Developing Accessible Collection and Presentation Methods for Observational Data

Drew Marie Williams

Marquette University

Recommended Citation
https://epublications.marquette.edu/dissertations_mu/836
DEVELOPING ACCESSIBLE COLLECTION AND PRESENTATION METHODS FOR OBSERVATIONAL DATA

by

Drew Williams, B.S., M.S.

A Dissertation submitted to the Faculty of the Graduate School, Marquette University, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Milwaukee, WI, USA

December 2018
ABSTRACT
DEVELOPING ACCESSIBLE COLLECTION AND PRESENTATION METHODS
FOR OBSERVATIONAL DATA

Drew Williams, B.S., M.S.
Marquette University, 2018

The processes of collecting, cleaning, and presenting data are critical in ensuring the proper analysis of data at a later date. An opportunity exists to enhance the data collection and presentation process for those who are not data scientists – such as healthcare professionals and businesspeople interested in using data to help them make decisions. In this work, creating an observational data collection and presentation tool is investigated, with a focus on developing a tool prioritizing user-friendliness and context preservation of the data collected. This aim is achieved via the integration of three approaches to data collection and presentation.

In the first approach, the collection of observational data is structured and carried out via a trichotomous, tailored, sub-branching scoring (TTSS) system. The system allows for deep levels of data collection while enabling data to be summarized quickly by a user via collapsing details. The system is evaluated against the stated requirements of usability and extensibility, proving the latter by providing examples of various evaluations created using the TTSS framework.

Next, this approach is integrated with automated data collection via mobile device sensors, to facilitate the efficient completion of the assessment. Results are presented from a system used to combine the capture of complex data from the built environment and compare the results of the data collection, including how the system uses quantitative measures specifically. This approach is evaluated against other solutions for obtaining data about the accessibility of a built environment, and several assessments taken in the field are compared to illustrate the system’s flexibility. The extension of the system for automated data capture is also discussed.

Finally, the use of accessibility information for data context preservation is integrated. This approach is evaluated via investigation of how accessible media entries improve the quality of search for an archival website. Human-generated accessibility information is compared to computer-generated accessibility information, as well as simple reliance on titles/metadata. This is followed by a discussion of how improved accessibility can benefit the understanding of gathered observational data’s context.
I want to thank Dr. Praveen Madiraju, Dr. Roger O. Smith, Dr. Stephen Guastello, Dr. Thomas Kaczmarek, and my dissertation advisor, Dr. Sheikh Iqbal Ahamed for mentoring me and assisting me throughout the process of writing this dissertation. Without your help, I certainly wouldn’t be where I am today. I would also like to thank Dr. Gary Krenz and all the teachers, faculty, and staff that I have gotten to know during my time at Marquette University.

Thank you to my brilliant friends, who kept me motivated and were happy to hang out at odd hours in the night and chat to help me de-stress. Furthermore, thank you to my wonderful mother, father, and brother for all their love and support throughout my time pursuing a Ph.D. at Marquette University – without their help, I would certainly have not gotten as far as I have today.

Finally, I would like to give a special thanks to Jordan Kozmary, for his love, support, and all those times he brought me coffee in the dead of night while I worked away on drafts of papers. I am fortunate to have someone in my life as perfect as he is.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ................................................................................................................. i
LIST OF TABLES ............................................................................................................................. vi
LIST OF FIGURES ............................................................................................................................. vii

CHAPTER
I. INTRODUCTION ......................................................................................................................... 1
   A. Data Collection for Decision-Making ...................................................................................... 2
   B. The Interconnection of Presentation and Analysis ................................................................. 4
   C. Creating More Usable Data Science Tools .............................................................................. 6
      1) Simplification of Data Acquisition and Presentation Process ......................................... 8
      2) Integration of Quantitative Data Acquisition ................................................................... 9
      3) Integration of Media Presentation and Context ................................................................. 11
   D. Research Questions and Aims ............................................................................................... 12
      1) Structuring Data Collection and Presentation with Hierarchies ................................. 13
      2) Integrating Quantitative Measures into Data Collection .............................................. 14
      3) Ensuring Media Context via Integration of Accessibility Data ...................................... 15
   E. Dissertation Outline ............................................................................................................... 15

II. RELATED WORK ......................................................................................................................... 18
   A. Data Visualization ............................................................................................................... 18
   B. Big Data and Data Analytics .............................................................................................. 21
   C. Information Visualization .................................................................................................... 22
   D. Decision Theory .................................................................................................................. 24
   E. Computer-Assisted Data Collection .................................................................................... 26
F. Universal Design as Usability Guidelines ...................................................... 29
   1) Creating a Universal Design ................................................................. 31

G. Differences Between Existing Work and This Solution .................... 35

III. STRUCTURING DATA COLLECTION AND PRESENTATION WITH HIERRARCHIES ................................................................. 37

A. Challenges in Observational Data Collection ............................................. 38
   1) Weather Severity .................................................................................... 39
   2) Software Accessibility for Section 508 ..................................................... 39

B. The Design and Development of OTFACT ............................................... 40
   1) User Interface Design ............................................................................ 41
   2) Scoring ................................................................................................... 43
   3) Filtering ................................................................................................... 45

C. TTSS Assessments: Past and Present ......................................................... 46
   1) RATE-IT ............................................................................................... 47
   2) MED-AUDIT .......................................................................................... 48

D. xFACT for TTSS Assessment Development ............................................. 48
   1) Main Application .................................................................................... 49
   2) Editor ..................................................................................................... 50
   3) Accessory Applications .......................................................................... 52

E. Case Studies for Using xFACT ................................................................. 52
   1) Severity of Weather ............................................................................... 53
   2) Accessibility of Software ....................................................................... 57

F. Discussion .................................................................................................... 59

G. Summary ..................................................................................................... 62
IV. AUGMENTING HIERARCHY-BASED DATA COLLECTION AND PRESENTATION WITH QUANTITATIVE INFORMATION

A. Data Collection using Mobile Device Sensors ................................ 64

B. Integration of an Assessment with Quantitative Data Collection ...... 66

C. The Access Ratings Project: Building Apps for Built Environment Accessibility................................................................. 67

   1) Creating the Mini-Tools .................................................. 68

   2) AccessTools....................................................................... 72

D. An AccessTools Workflow ..................................................... 75

E. Comparison of Data Taken with AccessTools ............................. 77

F. Discussion ............................................................................. 80

   1) Accessibility ....................................................................... 81

   2) Integration of AccessTools Data with Consumer Website ..... 81

G. Summary ................................................................................ 84

V. ENHANCING MEDIA COLLECTION AND PRESENTATION VIA UNIVERSAL DESIGN ........................................................................ 86

A. Understanding Media Content and Context ............................... 88

   1) Understanding Image Content .......................................... 88

   2) Understanding Audio Content ......................................... 89

   3) Understanding Video Content ......................................... 89

B. Universal Design and Media ................................................... 90

   1) Accessibility and Images ................................................. 91

   2) Accessibility and Audio .................................................. 93

   3) Accessibility and Video .................................................. 94

C. Improving the Searchability of Information by Including Accessibility Information ............................................................. 95
D. Results ........................................................................................................... 101
   1) Research Quality ................................................................. 101
   2) Universal Design ................................................................. 102
   3) Media Context ......................................................................... 102
E. Discussion ................................................................................................. 106
F. Summary ................................................................................................... 107

VI. DISCUSSION ............................................................................................. 108
    A. Implications of Work .......................................................... 110
    B. Strengths of System Design .............................................. 111
    C. Weaknesses of System Design .......................................... 112
    D. Future Work ........................................................................... 114

VII. CONCLUSION ......................................................................................... 116
     A. Contributions to the Field .................................................. 116
     B. Broader Impact ...................................................................... 117

VIII. REFERENCES .......................................................................................... 118

IX. PUBLICATIONS ........................................................................................ 131

APPENDIX 1: EXPLORING THE WORKFLOW OF THE xFACT TAXONOMY EDITOR ................................................................. 135

APPENDIX 2: DEVELOPMENT OF A TAXONOMY FOR OBSERVING WEATHER SEVERITY ................................................................. 137

APPENDIX 3: DEVELOPMENT OF A TAXONOMY FOR ASSESSING VPAT COMPLIANCE ................................................................. 141

APPENDIX 4: A SAMPLE OF IMAGE DATA FROM THE FRED SAMMONS ARCHIVES ................................................................. 145
LIST OF TABLES

Table 2-1: Examples of universal design provisions and populations helped.............33
Table 3-1: A sample of other projects using a TTSS-styled assessment for data collection and their complexity ..................................................46
Table 4-1: Mini-tools and sensors used from the past and present.........................68
Table 5-1: Media types and counts available in the Fred Sammons Archive.............92
Table 5-2: Media, related accessibility information, and searchability.....................92
LIST OF FIGURES

Figure 1-1: Depiction of the potential types of data available from a smartphone………3
Figure 1-2: A visualization of this work’s research goals……………………………….16
Figure 2-1: Playfair’s chart of imports and exports to/from Denmark and Norway……18
Figure 3-1: A screenshot of OTFACT’s assessment view…………………………………41
Figure 3-2: A diagram illustrating branching and scoring in a TTSS-styled assessment..44
Figure 3-3: A screenshot of the assessment view from RATE-IT…………………………..47
Figure 3-4: A screenshot of the taxonomy selection screen from the xFACT taxonomy editor…………………………………………………………………………………….51
Figure 3-5: Example of a weather severity taxonomy in xFACT via the taxonomy editor……………………………………………………………………………………56
Figure 3-6: A diagram illustrating the filtering mechanisms built into an xFACT assessment……………………………………………………………………………59
Figure 4-1: A diagram illustrating the data complexity involved when taking 1-2 measurements for a series of questions in an assessment…………………………………63
Figure 4-2: An example of the capabilities of Google Translate, an app for the iPhone..65
Figure 4-3: Various screenshots from AccessSlope…………………………………….70
Figure 4-4: A screenshot of the main assessment view from AccessTools……………….73
Figure 4-5: Example of slope measurement within AccessTools……………………….74
Figure 4-6: A graph representing the measurements, comments, and data collected using AccessTools……………………………………………………………………77
Figure 4-7: Screenshots of two completed assessments in AccessTools………………….79
Figure 4-8: Preliminary consumer view for AccessTools Data…………………………83
Figure 5-1: A screenshot of a TED talk video available on Youtube, with captions…….94
Figure 5-2: Example of image available from Fred Sammons Archive………………….98
Figure 5-3: A screenshot illustrating the integration of accessibility options into an artifact view

Figure 5-4: A screenshot of the front page of the Fred Sammons website

Figure 6-1: A representation of the work carried out by a TTSS system
I. INTRODUCTION

People nowadays traverse a world that bends to their likes and dislikes. Music streaming services generate playlists based on a users’ date of birth [1] and listening history. [2] New brunch and dinner spots can be discovered based on a user's current location and desired filters. [3] Health insurance rates dip relative to one’s wellness, as reported by sensors worn on their body. [4] While each of these applications are entirely different, all of them are a direct result of the ability to analyze quantities of data and derive conclusions from it.

Indeed, data science has become quite popular in recent years, and offers the individual quite a few benefits! As a result, many people are in search of professionals to help them collect, clean and analyze data. Data scientists have enjoyed having the "best job in America" for several years now. [5] However, as the field grows in popularity, concerns about overreliance on data mount as well - especially as incidents of bias in algorithms mount. Though many believe models derived from data will prevent biased results, these models do not necessarily remove bias from their data. [6] How can one improve the process of deriving decisions and predictions from data, to ensure that data science continues to benefit humanity instead of burdening it? Furthermore, as data science is an interdisciplinary field, how can one improve the user experience such that all people can effectively gather and process data, regardless of their experience with computers?
An excellent place to start is by enhancing the foundation of the data science process - the act of collecting data for decision-making. If data gathered is sound, conclusions drawn from data are more likely to be sound as well. Ensuring sound data requires a good look at methods of collecting and presenting data.

A. Data Collection for Decision-Making

Algorithms are reliant on the ability to gather and parse large amounts of data. Delving into data to make judgments and predictions has become more popular as computers' available bandwidth, memory, data storage, and processing power became cheaper, facilitating the collection and analysis of large quantities of data. [7] In particular, it became easier to collect data that is plentiful (high-volume), quickly gathered (high-velocity), and diverse (high-variety) - often called "big data." [8] The rise of mobile computing and the internet of things also fed into this explosion of data, by allowing a user to collect types of data that had been unavailable in years prior, such as location information [9] and user activity. [10] Indeed, with smartphones as ubiquitous as they are nowadays, [11] collecting data observed by a user about their status, or the state of something in their environment, can be initiated quickly and easily.

