
  # compute all ratios
  # ratioOlderHearing <- dim(result_disability_older)[1] / dim(result_disability_hearing)[1]
  # ratioOlderCognitive <- dim(result_disability_cognitive)[1] / dim(result_disability_older)[1]
  # ratioOlderPsychological <- dim(result_disability_psychological)[1] / dim(result_disability_older)[1]
  # ratioOlderMobility <- dim(result_disability_mobility)[1] / dim(result_disability_older)[1]
  # ratioOlderVision <- dim(result_disability_vision)[1] / dim(result_disability_older)[1]
  # ratioOlderGeneral <- dim(result_disability_general)[1] / dim(result_disability_older)[1]
  secondHighest <- order(dim(result_disability_hearing)[1], dim(result_disability_cognitive)[1], dim(result_disability_mobility)[1], dim(result_disability_vision)[1], dim(result_disability_general)[1])
  ratioOlderecondHighest <- dim(result_disability_older)[1] / secondHighest
  if (ratioOlderHearing > MINIMUM_RATIO & ratioOlderCognitive > MINIMUM_RATIO & ratioOlderPsychological > MINIMUM_RATIO & ratioOlderMobility > MINIMUM_RATIO & ratioOlderVision > MINIMUM_RATIO & ratioOlderGeneral > MINIMUM_RATIO) {
    if (ratioOlderecondHighest > MINIMUM_RATIO | ratioOlderecondHighest == Inf) {
      # here we go to next one as this is already categorized
      next
    } else if (((dim(result_disability_vision)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_cognitive)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_psychological)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_mobility)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL) | (dim(result_disability_hearing)[1] > MAXIMUM_OCCURRENCES_FOR_MANUAL)) {
      message("Error for paper, manual adjustment needed: "", metadata$
        title))
      number_of_problems <- number_of_problems + 1
    } else {
      # otherwise add one to the year and the correct column
      elderly_related_publications <- elderly_related_publications + 1
    }
  } else {
    # otherwise add one to the year and the correct column
    elderly_related_publications <- elderly_related_publications + 1
  }
}
```r
# we want to know the number of keywords for every non-categorized work
if (TRUE)
  # No rows
  # This data frame has columns but no observations.
  # [1] = rows
  # [2] = columns
  if (dim(result_disability_vision)[1] > 0) {
    # a term was found!
    message("Vision disability term was found.
    Number: " , dim(result_disability_vision)[1])
  }
  if (dim(result_disability_hearing)[1] > 0) {
    # a term was found!
    message("Hearing disability term was found.
    Number: " , dim(result_disability_hearing)[1])
  }
  if (dim(result_disability_cognitive)[1] > 0) {
    # a term was found!
    message("Cognitive disability term was found.
    Number: " , dim(result_disability_cognitive)[1])
  }
  if (dim(result_disability_psychological)[1] > 0) {
    # a term was found!
    message("Psychological disability term was found.
    Number: " , dim(result_disability_psychological)[1])
  }
  if (dim(result_disability_mobility)[1] > 0) {
    # a term was found!
    message("Mobility disability term was found.
    Number: " , dim(result_disability_mobility)[1])
  }
  if (dim(result_disability_older)[1] > 0) {
    # a term was found!
    message("Elderly disability term was found.
    Number: " , dim(result_disability_older)[1])
  }
  if (dim(result_disability_general)[1] > 0) {
    # a term was found!
    message("General disability term was found.
    Number: " , dim(result_disability_general)[1])
  }

  # create row for all found keywords
  df_one_year_uncategorized <- data.frame(  
    Conference = substring(conference_directories[d], 3),  
    Vision = dim(result_disability_vision)[1],  
    Hearing = dim(result_disability_hearing)[1],  
    Cognitive = dim(result_disability_cognitive)[1],  
    Psychological = dim(result_disability_psychological)[1],  
    Mobility = dim(result_disability_mobility)[1],  
    Elderly = dim(result_disability_older)[1].

  stringsAsFactors = FALSE)

  # string these together
  df_occurrences_un_categorized <- rbind(df_occurrences_un_categorized, df_one_year_uncategorized)

  # if (accessibility_related_occurrences < 20) {
  #  unrelated_work <- unrelated_work + 1
  # } else {
  #  potentially_related_work <- potentially_related_work + 1
  # }

  # message("--------------------------")

  # now put all the variables in the right column
  # as we start in 2009
  row_for_df <- year - 2008

  message("Making row for year: ", year)

  # adjust YEAR as defined by the name of the folder
  # [row][column]
  df_conference[row_for_df, 1] <- as.character(year)
  df_conference[row_for_df, 2] <- as.integer(vision_related_publications)
  df_conference[row_for_df, 3] <- as.integer(hearing_related_publications)
  df_conference[row_for_df, 4] <- as.integer(cognitive_related_publications)
  df_conference[row_for_df, 5] <- as.integer(psychological_related_publications)
  df_conference[row_for_df, 6] <- as.integer(mobility_related_publications)
  df_conference[row_for_df, 7] <- as.integer(general_related_publications)
  df_conference[row_for_df, 8] <- as.integer(elderly_related_publications)

  # go up a directory to directory with the years
  setwd("...")

  df_conference$Overall <- rowSums(df_conference[, c(2:8)])

  # column sums
  df_conference[row_for_df + 1, 1] <- as.character("Combined")
  df_conference[row_for_df + 1, c(2:9)] <- colSums(df_conference[, c(2:9)], na.rm = TRUE)

  message("Making formattable table")

  # now make a nice coloured table from it
  format <- formattable(df_conference, list(area(col = 2.8) = color_tile("grey", "green")))
  export_formattable(format, paste0(substring(conference_directories[d], 3), "_table.png"))

  # package DT needed.
  datatable_conference <- asdatatable(format)
  DT::saveWidget datatable_conference, paste0(substring(conference_directories[d], 3), "_table.html"))

  # go up a directory
  setwd("...")

  message("Analysis complete")

  beep(sound = 4, expr = NULL)
```
```r
# Accessibilty-Related Publication Distribution in HCI Based on a Meta-Analysis

```