So while ‘big data’ involves the collection of data that is too voluminous to handle with everyday tools, such as genomic data or clickstream data, [12] this is not the only sort of data collected. Crowdsourced data – data one accumulates via multiple users collecting data simultaneously – is another popular method of aggregating large
quantities of data, such as the mPING project’s crowdsourced observations of winter precipitation data gathered to analyze the performance of their algorithms. [13] This data relies on average people for collection, instead of trained data scientists. However, big or small, complex data is typically collected nowadays, rich in information and easy to obtain via smartphones. Images, audio, tweets, logs and sensor readings are all fair game for data collection and subsequent processing. [12] [14] Likewise, capturing qualitative data remains essential even in a world of automated data capture from sensors - especially in the form of customer interviews and comments, or social media posts. [14]
B. The Interconnection of Presentation and Analysis

Gathering this data is just the start of this process - as Gandomi and Haider (2015) stated, “big data are worthless in a vacuum.” [8] After one collects data, two steps lead to uncovering meaning: cleaning and presenting data, and analyzing data, wherein one develops a model and interpretations of the data. [8] Data visualization is not to be confused with the presentation step: visualization is often conducted with a hypothesis in mind, while presentation is not. Presentation may include acts such as data profiling; wherein one examines data from an existing source and produces summary statistics about that data. [15] Unfortunately, profiling, (among other methods of processing data before analysis) is designed first for stationary and structured data, typically in a table. [14] This makes it difficult for use with complex data. Additionally, critics of summary statistics fear they may hide interesting features of data, such as skews and distortions. [16]

Though it is not as glamorous as model creation, presenting data can help a user understand collected data, before making decisions with it. In doing this, hidden patterns and trends that could spark hypotheses or conclusions may also become apparent in the data. An example of this in practice is in the field of business intelligence, which focuses on the analysis of data captured by companies. [17] Say an auto insurance company leveraged data collection to understand how they might charge people in different locations for insurance. By visualizing the occurrences of claims and their sites (perhaps claims mapped to zip codes) they could see if significant discrepancies existed and make
a decision on where to charge more money and adjust their pricing accordingly (something insurance companies do in practice). [18] This use of data presentation is also useful in cases when one has to explain data to an interdisciplinary audience, given proper attention to user experience. [19] Rather than stating that one should charge a location more money, a data analyst can provide charts and graphs of real data to business executives as proof.

Understanding collected data also lends a hand in preventing (or at least understanding) potential bias in data. There is a risk in gathering the wrong data or insufficient sets of data, and proper data preparation or visualization can help catch cases such as these. [20] If the information is problematic, the conclusions or predictions made from that data will be weak – a principle often referred to as “garbage in, garbage out.” [21] Zunger (2017) brings this point forward when discussing how data relates to models. [22]

“AI models hold a mirror up to us; they don’t understand when we really don’t want honesty. They will only tell us polite fictions if we tell them how to lie to us ahead of time.” [22]

Data presentation is also a useful technique to utilize in the case of problems that machine learning cannot effectively address; typically those of summarization of existing data, or the examination of the state of a set of data. A good example of where this is important is in parsing and summarizing qualitative data. Much work has gone into developing automated text summaries, or “AI-powered tl;dr,” but researchers express that if anything, their efforts in this field have illustrated the limitations of relying solely on
machine learning for summarization efforts. [23] [24] In this case, viewing some structured summary of collected data would be beneficial. While data summaries of quantitative data are commonly done as a part of data profiling, [15] summarizing qualitative data can be quite a bit trickier for individuals to do. However, these data summaries can come in handy in cases where the user wants to feel in control of their data, and conclusions drawn from it – something interactions controlled by a computer often don’t offer. [25] One such example is the field of assistive technology, where doctors prefer to draw their conclusions from patient data. Here, many patients with the same diagnoses may end up preferring different solutions, due to their lifestyles and preferences. Although over time (and with additional data collection) systems may be able to predict preferred solutions automatically, consultation is a required step in the process of assigning a patient with assistive technology at this time.

C. Creating More Usable Data Science Tools

While many tools exist which help a user interested in data analysis, [26], [27] an opportunity exists for standardizing a process for gathering complex data and guiding a user with a goal through the means of data collection and presentation for rich hypothesis generation prior to modeling. [16] Though data collection is increasingly automated, some data needs, such as complex and crowdsourced data, rely on collection by people.

For this work, the problem of collecting and presenting observational data – data gathered when observing a state – is the focus. Observational data can be qualitative,
such as in the case of field notes and media recordings. [28] It can also be quantitative, where one uses standardized instruments to collect data during observation. [28]

The user experience related to this part of the data analysis process is especially important, as many times those who collect data (and thus, those who need to understand what they are gathering) are not trained data scientists. Consider a doctor taking medical data from a patient – a robust tool for standardizing the process of symptom data collection and presenting the data collected would help them proceed in choosing a treatment. The doctor could add this data and the resulting treatment to a more extensive database, for easy comparison of symptom data and treatment information across multiple patients. The standardized nature of the procedure would help ensure the quality of the data collected.

Through focusing on accessibility, an opportunity also exists to improve the accessibility of the data collection and presentation processes, making the process of data collection and visualization easier for individuals with and without disabilities. While strides have been made in setting requirements for accessible data visualizations, [29] the tools used to create those visualizations are often quite complex. To simplify the user experience – for both data scientists and non-data scientists - the idea of universal design is integrated. Chisholm and May (2008) describe universal design as being usable by “…all people, to the greatest extent possible.” [30] Focusing on creating a universal design develops a method of collecting data that is usable by everyone, regardless of
background or ability – thus improving the experience not just for people with disabilities, but for all.

Finally, standardizing the process of collecting observational data would give those collecting data a format to become accustomed to – and thus use in the future. Consistent UI is more usable, [31] and a consistent interface will help non-data scientists feel more at ease during the data collection process.

In creating a proper data collection and presentation tool, the focus should be on the following goals: a simplified data acquisition stage combined with data presentation, integrating complex data into the collection process, and the integration of accessibility information for context preservation. Achieving these goals structures the process of data collection for a positive user experience.

1) Simplification of Data Acquisition and Presentation Process

The first problem with data collection is that of planning: how does one guide a user through the process of collecting data?

The chosen methodology for this problem centers around a structure designed by Smith (2002), called a trichotomous tailored sub-branching scoring structure. [32] Typically, interpreted observational data falls into one of three categories: a positive conclusion, a negative conclusion, or an uncertain one. This sort of scale is trichotomous, or one with responses which fall into one of three types. [32] For observational data, this scale can simplify a plan of data collection into an assessment. Most things in the world
are, are not, or perhaps are something, making these answers surprisingly effective classifiers. If the data collector is uncertain about whether or not something is or is not a way, further details can be requested to help with decision-making. Following this method also has the following effects on data collection:

- Building the data collection plan as an assessment develops a protocol for data collection from the start.
- The state of something can be easily quantified: the “yes,” “no,” and “maybe” answers to questions can easily convert to 2/0/1 scores, respectively. [32]
- Collected data forms a hierarchy, where questions branch to have details if needed. [32]
- During data collection, items not relevant to immediate needs are filtered – there is no need to ask for additional detail if certainty is possible. [32]

Perhaps the greatest boon of using such a system is the fact that the hierarchical assortment of data allows one to quickly see “summaries” of data, and drill down into details. Summaries – held at the top parent nodes – indicate patterns of weak or strong scoring in their sections. If a user wants to see detailed scoring information, they can drill down into particular sections to see specific problems. [32]

2) Integration of Quantitative Data Acquisition

There is a case where the method described prior could be a problem: the case of adding complex data to this collection process – such as quantitative measures, and media. Imagine a question with the following hierarchy:
• Is this ramp accessible?
  ▪ Is the slope of the ramp accessible?
    o Is the height of the ramp within an accessible range?
    o Is the length of the ramp within an accessible range, relative to the height?

These items can be answered in a trichotomous manner, but it would make more sense to avoid the two most detailed questions and ask the user for a slope measurement, to prevent user error. [33] In this way, a TTSS assessment can structure the collection of quantitative data, not just observations. [21] Furthermore, the hierarchical structure of the assessment can help provide additional context to data collection, reflecting the categories/hierarchies that are part of a data collection step.

Building quantitative data collection into a data collection approach goes beyond merely making room for a user to input measurements, however. Many methods of retrieving measurement data from the environment exist, such as:

• Sensors and sensor networks (IoT networks)
• Mobile phones
• Other data streams

This sensor data collection can be integrated into the assessment process. For example, a measurement result might assist in answering a question in an assessment (whether or not a ramp is accessible), while retaining proof of why the response was marked as such. Folding complex data collection such as this into the overall assessment
ensures a simple interface for the user that does not involve any outside step of data collection via their means.

3) Integration of Media Presentation and Context

When integrating complex data, there is often a need to integrate media into a data collection process – images, audio, video, and the like. Media can be useful in the case of observational data, as “proof” of stated claims – similar to how data is used to answer questions definitively. However, while solutions exist nowadays to automatically recognize the contents of images and create alt-text accordingly, [34] [35] in addition to automatically captioning video, [36] discerning the context of an image is a remarkably hard problem. Media can have many reasons for being included in data collection – historical purpose, offering proof of a statement, or giving additional background. However, taking a simple course of action can often provide the amount of metadata needed to replicate the context of the media addition: ensuring that accessibility data exists for the added media.

Accessibility refers to how usable something is by individuals with disabilities – an act of designing that takes into consideration that many people have abilities that are not the norm. This definition has also been in focus in recent years as a recent ruling deemed that if the website of a business was not accessible, the business could be sued under the ADA. [104] Thus, in order to avoid alienating a portion of the worldwide population (and being sued), one should set aside time for implementing accessible features in their applications.
The census categorizes disability in three distinct domains: communicative disability (such as blindness, deafness, or difficulty having one’s speech understood), cognitive impairments, and disabilities in the physical domain (such as having trouble walking, or use of a wheelchair). [37] Physical disabilities may also extend to difficulties using the hands and arms, potentially caused by to paralysis, stroke, or missing limbs. [37] Typically, data representation accessibility has involved activities like ensuring that visualizations are understandable by individuals who are colorblind, have decreased visual acuity, or other vision impairments. With the advent of the internet and interactive methods of data representation, considerations when creating data representations increases in scope. However, one should avoid design for one disability or one persona. Disabilities may occur in multiple domains for one patient [37] – someone can be blind and have mobility impairments, as an example. Accessible design should consider a spectrum of ability, rather than an all-or-nothing approach.

It is also worth noting that a focus on accessibility can reduce the cognitive complexity of an interface, proving the universal design adage that designing for disability is designing for everyone. As a good example: consider a case where one wants to change the background and foreground (text) colors of a webpage – the result can be beneficial for users with vision disabilities, in addition to folks with dyslexia. [38] However, the result is just as useful for users who may want to reduce eye strain.

D. Research Questions and Aims
With this all said, the main research problem of this work is simple: understanding how to improve the processes of data collection and representation for all users with improved user interfaces. This problem can be broken down into three questions:

- How might one ensure a data collection process that allows efficient visualization of collected data?
- How might one visualize captured numerical/quantitative data?
- How might one integrate media into this system while preserving media context and universal design?

A system that answers these questions improves the overall act of data presentation and collection for all users, including those who may not be data scientists. Designing a method of collecting only relevant data (so to avoid overcomplicating the data presentation) and presenting complex data with principles of universal design at the forefront ensures that the presented data is available to extract meaning from, regardless of the users’ ability or background.

This system is constructed via fulfillment of three important aims of research over several projects.

1) Structuring Data Collection and Presentation with Hierarchies
Complex data can often be difficult to browse or summarize, primarily due to the amount of data that can be part of a system. Thus to satisfy the first research aim of this work, an application was created for collecting and viewing assessments in an easy-to-understand fashion. The assessments created using this program organize data via a trichotomous tailed sub-branching scoring system for collecting data as it is collected. Assessments each have a specific goal, and the questions within them are designed to obtain context about a users’ goal regarding gathering data at some given time. The scoring reflects an overall picture of a situation, rather than a determination about action to take, such that the user can understand the data gathered and take action to determine how to proceed. Furthermore, the structures of the assessments themselves allow a user to understand the hierarchy of data that they are collecting, and perform a preliminary visualization.

2) Integrating Quantitative Measures into Data Collection

Using the previous system as a base, a strong system for data collection might obtain data from the environment to help the user make a judgment without user intervention. Allowing environmental and quantitative data to be factored into the assessment system also improves the simplicity of the overall system, allowing the user to use it with less prior training than otherwise might be necessary. Thus, to achieve this aim, the development of a mobile application using a TTSS-compatible assessment is discussed. Such an application integrates observational data with environmental measures using mobile device sensors for determining the accessibility of the built
environment. By supplementing assessment questions with external data, the user does not need to remember information like acceptable measurement ranges and can focus on the data collection and the resultant summary for identifying problematic environments for accessibility.

3) Ensuring Media Context via Integration of Accessibility Data

Finally, there is the problem of media context in this system: while the first two aims covered the availability and simplicity of the data collection and presentation process, the problem of integrating data context within the system still exists. The solution is the integration of accessibility data. For a system to fully work for a wide variety of people, it should be fully accessible – that is, have provisions for media to be accessed in multiple ways, so one is not reliant on any single sense. Rather than design for specific disabilities, which is bound to miss a case, each bit of media is accompanied by alternate displays, for the broadest audience it can reach – adding voice transcriptions for handwritten items, captions and audio descriptions for videos, and the like. Equivalent text descriptions are also preferred over alt-text for images, in order to accommodate the full range of context one could draw from a particular image. A visible improvement in the understanding of data is seen when adding this information!

E. Dissertation Outline

In this dissertation, the three aims outlined here are achieved by deploying an assessment system for the collection of observational data, presenting it, augmenting
questions within an assessment with quantitative data from the environment and modify any media within the system to be fully accessible. Following this approach allows the widest variety of people to take advantage of data presented within such a system, and ensures that all data presentations created with this system get their messages across.

In chapter two, related literature will be discussed – notably, related work in the fields of information representation, visual analytics, decision analysis, and universal design and usability. The differences between these approaches and this proposed approach will be summarized.
In chapter three, the creation of a TTSS framework for data assessments, and how it helps ensure that data collection and presentation processes are accessible, will be discussed. Several case studies of designing TTSS assessments for data collection will be provided as examples.

In chapter four, the augmentation of this system with numeric data from the environment is discussed, along with how that assists in improving the simplicity of the system design. As an evaluation of this approach, a software suite for gathering data from the environment regarding built environment accessibility is presented.

In chapter five, accessibility is discussed – specifically how it can improve retaining data context. A case study is provided where a similar approach was used in developing an archival website, and it significantly improved the searchability of media stored as part of the site.

Finally, in chapter six, concluding thoughts regarding the type of design one should consider for accessible data presentation, and discuss this work’s contributions and broader impact.
II. RELATED WORK

Lessons about this proposed approach work can be found in the fields of data visualization, information representation, decision analysis, computer-supported data collection, and universal design. In this section, relevant work is reviewed in various fields. A discussion follows about how the proposed approach builds on and moves beyond this work.

A. Data Visualization

Work in data visualization dates back to 1786 when William Playfair created one of the first visualizations of quantitative information. [39] Playfair preferred his charts to
the tabular methods of data presentation of the time, as the graphics could convey a shape of data that tabular displays could not, saying the following in particular: [40]

“…[o]n inspecting any one of these Charts attentively, a sufficiently distinct impression will be made, to remain unimpaired for a considerable time, and the idea which does remain will be simple and complete…” [40]

Playfair’s graphics led to the creation of classical charting methods.

Strides in the field were few until Bertin, a cartographer from France, developed and published a theory of graphics called *The Semiology of Graphics* (1967) that identified various elements of diagrams, and described a general framework for designing graphics. [39] In the early 1980s, Edward R. Tufte published his first in a series of books in which he discussed information visualization and displaying quantitative discoveries. [40] Among other things, Tufte discussed requirements for graphical displays, such as the necessity for them to show data, make sets of large numbers easily digestible, encourage comparison, and above all reveal data. [40] In follow-up books, he discussed the use of particular techniques – such as the juxtaposition of micro/macro sizing to reveal data comparisons and prudent use of shades of grey instead of overuse of color. [41]

Tufte was also well-known for arguing the importance of spending time to ensure proper information representation. [43] A well-known story illustrating the importance of proper data analytics is that of the spaceship ‘Challenger,’ and the role information visualization played in the famous and deadly complications the shuttle met with on launch day. In January of 1986, engineers onsite asked to reschedule the shuttle’s launch in light of cold temperatures at the launch site. [42] These temperatures, engineers
argued, would damage rubber seals on the shuttle, which could cause hot gas to escape the shuttle and create a dangerous situation for astronauts. [42] NASA officials were reluctant to delay the event, and in order to persuade NASA officials to stop the launch, a data visualization was constructed in order to link O-Ring failure with temperature. [43] Unfortunately, the visualization failed to convey its message effectively, and the shuttle launch went forward – ending in the untimely deaths of all astronauts aboard the shuttle. [42] This is, perhaps, one of the most dramatic examples of a tenant of data analysis: if the data cannot be understood, it is virtually useless. Edward Tufte (1997) stated this in a different manner in his work “Visual and Statistical Thinking.” [43]

“…there are right ways and wrong ways to show data; there are displays that reveal the truth and displays that do not.” [43]

In 1977, Tukey (1977) began an investigation into the potential of exploratory data analysis – emphasizing the ability of images to serve as vectors for quick summaries of data. [39] Much work following this pursued the development of visualization strategies for multivariate and multidimensional data. [39] “An initiative on scientific visualizations” was launched in 1985, in an attempt to understand large data sets that had recently begun to crop up in the physical sciences. [44] Systems were developed to do more complex visualization of these data sets, and studies focused on user interaction with huge datasets. [44] In an effort to categorize the growing number of visualizations, Shneiderman [151] developed a taxonomy to understand them based on the task that is being accomplished in the data visualization process. This work produced a “mantra” of visual information, known as the “Visual Information Seeking Mantra:” [151]
“Overview first, zoom and filter, then details-on-demand.” [151]

To this day, this remains a strong contender for the most significant rule of interface design.

B. Big Data and Data Analytics

As computer use became more widespread, the field of information visualization became increasingly tangled in the field of business analytics – especially as the advent of big data lent itself to helping businesses make decisions and target customers. From 2001 until 2008, researchers developed definitions of “big data” and software applications to handle large datasets. IBM estimates that a whopping 2.5 quintillion bytes of data are generated daily, and as a result, 90% of the data available today was made in the last two years. [45] This has led to the creation of industry-led visualization projects, such as Tableau and TIBCO, and academic visualization projects to spin up alike. [45]

Due to the impressive amount of information contained in big data datasets, new methods of processing are often required in order to organize and understand the data correctly. In many cases, this involves bringing additional context into the data analysis process via the addition of metadata. Unfortunately, no standard method of metadata storage and access exists – causing efficiency problems as people are unable to understand what data has already been collected, and leaving individuals unable to control access to data. [46] A number of solutions to the problem of metadata have been offered. Hellerstein et al. (2017) suggested the use of application context, information
about how data was created and used over time, and versioning history as metadata. [46] Zhang and Huang (2013) suggested incorporating a 5Ws model, which would add additional context to collected big data, which includes: [47]

- What the content of the data is, [47]
- Why the data is occurring in the manner it does [47]
- Where the data originated [47]
- When the data originated [47]
- Who completed the data transference [47]
- Method of data transference [47]

By focusing on questions such as these, one can ensure not only the visualization of advanced data patterns but patterns in the density of the data. [47]

Another concern regarding big data is that methods of storage, privacy concerns, and improper infrastructure can make accommodating big data difficult for businesses. [48] A lack of individuals with the skills to parse large data sets can also add to this problem, though schools have begun remedying this problem with specific programs in data science. [48] Additionally, critics of data-focused visualizations claim that use of things such as summary statistics can hide interesting components of data, such as distortion due to outliers. [16]

C. Information Visualization
Card, et. Al (1999) defines information visualization as a subcategory of (data)
visualization: the use of computers to “generate visual representations of abstract data to
amplify cognition.” [39] Rather than visualizing data in the form of images, the field of
information visualization deals with improving user cognition via the use of visual aids.
A ubiquitous example of this is writing down a multiplication problem, versus solving the
problem mentally: in a study, allowing someone to write a multiplication problem down
on paper brought the time to solve it from 50 seconds to less than 10. [39]

Amplifying cognition is tied into the crystallization of knowledge, a process that
has the following steps: [39]

1. Collecting information on a particular topic, [39]
2. Categorizing information via a representation, or a schema, [39]
3. Adding collected data to a given schema, [39]
4. Analyzing the data as-is, modifying it for efficiency, [39]
5. Re-visiting the schema and adjusting it as needed, [39]
6. Developing final conclusions [39]

In practice, this often consists of developing structures to hold data in and probing
those structures to ensure they reflect aspects of a given data set effectively. Maps,
diagrams, and charts are common methods of conveying information in this way. [39] A
popular example is creating a table of data, and then constraining the data in a table to a
view, or using overview and detail views to drill down into further details about a row.
[39] The development of an ontology for organizing data may also help with knowledge
crystallization, allowing a user to construct a taxonomy reflecting relationships between items in their ontology for organization’s sake. [39]

A few structures oft-used in information visualization are of specific interest. Hierarchical structures can be quite beneficial, although to display them traditionally, printing all leaf and branch nodes, can often result in many lines of text! Shneiderman (1999) suggests the use of tree-maps, which display hierarchies as squares-within-squares to solve this problem. [50] Dynamic queries are also often used in information visualization, which allows a user to use interactive controls to view a rapid and visual display of potential database results – resulting in interactive, on-the-fly filtering. [51] Users tend to enjoy using dynamic queries, as they give them a strong sense of control over their data and are generally perceived as being “fun” when compared to, say, SQL queries. [51] Overview and detail systems are likewise quite popular – such as MIT’s Dataland system: comprised of a wall-sized display and a touch-sensitive visual control, the overview display contained a movable rectangle indicating where in the data the user has currently zoomed in to. [152]

D. Decision Theory

Decision theory deals with making decisions in the face of insufficient data (or access to data) – and creating quick and efficient assessments for collecting the best possible data in certain situations, and discovering best estimates instead of perfect answers. [52] This differs from data theory, which assumes all data that will be needed by the system is available. [52] Decision theory can be a more realistic approach to
making decisions, as often in real-world decision making, one does not have all the data required to make a choice. [53]

Computers are frequently used to assist individuals in making decisions. They are able to aggregate and present data that will impact decisions in a fast and cheap way, and using data collected by a computer is often thought of as being more reliable and less prone to human error. [52] Using computers to support decision-making can refer to the use of computers to access informative displays of information for helping with decision-making, or monitoring human activity and providing assistance with tasks and process control as needed. [53] The latter is seen frequently in fields where a user may have to split their attention across many tasks and may benefit from computerized support, such as in flying aircraft. [54] Heads-up displays are another commonly seen construction of this approach. [55]

One example of a decision-support system would be expert system software. Software designed to be an “expert system” functions like a human expert when dealing with a particular set of problems – they might determine proper courses of treatment for medical purposes, determine failures, or help a user make choices. [52] Business intelligence has also been deemed a type of decision support tool; decision support also is considered to be highly compatible with the use of big data. [57] As a result of this interest, decision support research is expected to improve in quality. [57]

Techniques have been developed in the multi-criteria decision analysis field for assisting with decision-making that, in many cases, have been extracted to software. [58]
These techniques are often organized by the problems they are trying to solve – decision problems include choice problems, sorting problems, ranking problems and description problems. [58] Approaches include full aggregation approaches, where scores are evaluated for each criterion and assimilated into a total score, outranking, where a bad score may not be compensated for by a better score and incomparability is allowed, and goal/aspiration approaches, where a goal is defined for each criterion, and worked towards. [58] Of these, goal programming would probably the most related to this work: goal programming focuses on the optimization of multiple goals by minimizing deviation from an objective. [112] The technique sees frequent use in quality control and management. [112]

E. Computer-Assisted Data Collection

In the 90s, a number of projects investigated the use of computers for directly collecting observational data. [146][147][148] In behavioral research, systems that collected observational data dealt with collecting data such as response frequency, duration, latency, time samples, and so forth – working on both Mac and Windows machines, and collecting data via laptop and hand-held computers. [146] One software title in particular, called “The Observer,” [147] worked on both PC and handheld device, and allowed grouping of events in classes, note-taking, and in an extension of their work, [148] timing their observations alongside recorded video.

Nowadays, mobile computing and smartphones provide the chance to collect varied data from the environment at any time. Using this data, maps of an individual’s general activity can even be collected, as they go about their day. [10] However,
computers can also directly assist with data collection – a practice that has been popular for years. This is typically done via disbursement of an assessment, either via phone or computer, that an individual is asked to complete. [59], [60] Assessments typically have two end goals. Formative assessments aim to help with decision-making, guiding potential interventions and the like - however, summative assessments directly act as a depiction of the current state or general decisions that might be made regarding whether a change has affected someone (or something). [52]

Assessments can be quite useful at data collection, with the ability to collect both qualitative data (such as comments) or quantitative data (in the form of measurements or multiple choice questions). Multiple choice questions are often scaled according to the Likert scale, which states that all that is needed for someone to make a judgment rating is a five-point scale with the following options: strongly disagree/agree, disagree/agree, and neutral. [52] This scale can be adapted to other systems as needed and is scored in an additive manner. [52] Likert-scaled assessments can then be integrated into a scoring system or analyzed merely for frequency of answers. In some cases, such as when assessing the function of people with disabilities, the use of seven-point scales is encouraged. [52]

Successful assessments for data collection require the assessment creator to be able to elicit knowledge about the topic and break it down in an effective manner. One way to do this might be via procedural knowledge about the topic of the assessment from an expert (be it themselves or another individual), and then write in a manner that elicits
declarative knowledge from the data collector. \[114\] When building the assessment, investigative interviewing, combined with understanding the sequence of events it takes for a state to change would be the best course of action. \[114\] Performing a hierarchical task analysis to decompose tasks into sub-tasks would also be appropriate – the focus of breaking a problem into sub-problems, and working through what it would take to solve each of these fits perfectly what observational data collection. \[115\]

Using assessments for data collection always introduces the potential for bias to be introduced into the system. When a question asks a person to rate anything – an observation, or make some quality judgment – humans will bring their own biases, beliefs, known diagnoses and so forth with them in making their judgment. \[52\] Various solutions for this problem, such as pre-scoring based on previously completed work may be an option for dealing with this concern. Scaling might require an adjustment in cases such as these as well, as different types of scaling could improve assessment resistance to bias.

This all said, using a computer to filter/tailor questions and collect data isn’t a new idea. Computer adaptive testing (CAT) sorts items on an assessment via level of difficulty and assesses a user until some error tolerance is exceeded and the system proceeds to the next question (dependent on the prior question’s answer). \[52\] Furthermore, nowadays, lots of software exists to assist users in collecting data from users via assessments. \[61\]–\[64\] If anything, the standardization of an approach that for assessment-driven data collection would be greatly beneficial, especially in the case of a
field such as assistive technology, where hundreds of different assessments exist due to a lack of standardization. [52]

Much of this work will be extending that of Smith (1993), who set out to create a standard evaluation – usable regardless of population, practice or place – for functional assessment. [52] The assessment was designed to be flexible and based on the previously-mentioned field of decision theory. [52] The assessment also utilizes a branching process similar to the CAT, ensuring that therapists did not have to answer (or even see) questions that weren’t relevant to their patients’ needs and their particular practice. [52]

F. Universal Design as Usability Guidelines

This work considers usability a primary goal – and the best way to ensure usability is to focus on implementing a universal design. Burgstahler (2015) outlines seven Universal Design Principles (UDP) designed to help individuals use Universal Design in their own product design. [65] These principles are as follows:

- Products should be usable by people with the widest variety of abilities. [65]
- Designs should incorporate a range of preferences/abilities that users may have. [65]
- Users shouldn’t have to have any prior experience or knowledge to use the design. [65]
- Using the design should be easy, regardless of user ability or background experiences happening around the user. [65]
• The design shouldn’t result in hazards if the user performs any unintended actions. [65]

• The design should require a minimum of physical effort, and not result in fatigue. [65]

• The device should be usable regardless of the users’ body size, posture, or mobility. [65]

These tenants allow for the design of applications, assistive technology, housing, and more with the broadest number of abilities in mind. Taking these ideas into consideration from the start allows for the construction of accessible and available solutions for individuals from the start, costly renovations and redesigns becoming a thing of the past. Any data presentation following these tenants, likewise, will allow the most significant number of people possible to take advantage of the data reflected in the presentation.

A quick note on the idea of availability: accessibility and availability are both essential concepts, but can sometimes be used in very similar ways. To say that something is accessible, in this work, will always mean “available, regardless of a person’s abilities” – built in a manner such that people with disabilities are able to access that material. However, availability is also an important concept – material should be accessible (that is, available) to people regardless of their background, a fundamental tenant of this system design. Software might be accessible for people with disabilities, but could assume a data science background, thus not being available to people who have
a different background and are interested in collecting and presenting data. Software, likewise, may be available (assuming nothing about a person’s background or prior data science knowledge), but may not be accessible for people with disabilities. Universal design encompasses both of these requirements, by ensuring that software is both accessible and available for people, regardless of background, by focusing on simplicity and a design that is usable for everyone.

1) Creating a Universal Design

Prioritizing universal design can have brilliant side effects. For example, a particular feature may end up being beloved by user groups who have a disability and those who do not: a famous example is the creation of VoiceOver, a feature debuted on the iPod Nano that allowed both individuals with limited mobility (as well as those who were able-bodied but were currently involved in an activity) to hear the title of a song without looking at the screen of the iPod. [67] In the world of fashion, eyeglasses have become fashionable to the point of having high-end brands make all sorts of styles and colors of frames, despite their beginnings as assistive technology. [68]

This said, it’s important to remember that the following summaries of accessibility provisions do not merely apply to users with disabilities: not everyone will be able to take advantage of their full range of abilities at all times – and this may not always be due to disability! Some users might not be able to use a keyboard or mouse, or understand instructions due to not speaking the language instructions were written in,
having a cognitive impairment, or just being busy at the time instructions were given. [69] Edwards (1995) writes regarding this: [153]

“A person who has aphasia following a stroke cannot remember the names of objects, but most human beings exhibit mild symptoms of this nature.” [153]

Typically, when designing websites or software, developers break accessibility concerns up into the following categories and offer the following suggestions. Note that these suggestions are in no way exhaustive.

**Vision Impairments/Blindness.** Vision impairments may range from partial vision or impaired visual acuity [69] to colorblindness. [66] Blind users tend to use a lot of different assistive technology devices, including screen readers (which read text aloud from a screen), speech recognition (for completing data input) and refreshable braille displays. [66] To ensure data is accessible for these users, text is implemented that is

<table>
<thead>
<tr>
<th>Accessibility Provision</th>
<th>Target Population Helped</th>
<th>Secondary Population Helped</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operating a Smartphone via Speech</strong></td>
<td>Individuals with mobility and vision impairments.</td>
<td>People holding items and unable to use a touchscreen temporarily, distracted individuals.</td>
</tr>
<tr>
<td><strong>Closed Captions on a Video</strong></td>
<td>Individuals with hearing loss, deaf individuals.</td>
<td>People watching the video in a very quiet environment (i.e. cannot turn on audio.) People watching the video in a loud environment.</td>
</tr>
<tr>
<td><strong>Picture Context Explained in Alt-Text [105]</strong></td>
<td>Individuals with vision impairments.</td>
<td>People who may be unfamiliar with the source material/background of the picture.</td>
</tr>
<tr>
<td><strong>Consistent User Interfaces Throughout Website [66]</strong></td>
<td>Individuals with cognitive impairments.</td>
<td>People who are unfamiliar with computers/uncertain if they're still on the same website.</td>
</tr>
</tbody>
</table>
equivalent to graphic features, clearly label UX elements, clearly title links, and describe all navigational structures well. [66] For media, audio descriptions or transcription may be necessary.

**Hearing Impairments/Deafness.** Users with hearing loss or who are deaf often rely on closed captions to access audio-based media. Closed captions only appear when support for them is offered by the software or hardware screening the video, and they are provided separately (i.e., in a file). [66] Youtube has support for this style of captions, for example. Open captions are coupled directly with the video, and thus may lose quality as the video is compressed. [66] Users with hearing impairments can also take advantage of sign language access to spoken material, but automatic translation and the development of 3D avatars poses a problem here. [69] Any audio-based alerts should likewise offer an audio-free option for users.

**Physical/Mobility Impairments.** Mobility impairments range from slight to severe, temporary to permanent – their causes may be anything from old age to injury. [66] Users with physical impairments may suffer from paralysis, ataxia (a loss of gross coordination), or tremors. [69] These impairments are often exacerbated by fatigue and pain. [69] Assistive technology required by those with mobility impairments ranges from utilizing alternative mice and keyboards to using head mice (which translate head motions into mouse movements) and switches. [66] Taking care to have material usable without a mouse, or compatible with a switch, helps ensure a design that considers this population.
Cognitive Impairments. Cognitive function refers to executive function, memory, attention, visual and special perception, emotions, and language. [69] The operation of these functions can be altered due to injuries (temporary and permanent), aging, mental illness, and so on. [69] Designing with consideration for cognitive impairments is a bit different than the other examples as cognitive impairments are typically similar to usability problems that would affect all users – just amplified. [69] Consistent designs and few distractions are the best for individuals with cognitive impairments. [66]

There’s also always a potential case of overlap. An example of a population that may have needs that overlap between these categories is that of aging users. Definitions of “aging users” vary, ranging from those over 40 to those over 58. [69] However, in general, gradual sensory decline starts between the ages of 40 and 55. [69] Aging users typically have three problems when using technology: barriers preventing them from being fully confident with technology (such as education or awareness), any physical or cognitive impairments they may have, and technology they are attempting to use having poor UX. [66] The latter two are the focus of this work. As one grows older, their vision can worsen, response times for complex motor tasks can increase, their hearing can become impaired, and they can encounter a decline in procedural and episodic memory. [69] To further complicate things, older adult users do not all age in the same way. Many aging users don’t appear or identify as a person with disabilities – despite any different functionality, needs or wants. [66]
G. Differences Between Existing Work and This Solution

The most crucial difference between existing work and the proposed solution is that the proposed solution does not attempt to make a decision on behalf of the user. As mentioned, most solutions do try this approach – visualization is typically centered around displaying data that supports a hypothesis, rather than merely showing what data was collected. Decision support often refers to the act of making a decision for the user, or at least strongly suggesting a decision – at this point in time, the proposed solution organizes data instead and leaves the decision regarding what to do with it up to the user. Rather than visualizing data after the fact, the proposed solution allows for the visualization of complex data during collection. To do this, it models the data collection (and thus the arrangement of data) after an assessment.

The proposed solution also focuses on complex data rather than big data. Crowdsourcing data collection might result in many assessments being completed (and thus, hundreds of thousands of data records). However, the main focus of this work is on data collection carried out by a person at this point in time – and the organization and structure of this. To lighten the burden on the data collector, the proposed solution integrates complex data capture into the stated organization method. While streams of data could theoretically be integrated, at this point in time the proposed system focuses on the reduction of human error that comes with using mobile technology/sensor data capture.

Approaches close to this approach come from the information visualization field – specifically work surrounding hierarchical data arrangements [50] and overview/detail
views. The proposed approach utilizes a hierarchical view, but in the form of a collapsing outline. The focus is on a notion of dynamic queries, in allowing a user to drill down for details as needed – but the extensive filtering explored in Shneiderman’s work on the subject [51] is not retained.

Finally, aspects of universal design are found throughout the proposed solution, but nowhere as much as in the integration of media. This sets this work apart as well – other solutions for handling media metadata in the context of data collection do not integrate accessibility data for this purpose. While this does put more work on the data collection team, it’s worth noting that various legislation requires accessibility to be a crucial part of software. Notably, Section 508 requires that all federal employees and members of the public accessing federal resources are able to access the same information and data – or at least data that is comparable. [69] This said, integrating accessibility information for data collected for federal establishments is a practice that should be taken up anyway!
III. STRUCTURING DATA COLLECTION AND PRESENTATION WITH HIERARCHIES

In 1985, a grant was awarded to the University of Wisconsin-Milwaukee for the development of a standard assessment for occupational therapists - a project that brought to light some of the most significant problems in data collection and examination by practitioners. [52] Hours were spent in the collection and presentation stage, and over half of the data sets generated resulted in mathematical errors. [52] Some method of organization was sorely needed.

To amend these problems, Smith (1993) developed a functional assessment for people with disabilities, designed to be flexible and robust. [52] The software, OTFACT, took inspiration from decision theory in its design, rather than relying on the presence of data (something that couldn't be guaranteed). [52] In the years following its debut, the assessment would be used to evaluate outcomes of assistive technology in the literature [71][72] and to develop treatment protocols for injury [52] and as a basis for additional assessments for evaluating technology [73] and environments. [74]

While one might assume that such problems would fade over the years, in many cases analysts still encounter problems such as those OTFACT sought to solve - namely, understanding where to begin data collection and developing a standardized protocol for the act. Consider for a moment all of the factors that are considered when determining whether some assistive technology is a ‘best fit’ for someone. While, as an example, there are plenty of glucose monitors – ones with bright screens, ones that speak measurements aloud – each one is best for a different sort of person. [84] In order to
understand what type of person would be the “best fit” for each type of glucose meter, one needs to gather data on whether the glucose meter fits the requirements of the person who is interested in it. [84] This means collecting data on how easy it is for a person to use it who has a sight impairment, hearing impairment, mobility impairment, and so forth. [84] This is a large amount of complex data to gather!

In this chapter, the process of gathering observational data as it currently stands is discussed. OTFACT is investigated as a functional assessment for people with disabilities which solved the challenge of observational data collection in a novel way. [52] The prospect of standardizing the OTFACT approach for other types of data collection is elaborated upon, and evaluated by working through examples of creating TTSS-styled assessments.

A. Challenges in Observational Data Collection

A number of problems exist when collecting observational data today: for one, developing a protocol for data collection can be difficult, as even gathering data for the same end goal can mean different steps on a case-by-case basis. Context needs to be gathered for measurements as well, in order to understand why data was collected. Finally, protocols created should be reusable – but this can often cause protocols to be too broad, and contain steps that the average data collector will merely skip.

To illustrate the current challenges that one may encounter during the process of data collection, consider the following case studies.
1) Weather Severity

First, consider the problem of assessing the severity of weather on a particular day. Understanding both what constitutes severe weather, and what amount of those factors affects the severity of the weather, are required to solve this problem. Weather metrics and interpretations for what constitutes severe weather could be collected, data related to the current weather forecast arranged in a table, and indicators for those metrics that fell within the range for a severe weather event could be marked. However, completing this assessment would require the measurement of each metric each time one wished to understand the status of the current weather forecast. There wouldn't be a mechanism in place for capturing uncertainty. Furthermore, a spreadsheet approach may be confusing for those interested in using the method in the future.

2) Software Accessibility for Section 508

The Rehabilitation Act of 1973 addresses computer accessibility for federal agencies – in particular, Sections 501 and 504 require federal agencies, and entities receiving federal funds, to offer reasonable accommodations for employees, and people using their technology. [122] Section 508 is more specific, requiring that all federal employees and members of the public have access to the same data – a goal that requires implementation of universal design and accessible technology. [122] In an effort to make the process of ensuring compliance with Section 508, companies often use VPATs, Voluntary Product Accessibility Templates, which assess how strictly products comply with Section 508 regulations. [122] A VPAT should be completed for software that was
developed for federal use, or software a federal user is considering purchasing for an organization. VPATs require in-depth knowledge of operating systems, interactions between a software title and a given operating system, and interactions between multiple software titles. Researching the different cases, and filtering the guidelines for a particular OS and software situation, could take hours.

With these cases in mind, one can explore how OTFACT - and more specifically, the trichotomous tailored sub-branching scoring technique the software is centered on - can assist with the creation of data collection protocols, and how the presentation built into the system can benefit data collectors.

B. The Design and Development of OTFACT

OTFACT, developed by Smith (1993), was a software system that set out to create a standard evaluation – usable despite population, practice or place – for functional assessment. [52] The assessment was designed to guide a user through questions and request detail when appropriate, allow a user to skip branches of the assessment not directly relevant to their use case, and be flexible, ensuring that therapists did not have to answer (or even see) questions that weren’t relevant to their patients’ needs and their particular practice. [52] To obtain these goals, OTFACT had two novel aspects to its design: the trichotomous, tailored, sub-branching scoring (TTSS) system, and the reliance on a computer for presenting this format of assessment. [32] Also prioritized was a simplistic user interface. Simplicity in design was key for creating an assessment that
occupational therapists would enjoy using over their preferred method of data collection: hand-logging, which was (and often continues to be) the default method of data collection for occupational therapists. [70]

1) User Interface Design

Figure 3-1 displays the assessment view of OTFACT - comprised of a toolbar of reporting and other options on the far left side of the screen, an overview of the assessment overall to the right of this, and details and response options for their selected assessment item further to the right of this. The screen is designed to allow a user to answer the fundamental questions of navigation without issue: where they are, where
they have been in the assessment thus far, and where they have yet to go. [76] The toolbar of options on the right side is mainly used to access reporting features and other parts of the application but offers a few advanced features for progressing through the assessment. An example is the 'Next NE' option, which refers to traversing to the next unexamined item in the assessment.

The overview of the assessment is constructed as a collapsing outline, in order to aggregate collections of items that are not relevant at a particular moment. The outline starts numbering with Roman numerals, proceeds to letters, and finally digits. The outline's sections indent to represent hierarchies - for example, in the screenshot displayed, 'Cleanliness, Hyg. & Appear.' is a sub-section of 'PERSONAL CARE ACTIVITIES,' itself a subsection of 'ACTIVITIES OF PERFORMANCE.' Hierarchical views are appropriate for situations such as this, where parent/child relationships are presented between items. [77] Only one section of each level can be open at a time, to minimize distraction. When an item is selected, the color for that item inverts, and connects to the detail view on the right - itself inverted in color scheme compared to the rest of the application - to draw a visible "connection" for the user, and avoid distraction.

The detail view provides detailed information and scoring options for a selected item. In Figure 3-1, for 'Medication Routine' additional notes about what this refers to are given: 'Obtains medication; accesses container...' and so forth. [52] A note regarding the response shows that this item is worth 14 points, but has been answered with a 0 (indicating a total deficit in the person this assessment is scoring, an inability to perform
The three standard options of a trichotomous assessment are always available – the "yes" (here, "No Deficit"), "no" (here, "Total Deficit") and "maybe" (here, "Partial Deficit") options.

One might question why this item is worth 14 points. This is a nod to the fact that this item can branch, a fact shown visibly in the outline view via the existence of a \(+\) icon adjacent to the item text. There are seven child items for this item, totaling in a potential score of 14 for this particular item.

2) Scoring

Part of the TTSS structure’s name comes from the scoring of the system: the scale is trichotomous, with two items acting as quantitative measures (yes/no) and one as a screening question (maybe) and allowing an assessment to branch, and acquire additional details from additional items. [52] These definitions in mind, the following transitions from one question to another are available when scoring a question.

An item in OTFACT can be answered with 'No Deficit,' 'Partial Deficit,' or 'Total Deficit.' Additionally, an item can be a leaf or a branching item. [52] For certainties of no deficit or total deficit, the question is awarded 0 or 2 points – if the question is a branch question, all of that branch’s leaves are scored accordingly, and skipped as the assessment moves on to the next question at the same level as the branch. [52] For leaf questions, the assessment continues to the next question as the branch, as a leaf question does not have leaves of its own. [52] If a leaf question is scored as a ‘maybe,’ the assessment also continues – but scoring a branch question with a one causes the
assessment to branch, and additional detail to be requested. [52] Leaf questions offer the numerical scores that are ultimately tallied: while the questions are asked in a top-down manner, question scores are tallied bottom-up. [52] Figure 3-2 shows an example of this scoring as a flowchart.

This scoring approach has many benefits. For one, uncertainty is clarified via branching: if a user is not sure, they can mark a question as such, and obtain certainty via the answering of additional questions. [52] A three-point scale minimizes potential cognitive load on a user, which would otherwise increase as the number of potential
answers increases. [52] Finally, assessment sensitivity is maintained in the long run - section scales may have a wide range, while single questions never go beyond a trichotomous scale. [52]

3) Filtering

There are two ways an assessment is dynamically filtered. The first is the scoring mentioned in the preceding section: done correctly, a user who is confident of an item can skip all of the children for that item, reducing the number of questions they ultimately need to answer. [52] A question can also be marked non-applicable to remove it from scoring entirely, filtering an assessment and the resultant scores to only those that are needed by an individual. [52] By doing this, the user is assisted in focusing on the task at hand, and not be distracted by questions that may have little to do with their goals.

Filtering is especially vital in the case of TTSS assessments because the filtering of a question set is expected. Question sets are exhaustive, numbering in the thousands, separated into a number of categories that may or may not be pertinent to a given situation. It’s expected that some questions will likely be skipped due to specific answers earlier in an assessment, or perhaps specific questions will not be applicable to a user’s particular situation. This approach allows one to construct a comprehensive data collection protocol for a topic, and then whittle down the assessment depending on the needs of the assessment-taker. Doing this makes reusable data collection protocols. The only downside is the time (and potentially, the interdisciplinary team) required to ensure that a question set is genuinely comprehensive.
C. TTSS Assessments: Past and Present

TTSS techniques were shown to be resistant to rater bias when compared with traditional functional assessments at low (<40%) levels of rater bias. OTFACT proved itself to be quite popular [32] and used in many studies as a method of measuring outcomes for assistive technology. [71][72] As a result, several assessments based on the principles of TTSS were also developed for various goals. Table 3-1 illustrates a number of these assessments and their question counts. RATE-IT and MED-AUDIT are great examples of these related assessments, and their typical workflows and goals.

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcomes Studies</th>
<th>Number of Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTFACT [52]</td>
<td>Functional assessment for assistive technology outcomes.</td>
<td>965</td>
</tr>
<tr>
<td>School-Fact [117]</td>
<td>Assessment to determine &quot;school function skills&quot; for children and young adults.</td>
<td>1015</td>
</tr>
<tr>
<td>MED-AUDIT [73]</td>
<td>Evaluation of medical equipment accessibility.</td>
<td>870</td>
</tr>
<tr>
<td>HAFT [118]</td>
<td>Comprehensive home assessment for environmental, architectural and functional concerns.</td>
<td>257</td>
</tr>
<tr>
<td>RATE-IT [78]</td>
<td>Evaluation of restaurant accessibility.</td>
<td>964</td>
</tr>
<tr>
<td>AAC-FACT [120]</td>
<td>Functional assessment of outcomes related to augmentative and alternative communication.</td>
<td>797</td>
</tr>
<tr>
<td>Trans-Fact [121]</td>
<td>Evaluation of public transit accessibility.</td>
<td>284</td>
</tr>
</tbody>
</table>

Table 3-1: A sample of other projects using a TTSS-styled assessment for data collection, and their complexity.
1) **RATE-IT**

RATE-IT was designed as an assessment of the accessibility of restaurants - not all restaurants follow appropriate accessibility provisions, and thus having a method of determining a state of accessibility of a restaurant would be beneficial to both those with disabilities and those interested in maintaining compliance with the ADA. [78]

Assessments for ensuring accessibility of built environments exist, but often expertise is required to complete them successfully - RATE-IT opted to be a practical assessment. [78] Indeed, participant groups reported that using RATE-IT was significantly more

![Screenshot of RATE-IT](image.png)

*Figure 3-3: A screenshot of the assessment view from RATE-IT.* [56]
user-friendly than the use of the standard ADAAG checklist while also finding that they spent far less time completing the assessment. [78]

2) MED-AUDIT

The MED-AUDIT project aimed to rate the accessibility of medical instruments in the field - including instruments found in hospitals, nursing homes, and clinics. [73] A TTSS-styled assessment was utilized in procuring outcomes for this end, due to the efficiency and flexibility inherent in a TTSS system. [73] Two methods of scoring accompanied the taxonomy: one for adapting the xFACT system to be used as an 'Expert User Taxonomy' with true scores revealed after completion of the assessment, and another 'Black Box System Taxonomy' scoring approach, which weighed items in correlation to medical device features. [73]

D. xFACT for TTSS Assessment Development

Over time, work was completed in generalizing the application into a taxonomy editor and presenter, called xFACT. [80, 81, 82] xFACT was designed as a cross-platform application suite with the following features.

- Ability to score TTSS-styled taxonomies
- Ability to indicate question relevance for taxonomies
- Ability to score question validation for taxonomies
- Ability to create, edit or import new taxonomies from Excel spreadsheets
xFACT is released as an application suite, composed of several main elements: the primary application and the editor [80] and two accessory applications. [81] [82]

1) Main Application

The main xFACT application [80] allows a user to score TTSS-styled assessments. xFACT, available on both Windows and Mac platforms, works as a desktop application in its current state. [79] Surveys are completed using a mouse and keyboard, in a way that mirrors OTFACT and survey completion in this software.

Completed assessments can be saved or loaded. For example, if a user wanted to complete a RATE-IT taxonomy, rated restaurants could be saved as individual files that are loaded and saved at their discretion. Taxonomies are offered as separate instances of the xFACT application – xFACT is designed to offer one taxonomy at any given time, with the exception of the editor (which may offer multiple for scoring, as it allows a user to test out taxonomies for use after editing). In any case, the user's interface only changes in color depending on the taxonomy selection - interface elements do not change or move between assessments. [79] A consistent structure to the application pages is maintained.

The TTSS format of OTFACT is maintained, with trichotomous answer sets being displayed for each item in an assessment. Scores are tallied in the TTSS manner discussed earlier in this chapter. Thus, higher scores can indicate more compliance with guidelines, while lower scores may indicate noncompliance or uncertainty. Depending on the assessment, xFACT can also act solely as a data collection tool (hiding the scores of previously-answered questions from the outline), but also as a data presentation tool.
Scores can be displayed adjacent to the item text on the left side of the screen in the outline, allowing for items to be "drilled down" into to assess which items are the sources of lower scores. In this way, the collection of data is guided via a hierarchical structure and presentation of the data is offered in a summarized way. Problems in this observational data collection (such as noncompliance with guidelines) can be caught at a glance, and details as to why these problems are occurring can be confirmed by drilling into given data sets.

Questions can be answered in any order. For professional or experienced users, this can be beneficial - a user may be able to make snap judgments on the state of something, and in an order not guided by the assessment. [79] This is supported. Likewise, a new user may wish to use xFACT more like a checklist, going through each item dutifully. This is also supported.

2) Editor

xFACT allows a user to create TTSS assessments not only in Excel (and add to the software via import) but in an application included in the software suite itself. Using the xFACT Taxonomy Editor (XTE), a user can edit and add xFACT taxonomies that keep to the spirit of the TTSS-styled assessment automatically.

The XTE allows a user to add taxonomies and response sets as needed. Taxonomy items are added in the hierarchical manner that OTFACT arranged them in - large sections, with detailed items within broader sets so that a user can drill down into details. The user interface allows item details and response sets to be added to individual items as
needed, in addition to prompts - eye-catching alerts, to warn a user of something such as a response set change.

Items are coded with response sets. Response sets reflect the TTSS methodology, containing a trichotomous answer set at the very least, and including various other filtering mechanisms if needed. However, numerous types of response sets can be used in this format (i.e., Yes/No/Maybe, Exists/Partially Exists/Does Not Exist, No Deficit/Partial Deficit/Full Deficit, etc.) The XTE also allows assessments to be imported from Excel, such that new taxonomies can be added even by those who are more

Figure 3-4: A Screenshot of the taxonomy selection screen from the xFACT taxonomy editor
comfortable using software such as Excel or a text editor for construction of their assessments, for maximum flexibility. [79]

Finally, the editor allows a user to test out their newly-created taxonomy, via the same scoring view found in the main application. Users can toggle additional features in the editor for their taxonomy, such as representing scores as fractions, or percentages. An example of the specific screens involved in using the xFACT Taxonomy Editor is included in Appendix One.

3) Accessory Applications

Finally, many accessory applications are available for xFACT compatible assessments. One is the xFACT Relevator [81], which allows one to rate the relevance of questions in a taxonomy to the task at hand. This is fantastic for the creation of new taxonomies in interdisciplinary work, where the approval of a trained expert may benefit the construction of a taxonomy. The other is the xFACT Validator [82], which seeks to rate how appropriate the language and relevance of an item is for an xFACT taxonomy. [83]

E. Case Studies for Using xFACT

In this section, the appropriate use of xFACT in creating TTSS-styled assessments will be demonstrated via the creation of two assessments. In the first, the collection of data from the environment is explored, to assess the severity of current weather, while noting some weaknesses on the part of xFACT for completing this task. This assessment
serves as an example of an observational assessment, where weather conditions are examined for severity. In the second, software accessibility is assessed, as per the guidelines outlined in the Voluntary Product Accessibility Template (VPAT) [122]. This assessment serves as an example of a special assessment, where features of software are examined for accessibility features as per the VPAT.

1) Severity of Weather

The first step to creating this taxonomy is to understand the problem thoroughly. What does severe weather mean? One can start by assessing what comprises a severe weather event. Considering common severe weather events, one might come up with the following list: precipitation, extreme temperature, wind, and risk of dangerous conditions. (Note that this list doesn't have to be exhaustive for the sake of this demonstration.) These work well as categories for this TTSS assessment.

These categories will need to be broken down to develop this assessment. The suggested method for breaking down these categories is via a task analysis. [108][115] Understanding the behaviors that must occur for a task to be completed, in addition to a list of things that are necessary for the task to be completed, are great ways to break down a task. [108][115] This is where a hierarchical task analysis of the procedure at hand comes into play, breaking down precipitation, temperature, wind, etc. into qualifiers for each item that confirms its hand in creating severe weather. [115] Consider the ‘Extreme Temperature’ item – certainly, this plays into severe weather, as too high or low of a temperature can result in a weather warning for the day. But what exactly should
constitute an extreme temperature? The temperature outdoors is a good start, but what about the temperature indoors? What about indoors, without some sort of temperature management? Thus, the temperature can be broken up by location: outside in sun's direct path, outside in the shade, indoors with temperature adjustment, and indoors without temperature adjustment.

I. EXTREME TEMPERATURE

A. Outside

1) Outside, In Sun’s Path

2) Outside, In Shade

B. Indoors

1) Indoors, with Personal Temperature Adjustment

2) Indoors, without Personal Temperature Adjustment

The task was broken up by assessing subtasks within the measurement of temperature severity task and categorizing them.

Next, one might consider what constitutes poor temperature in these locations? To do this, one can consider that when extreme temperature leads to severe weather, the National Weather Service requests that people stay indoors, with special emphasis on children and older adults. [109][110]

So now, the assessment is amended with these considerations:

I. EXTREME TEMPERATURE
A. Outside

1) Outside, In Sun’s Path
   i. Poor personal comfort
   ii. Danger to at-risk populations
   iii. Danger to all populations

2) Outside, In Shade
   i. Poor personal comfort
   ii. Danger to at-risk populations
   iii. Danger to all populations

B. Indoors

1) Indoors, with Personal Temperature Adjustment
   i. Poor personal comfort
   ii. Danger to at-risk populations
   iii. Danger to all populations

2) Indoors, without Personal Temperature Adjustment
   i. Poor personal comfort
   ii. Danger to at-risk populations
   iii. Danger to all populations

This is about as granular as this system could get in terms of describing the temperature conditions. However, in describing the temperature conditions in this way, a great amount of information as to the state of the temperature is offered at this point in time. This information consists of:
• If it matters if one goes indoors. (If the temperature continues to be difficult despite any sort of personal temperature adjustment, such as air conditioning, this is a problem.)

• If the user is simply uncomfortable, or if sensitive populations are at risk.

Thus, this assessment is an excellent example of an observational assessment: one works to observe the environment and categorize their observations using a TTSS assessment appropriately. A more comprehensive breakdown for this assessment is available as Appendix Two.

Figure 3-5: Example of a weather severity taxonomy in xFACT via the taxonomy editor.
2) Accessibility of Software

Another type of TTSS assessment is one that allows for preferential analysis. In this sort of assessment, two TTSS assessments are filled out with desired scores, and one with scores as they relate to something. In doing this, the user can compare what they want to what they have, and see where discrepancies lie. Note that xFACT is not a real decision support software title, and so no decision will be made for the user.

A taxonomy that might lend itself well to this sort of completion is one inspired by the Software Applications and Operating Systems VPAT (1194.21). [111] The questions here ask about categories of accessibility and provide detail for the user to consider when assessing each area – thus, they are very easily broken down into a TTSS assessment. [111] As an example, consider the second question in the sequence of questions related to software accessibility:

"(b) Applications shall not disrupt or disable activated features of other products that are identified as accessibility features, where those features are developed and documented according to industry standards. Applications also shall not disrupt or disable activated features of any operating system that are identified as accessibility features where the application programming interface for those accessibility features has been documented by the manufacturer of the operating system and is available to the product developer." [111]

Breaking this down, it’s evident that this question deals with the disruption or disabling of accessible products. This seems like an excellent top level item.

I. INTERFERENCE WITH EXISTING ACCESSIBILITY FEATURES
This can be broken down into features of other software and features of the operating system. Problematic software would disrupt or disable existing accessibility features. Thus, for this item (and after taking into consideration other items in the VPAT that line up with this heading), the TTSS assessment would look like this:

I. INTERFERENCE WITH EXISTING ACCESSIBILITY FEATURES

A. Interference with other software features.

1) Disrupts existing software features.
   i. Overrides User Selected Contrast
   ii. Overrides User Selected Colors

2) Disables existing software features.

B. Interference with operating system features.

1) Disrupts existing software features.
   i. Overrides User Selected Contrast
   ii. Overrides User Selected Colors

2) Disables existing software features.

To create an even more exhaustive assessment, one could even go as far as breaking down the individual operating system features that one should check for interference with. This would make the xFACT assessment for VPAT especially useful for non-computer savvy folks! In this way, the creation of a TTSS-styled assessment that thoroughly replicates the requirements of this portion of the VPAT has taken place. A more extended example of this assessment is available as Appendix Three.
F. Discussion

The most significant benefit of xFACT is how it helps users deal with very complex data tasks. A user starts by building an exhaustive set of questions, breaks those questions up into categories and sub-categories, and only deals with categories directly applicable to their needs. Due to its hierarchical organization, summaries of collected data are evident as the assessment is completed - potentially allowing a user to pick out problems with scoring, in addition to problems with scoring. This structure of user interface, dutifully following the Visual Information Seeking Mantra, [151] becomes a visualization in and of itself. Scores are shown alongside the questions being asked, such
that a user can view a summary of scores and, if a question is raised, drill down into the score set for a section to figure out the reason behind particular scores.

This approach helps users complete an assessment quickly, without distraction or delay resulting from the existence of questions that aren't relevant to their needs. [79] Filtering helps a user only concentrate on those questions relevant to their immediate tasks. Navigation is designed such that a user always has a sense of where they are in their progression throughout the survey. The outline view offers a clear understanding of what questions potentially remain, while a progress bar indicates a clearer understanding of where the user is in their current travels through the assessment. [79] Furthermore, a progress bar helps a user understand where they are in the assessment completion process even after marking sections as non-applicable.

The software also creates assessments that adapt to the needs of the user - both in the act of scoring and creating assessments. Surveys can be completed in any order, stopped, and restarted at a different time, making for a very flexible assessment structure. [79] New response sets can be constructed as needed for the best user experience. Utilization of both location probes (via hierarchical displays) and viewpoint controls (via the overview/detail design) help a user understand the context of a particular question, while also understanding the structure of the assessment they are completing. [39] Furthermore, utilizing the outline to construct a taxonomy can be useful for organizing knowledge regarding the act of data collection, and produce a reproducible protocol that integrates the display of collected responses thus far.
While reporting features are not yet complete, investigations have begun into the best way to produce comparisons using TTSS-styled assessments and how one might generate reports based on this. [84] OTFACT included reporting, including comparative reporting between assessments. Experimental work has been carried out by students to investigate this as a potential use. MediRank is a proof of concept for comparing two MED-AUDIT assessments, allowing drill-downs into both assessments simultaneously. [84] Two medical devices could be compared, or a medical device and a user-completed MED-AUDIT assessment acting as a users’ preferences. [84] In doing this, one can compare results to help them better make decisions regarding the use of one medical device over another. [84]
Another potential extension of xFACT for decision making might make use of
decision trees, wherein patterns of user choice at nodes are presented and used for
prediction as to what a user's conclusion will be. [77] In doing this, one can inject some
predictive capability into this xFACT system. TTSS-styled assessments do appear
similar to a decision tree, as such assessments are designed to have specific paths if
specific questions are asked, but in the case of the decision tree, questions are not
restricted to answer sets. Decision trees can be quite complex in their structure. [85]

G. Summary

In conclusion, xFACT stands as a flexible and efficient way to collect data and
present it as it's being collected. Flexibility is offered in many ways: allowing the import
of assessments from Excel or creation of them in-app, allowing the user to dictate how an
assessment is completed, and allowing tailoring to a user's use case. Tailoring also adds
to the efficiency of the workflow, wherein only questions that are relevant to a particular
user are even ever seen. Scoring occurs as a user continues through an assessment, and is
displayed alongside items for immediate feedback. The application is designed to
minimize distraction, structure data, allow mouse and keyboard input, and provide a
consistent user interface that a user will grow accustomed with as they use the application
to work towards stated goals of universal design. Reporting has not yet been
implemented but is actively being discussed.
IV. AUGMENTING HIERARCHY-BASED DATA COLLECTION AND PRESENTATION WITH QUANTITATIVE INFORMATION

A potential problem with the assessment outlined in the previous chapter can happen when the user is uncertain about something akin to quantitative measurement. Here, a drill-down event could be time-consuming, and border on increasing the complexity of an assessment. Furthermore, adding the measurement data to a data collection strategy – or multiple measures, for ensuring accuracy – increases the amount of data being stored and analyzed by a system. As illustrated in Figure 4-1, even taking 1-2 measurements for verifying answers for three or so questions can increase data points from 3 to 8!

For a more specific example of the problems, measurement integration can cause, consider the detailed items that may be used for ramp measurements for accessibility measurement. Numbers below a particular range and above a particular range might be

![Figure 4-1: A diagram illustrating the data complexity involved when taking 1-2 measurements for a series of questions in an assessment.](image-url)
considered inaccessible. This could lead to one of two potential outcomes for the person initiating data collection with a TTSS assessment:

- A user encounters multiple, tedious, ‘detail’ branching questions as the truth of the measurement is elicited, or
- A user breaks the ‘flow’ of the assessment to measure the ramp in order to answer the questions of the assessment.

While the latter option will certainly take less time, it still stands as a problem with this solution. Luckily, modern smartphones allow for the capture of information about the environment on the fly, using their onboard sensors. What if the capture of quantitative and other complex data could be integrated into the assessment automatically? This would allow a user to capture measurements from the environment without moving away from the data capture order outlined in the assessment. Furthermore, one could even use the data capture to answer trichotomous questions included in the assessment automatically, based on the results of the measurements!

In this chapter, the use of automatic quantitative data gathering to enhance the efficiency of the TTSS-styled assessment is discussed, and how this approach can benefit an assessment for gathering built environment data. The improved efficiency brought about by integrating automatic data collection is stepped through as a method of improving the user experience for the data collector, as well as the consumer who may be interested in the data behind the conclusions.

A. Data Collection using Mobile Device Sensors
Mobile computing has made the process of gathering data much easier than in years past. Smartphones, for example, allow for the collection of the following types of data: [86]

- Sensor-based data, (gyroscope, accelerometer, and compass data) [86]
- Proximity and light sensor data (sometimes integrated with camera) [86]
- Audio data (microphones, often multiple) [86]
- Wireless radios (Bluetooth, WiFi, GPS radios) [86]

While in some cases, data can be analyzed and processed on the same phone as it is taken, typically data is transmitted to a server for computation and analysis. [87] Luckily, this does not have to mean a slowdown in the overall process: the analysis of such data can seem near-instantaneous. Consider as an example Google's Translate application, which analyzes video of text, translates it to a user's desired language and

Figure 4-2: An example of the capabilities of Google Translate, an app for the iPhone. [88]
alters the video to overwrite the existing text with the translation. [88] Figure 4-1 displays a screenshot of the text and the text after processing. The process of translation and video replacement in total for the title of the book in this example took less than a minute, even when selecting the "quick" download of Spanish text information. [88] Note that the subtitle remains untranslated, most likely due to contrast problems.

However, data collection via sensors and wireless devices goes beyond the use of smartphones. Microcontrollers, such as the Arduino [89] and Particle [90] boards, can be configured as wireless sensors, collecting a steady stream of data and feeding it to a gateway, which also processes data collected. Such a system is a practical solution for data capture in industrial environments, or when measurements are frequently changing.

B. Integration of an Assessment with Quantitative Data Collection

For the integration of quantitative data collection to mesh well with current goals for this research, the following considerations are made.

• Even with the integration of data, this software will still not make an obvious decision for the user. At most, the software will show the data and offer an interpretation of the data, which allows the user to retain agency and control over final decisions.

• The process of obtaining the data should not distract the user too much from their assessment completion.

• When available, use things the user may already have. For example, using a microphone to detect sound levels of a room is a good choice. In the worst
case scenario, use a tool that can connect wirelessly, avoiding the case of user error infiltrating the measurements.

- The stated commitment to a usable interface should not falter, and any additions (such as this quantitative data) should improve usability if possible.

By sticking with these four goals, one can ensure that this application enhances a user’s experience. Keeping the collection of data as part of the flow of the tool and prioritizing usability in procedure helps TTSS assessments feel efficient. Using sensors the user may already have for data collection, such as a user's smartphone, increases the chance that a user will be able to take advantage of built tools. Additionally, sticking with tools a user may already have can save a user money.

C. The Access Ratings Project: Building Apps for Built Environment Accessibility

Like the xFACT Software Development Suite, the Access Ratings project is made up of multiple bits of software. The main application, AccessTools, is a TTSS-styled assessment for appraising the accessibility of built environments. The program runs on an iPad tablet as an app. The accessory applications, known collectively as the mini-tools, take measurements which can be determined using onboard sensors of an iPod Touch, iPhone, or wireless devices. The mini-tools can also be used on an iPad, although the graphic interface is not designed for it.

Built environment access was a very appropriate problem for a TTSS-styled assessment, and a great example of an assessment to fold quantitative data collection into. Assessing the built environment for accessibility involves a healthy amount of
quantitative data collection. Traditionally, the ADA guidelines [75] would be consulted in order to ensure that data is within accessible ranges. Unfortunately, this can be confusing, as there are quite a few guidelines to sort through. Checklists and toolkits also exist to facilitate the process, but still require a user to make an assortment of measurements to assess accessibility - making the process a time-consuming (and potentially expensive!) one. [91] Thus, the idea of using smartphones to obtain measurements via sensors was born.

1) Creating the Mini-Tools

The goal of the mini-tools was to provide a proof of concept that one can use the hardware sensors of smartphones to augment TTSS data collection and ensure that accurate answers are marked. Initially, there were four mini-tools: AccessSlope, [91] AccessSound, [95] AccessLight, [94] and AccessFont. [93] A fifth, AccessRuler, does not rely on smartphone-based sensors, but instead on a laser ruler with Bluetooth capability in order to ensure accurate measurements. [92] At the time of writing, technical difficulties had put the AccessLight and AccessFont applications on hold indefinitely, and they are not included in the current version of AccessTools. As a result, computation methods will only be discussed for AccessSound and AccessSlope. Prior work did dive into the computation of AccessFont [93] and AccessLight [93], but due to
changes in the data being presented by the iOS API, the desired accuracy was not possible with a sensor-based approach in the case of these tools.

**AccessSlope**: AccessSlope aims to carry out the measurement of cross-slopes and slopes of ramps and paths, as is required by the ADA for accessibility compliance. Tanviruzzaman (2012) showed that an excellent way to carry out this measurement is taking the weighted average of two measurements of slope: one determined via gravity acting on the phone, and one determined by the rotation of the phone about the x-axis. [91] The gravitational measurement requires a user to position a phone on a ramp such that the slope of the surface is equal to the counterclockwise rotation (noted by the
variable $\theta$) of the phone about its x-axis. \[91\] Assuming gravity to be $g = 9.8 m/s^2$, the z-axis of the phone is equivalent to the following equation \[91\] :

$$g_z = -\cos \theta$$

This means that the slope of the angel can be written as \[91\] :

$$\theta = -\cos^{-1} g_z$$

As $g_z$ is given by the iOS Core Motion Framework (CMF), the slope calculation function is able to directly use this information. \[91\] In the second method, the rotation is obtained about the x straight from the CMF (denoted as $\phi$), and then perform the following calculation to obtain a weighted average of the two measurements (and thus the slope): \[91\]
\[ \text{Slope} = \frac{(\frac{1}{\sigma_1})^2 \cdot \theta + (\frac{1}{\sigma_2})^2 \cdot \varphi}{\frac{1}{\sigma_1} + \frac{1}{\sigma_2}} \]

In this equation, \( \sigma_1 \) and \( \sigma_2 \) represent the standard deviations of \( \theta \) and \( \varphi \), respectively.

[91]

**AccessSound**: To accurately measure sound, Johnson (2015) showed that one could use the iOS API to capture 100 samples of sound intensity over a given timeframe using the front microphone. [95] The sound levels are then graphed over the period of time they were taken, to indicate the range of sound that was available to the user – the median is highlighted in particular for the user’s benefit. [95] Samples are needed due to the variable nature of sound in environments – a sudden pop every so often should be differentiated from a low droning noise, and obtaining an average will not reflect the pattern of sound over time accurately. Visualizing the sound helps the user make a decision based on the sound levels.

Now, one might question why one shouldn’t automate the collection of all the sensor data at once, instead of asking the user to select a preferred tool and use it if they feel it is necessary. Asking the user helps the user feel in control of the assessment completion and the data gathering, which makes the overall experience of using AccessTools more positive. Consider how happy individuals were when using the dynamic queries, which gave them full control over on-the-fly filters for controlling their data. [51] Helping the users remain in control of the computer and ensuring that software offers them choice is also a fundamental tenant of information visualization. [25]
2) AccessTools

The AccessTools application allows a user to select a location near to them for accessibility assessment, and complete a TTSS assessment that will offer insight regarding the accessibility of the particular built environment. The branching structure of a TTSS assessment is particularly helpful in the case of AccessTools, as building features each constitute a category and a given building may not have all features included in the exhaustive question set. [33] Locations are populated using a Google Maps API, and one location can have multiple evaluations completed for it at a given time. Evaluations, once complete, can be reviewed in-app. Figure 4-2 illustrates the main assessment screen included in AccessTools.

As mentioned, AccessTools uses a TTSS-styled taxonomy, consisting of topics corresponding to building elements. This design was created by an interdisciplinary team, which involved themselves in the study of ergonomics and accessibility guidelines before constructing their taxonomy based on said guidelines. [96] Details ask for specific information about a particular building element, some of which require the use of a sensor measurement to fully complete. [33]

Regarding the user interface, it has been designed to mirror that of assessments generated by xFACT, although with a darker color scheme and some resized elements due to the app running on an iPad and needing to be compatible with touch support. Instead of clicking between items in the assessment overview, they can be tapped in order
to open and close them individually. Plus symbols still indicate which items can be "drilled down" into for additional detail.

However, to further improve usability, an addition has been made to the user interface: in-application help. Assuming a design is intuitive is itself poor design. [33] If the user is confused at any point about what a section of the taxonomy is aiming to measure, they can view a video that discusses the aims of the section at length. If a particular question confuses the user, they can access additional help by tapping the information button in the upper right-hand corner of the detail view. [33] In addition to giving further details about a question, this button often will give information related to the standards a question was inspired by - be these standards from an ergonomics standpoint, or from the ADA.

![Figure 4-4: A screenshot of the main assessment view from AccessTools.](image)
Icons at the bottom of the item detail view represent media and sensor measurements that can be captured. The first two icons allow a user to capture images and video of the environment that they are assessing; useful for "proving" why a question was scored as such. The next three represent activating AccessRuler, AccessSound, and AccessSlope, respectively. The icons are only active on questions where using the measurement tools would make sense; even in these cases, using the tools is optional.

Some may wonder why one would not use sensor measurements for every question requesting input on the measurement of a built environment feature. In some situations, sensor tools (though innovative!) may not be necessary. Consider the case of measuring the accessibility of a ramp, only to immediately see that the ramp is broken - it

Figure 4-5: Example of slope measurement within AccessTools.
would not make sense to continue the measurement when the user knows it will not be accessible due to the crack down the center of a ramp! Allowing a user to take advantage of these tools, without requiring their use, allows AccessTools to help (and not hold back) a user. [33]

Since building features separate sections of the AccessTools taxonomy, a user can pick and choose building features that apply to the building they are assessing - an excellent example of the TTSS assessment's efficiency in the field. This also means that a user can customize an assessment's features to improve the measurement workflow for a building. [33] Sections that aren’t applicable to the building being measured can be marked as such. [96] This, coupled with the TTSS strategy of allowing an assessment to be completed in any order allows the user to have full control over what elements of the assessment are completed at what times, and what is directly relevant to the building they measure during that particular session. [33] It also allows experienced assessors to stick with their preexisting workflow (if one exists) without accidentally hindering it. [33]

D. An AccessTools Workflow

When using AccessTools, a user will start by selecting a location to evaluate. They can do this via either looking at existing evaluations and selecting one of these to continue or starting a new evaluation and selecting a location to assess. If they opt to select a new location, they will need to enter an evaluation name, a place name, a location and ensure that their assessor information is correct, before proceeding to the taxonomy screen.
The taxonomy screen allows a user to parse the questions available for various built features. Some of these features assume that there will be more than one existing - such as the 'Doorways' section, where the assessment is split into "Main Doorways," "Other Exterior Doorways," and "Inside Doorways." The assessment here follows the rules of a TTSS assessment: questions are answered in a trichotomous format, and a user can also drill down into any of sections marked with a plus symbol and answer additional detailed questions to ensure certainty.

The assessment can be saved or uploaded to the cloud at any time. Multiple assessments can be taken for a single location; the responses are not aggregated kept separate. Comments can also be added for any question; this section allows a user to elaborate on a given answer if a particular question needs it. If a question is confusing, or a section does not make sense, the user can view a video (per section) regarding the point of measuring that section, or detailed help for many questions in the assessment.

In some cases, such as when answering questions in the Ramps section, a user may be asked a question such as "Are the runs of the ramp accessible?" When it isn’t immediately apparent, the user may benefit from taking data. In this example, the user would need to place their device on the ramp, and press the slope tool button, followed by the 'new' button. On the next screen, a user can select what type of slope they are interested in obtaining, and wait for a moment as the data is collected from the iPad sensors. The result is displayed, along with an interpretation of the result. Upon saving the result, an indicator appears alongside the AccessSlope icon at the bottom of the detail
view, indicating that a measurement has been taken and can be viewed - not forcing the user to view or use the measurement!

Completed assessments can be viewed online, along with all associated measurements taken and media available from the assessment process. Assessments can also be exported as a CSV via email, for processing in other software. Reporting, as with xFACT, is in the works.

E. Comparison of Data Taken with AccessTools

To explore how AccessTools works in practice, data was collected for three restaurants in Milwaukee, Wisconsin near to the campus of the University of Wisconsin-Milwaukee. For each assessment, data, measurements, and media were collected as

![Data Count per Study Location](image)

Figure 4-6: A graph representing the measurements, comments, and data collected using AccessTools from three restaurant locations.
needed to understand the accessibility of the location. A “complete” assessment was finished for each location, meaning enough data was collected to understand the accessibility of each location.

As expected, total counts of data varied, despite using the same taxonomy in all three cases. In some cases, collectors chose to mark items as non-applicable, reducing the number of categories they had to consider and questions they had to answer. In other cases, the raters chose to use the measurement tools to collect data from the environment to check their answers. In Figure 4-7, two assessments in-application are depicted, each with different question score totals and data available. A few of the differences between these assessments include:

- Landing accessibility is marked NA for the Qdoba location, while it’s marked 3/4 at Noodles and Company – a nod to not only the fact that this question branches, but there is not a landing to rate at Qdoba.
- Due to items marked non-applicable, the score for Ramps is a 4/8 at Qdoba, while it’s a 15-20 at Noodles and Company.
- The raters decided that no additional measurement was needed to assess the slope at Qdoba, but took a measurement at Noodles and Company.

It’s evident that AccessTools fits a number of use cases – and the quantitative data integration is useful for those questions where a user is confused. Using the same taxonomy, two locations can have very different total items in the assessments of their accessibility – a testament to the power of a TTSS-styled assessment.
Figure 4-7: Screenshots of two completed assessments in AccessTools. Note that different questions have different totals, representing different building features existing in each example. The latter screenshot also indicates use of the slope tool to complete a question.
F. Discussion

With this approach, data collection is integrated into the system. In a previous system, the mini-tools would need to be accessed to take data, which could be distracting and difficult to maneuver. Opening the tools "in" AccessTools keeps the experience seamless. Furthermore, the user retains agency by not forcing the results of data collection on them - they can choose to score according to the results of data capture as they see fit. The use of automated data capture at all helps the efficiency of the assessment process, by allowing a user to remove uncertainty from their assessment process. Finally, smartphone sensors are used for data collection when possible, both to save the user money and ensure they can take advantage of this data collection system.

The organization of data using the TTSS assessments helps users parse complex data in two main ways. First, it helps users understand the context of quantitative measurements – the question this work was attempting to answer with each measurement. Second, it offers the interpretation of the data first and foremost, letting the drill-down technique be used to view data backing up a particular response. These two points in conjunction create a more usable interface than directly aggregating measures in a table.

Two considerations unique to this project are usability concerns related to the application running on a mobile device, and consumer-facing access. The former reflects the understanding that AccessTools is an iPad application, and as such specific usability concerns need to be taken into account to ensure users have a well-designed application. The latter involves a secondary part of Access Ratings Project: AccessPlace, a consumer-
facing website that allows users to view business with accommodations that fit their abilities around them, [154] and how one might integrate AccessTools data for such an audience.

1) Accessibility

One consideration that is particular to AccessTool’s nature as an iPad application is that of mobile device accessibility. This typically refers to ensuring that the application is compatible with assistive technology for tablets, such as switches, Bluetooth keyboards, and Apple's built-in voice synthesis system, VoiceOver. [98] It also requires the revisiting of some user interface elements, to ensure the application is touch-friendly. The tutorial videos also need to be captioned and potentially have audio descriptions added – something that the iOS platform does have support for. [98] This work is currently in progress. While steps have been taken to improve usability, even over the design of xFACT with tutorial videos and in-application help, switch compatibility and compatibility with iOS assistive technology (such as VoiceOver) are imperative to ensuring a universal design.

2) Integration of AccessTools Data with Consumer Website

AccessTools allows a user, typically trained, to view the data collected in the tool, using the TTSS-styled data representation to organize the measurements taken from the environment. However, this data is displayed to a consumer in an easy to understand format using the same principles! In creating a consumer TTSS-styled data view, some requirements of the system differed from the standard presentation requirements. A user
who is walking through AccessTools collecting data understands the fractions that occur with each line - they are familiar with the data being collected and have that as a reference to understand what is being presented back to them. Unfortunately, a user who did not collect this data might find themselves a little confused. How should they interpret a 2/6?

Thus, for a consumer, scores are converted to percentages - this conveys a "completion" message and makes more sense when one is trying to convey that some number of requirements of built environment access were met, or went unmet. However, another problem occurs when it comes to item sets where different permutations of answers could result in the same percentage score.

As an example, for four questions, one can receive up to 8 points - the maximum score for a branch node with some number of leaves is \( B = 2 \times L \), where \( B \) is the branch’s total score and \( L \) is the number of leaf nodes. Thus, a user can receive a 2/8 by:

- Having a single answer marked yes, and the rest marked no, or
- Having two answers marked maybe, and the rest marked no.

For this reason, it is a good idea to implement a small graph. Tufte was a proponent of "small multiples," which he offered as a good summary of data. [41] A small multiple was a tiny, easily-generated graphic that was repeated multiple times. A sort of "small multiple" is created in the form of a tiny bar graph, with coloring reflecting a stoplight: the first bar reflects the number of 'No' options, the second the number of
'Maybe' responses, and the final the number of 'Yes' options in the leaves of that node. If the node is a leaf, it reflects the answer for that leaf.

To reflect this in a view, data is arranged in a hierarchical format that mirrors the typical TTSS view, with a small bar graph on the far right side and a branching indicator on the left. The design here was developed to be mobile-first to ensure the largest number of people could access it; thus while the detail view still exists (and lists measurements, media, and a longer description for each item), it is on a separate page. Bar graphs always represent the children of a given node, or if the node is a leaf, the node's score. The illustration shown in Figure 4-8 is of a preliminary example of this user interface brought to life, in the final the scores shown as part of the graphs will be illustrated as a percentage.

Figure 4-8: Preliminary consumer view for AccessTools data. Data in this example has been mocked up for the sake of illustration. TTSS Branching Outline organizes data for display, allows for drilldown if needed. Mini-Graph illustrates yes/no/maybe count. ‘1’ is placeholder for percentage score.
G. Summary

AccessTools stands as a more efficient method of applying a TTSS assessment, with the goal of capturing data from the environment and providing it to a user so they can take accurate data. AccessTools only requires one "special" tool - a laser ruler - as opposed to toolkits of the past, which required one to use a multitude of different tools for assessing the built environment. [91]

The TTSS structure of the assessment allows for the efficient collection of data; only those questions that are relevant to the structures that a user is assessing are shown at any given time. [33] Branching, as always, allows a user to dive into detail for a particular question as necessary. [33] Furthermore, the collection of quantitative data is allowed via device sensors, so that a user can adequately answer questions in the AccessTools assessment without requiring to stop and take measurements. It should be noted that allowing for sensor-based data capture is not without its problems: two sensor tools were put on hold, due to changing APIs from Apple during the development process.

Finally, the problem of adapting the TTSS view for a consumer was investigated. Direct users of AccessTools will be familiar with the data, and thus more familiar with what the fractions representing data measurements are telling them. However, consumers may need additional context to understand what data has been taken. To meet these consumers' needs, a consumer view of the data was designed, keeping echoes of the TTSS-styled assessment view, while adding small graphs and a mobile-first design to
ensure that everyone is able to access and understand the data being presented, regardless of their data science background or device they are accessing the data with.
V. ENHANCING MEDIA COLLECTION AND PRESENTATION VIA UNIVERSEAL DESIGN

As storage cheapens and processing power improves, mobile devices can be used to capture increasingly complex varieties of data. In particular, media often can provide important context to the reasoning behind measurements taken. After all, when discussing something like repairing the built environment, sending an image of the deprecated structure rather than just marking it as inaccessible can often convey the extent of the inaccessibility, rather than numbers alone!

Media has proven helpful as supplementary data in other tasks that deal with complex data. For example, when disasters strike - such as the devastating 2010 earthquake in Haiti - it can be expected that many at the scene will have mobile technology, ready to take photos and video of the event occurring. [99] This allows people who live in places far from the event to understand the magnitude of what has happened (leading to donations), or help response teams organize requests for help via priority (such as those who have no food or water, or who are trapped). [99]

Some challenges are involved with the inclusion of media in a data collection process. A first-time user of this assessment application might have contextual questions that an image, by itself, wouldn’t be able to answer. These questions include:

- Why exactly was this particular image taken?
- Who can benefit from this information?
• What is this image telling the user?

While some images indeed speak for themselves, others may require a bit more work to ensure that the information offered by the image is apparent to the user. Some context will be provided by any questions or text accompanying the image, but this may not present what was intended by the user related to a particular image. Proper titles and descriptions can also improve a user’s understanding of the context surrounding an image, but these also may not carry the full meaning intended by the user who initially captured the media. Finally, even if one user can discern why a particular image is included with data collection, the point remains that not everyone will be able to.

Universal design ensures that most people can enjoy the use of an application and understand it, regardless of their ability. For example, implementing alt text for website images is an excellent example of universal design. The media may not load for users who are accessing a system on a poor internet connection, causing them to become reliant on the alt-text. However, an individual trying to access the image may also have poor visual acuity, and thus might rely on the existence of alt-text to understand the context and reasoning behind the inclusion of an image. Consideration of features to include to improve the system for all folks, regardless of their background or abilities, will help the overall usability of this application in the long run.

As discussed, two leading research problems crop up when discussing the use of media in data collection systems: (1) how to standardize the incorporation of additional contextual information when gathering media, and (2) how to ensure the
additional information adds to the users’ understanding of the media gathered, regardless of their background or ability? In this chapter, it’s described how the inclusion of accessibility information for gathered media can solve these problems.

A. Understanding Media Content and Context

As mentioned in the second chapter of this work, retaining context when gathering massive amounts of data can be a challenge. This especially becomes the case when data is gathered that was not coded for research purposes, but has now become available for integration into systems – such as phone call data, and social media data. [128] Typically, for this new data, content analysis strategies are applied, and context is extracted related to the creation of this data: timestamps, locations, and in the case of social media, the relationships between users. [128] However, the content of media is another thing entirely. The advent of machine learning has made understanding the content of media easier, but in all cases, machine learning is not up to par with human coding in terms of context extraction from media. As an example, consider current methods of audio, video and image content analysis.

1) Understanding Image Content

Understanding the content of images goes hand in hand with the fields of object recognition and detection, both of which have gotten easier with the advent of machine learning systems, such as Google’s AutoML [129] and Wolfram’s ImageIdentify [130] projects. Deep neural networks can be trained on vast quantities of labeled images, [131] which can pick out features of labeled images and thus acquire an understanding of what
a dog looks like versus what a cat looks like. Wolfram’s ImageIdentify project is even able to delegate sub-classifiers, such as dog breeds, [130] while Microsoft’s CaptionBot is able to label images based on what people in the images are doing. [142] However, elements of image context – such as identifying the historical context of an image – still require manual coding. Understanding image content has been of interest to those wishing to automatically generate captions for users with vision impairments – however, as it stands, some users are still disappointed with the level of detail and mistakes made in automatic alt-text generation systems. [35]

2) Understanding Audio Content

Understanding the content of audio requires differentiation between spoken word content and musical content, first and foremost. [132] Some opt for further classifications, such as separating audio features into speech, silence, laughter, and nonspeech, or speech, music, silence, and environmental noise. [132] No matter the categorization method, once audio has been separated the spoken word is typically where the content of the audio lies – and this can be analyzed via speech recognition. [135] This can also be performed on real-time audio, for real-time captioning systems. [135] Unfortunately, modern automated audio transcription isn’t just yet a perfect science - automated captions generated by the popular website YouTube, for example, are only accurate 60-70% of the time. [136] Text for indexing can be captured from an audio file, but it cannot be used to generate entire caption tracks. [137]

3) Understanding Video Content
Understanding video content requires a mixture of these analysis techniques – understanding the audio, and understanding the content of keyframes in the video. However, additional context is available via the construction of scenes in the video – groups of shots that share some content, be it a location or a theme. [147] Analyzing this with the other components of a video relies on the use of multimodal processing algorithms. [147]

In all of the cases detailed prior, the automatic analysis is just a tad lacking – it does not differentiate contextual information about the media and gives a fundamental overview of what the media includes. While this will change in the future, for now, to ensure all of the meaning of media is understood, it’s best to derive context manually. However, what’s the best way to do this? Integration of accessibility data can help significantly in this process, allowing users to ensure that media is accessible for data scientists and users with disabilities, and allowing vital contextual data about media to be added into the system for further analysis.

B. Universal Design and Media

Accessibility is a hot topic in web and application development. Many software companies – large and small – put great emphasis on ensuring a design that is usable for all of their users. With the release of Windows 10, Microsoft created a separate website outlining amendments one can make to adjust the operating system for users with low vision, hearing loss, mobility impairments, and neurodivergent users. [100] Likewise, Apple designed ways for their popular AirPod earbuds to be used as assistive technology.
for those with hearing loss, and ensured that their devices are compatible with switches for those who have mobility impairments. [101] Independent developers are also pushing accessibility in software, releasing free tools for brain-computer interfaces [102] and screen readers [103] for those who need it (and for software testing).

There already exist some excellent reasons for folding accessibility into work, such as avoiding lawsuits [104] and ensuring no portion of a customer base is accidentally abandoned. However, in providing additional context for users who may have disabilities, the developer also provides that additional context for everyone. Indeed, what are typically considered accessibility provisions for media often end up helping more populations than intended – a perfect example of universal design. In this section, methods of ensuring accessible media are explored, along with whom they help and how they could affect the context offered for each type of media.

1) Accessibility and Images

Non-text content, traditionally, requires the use of a text alternative to be considered accessible – this is most notably the case in the Web Content Accessibility Guidelines, wherein it is stated that “all non-text content that is presented to the user has a text alternative that serves the equivalent purpose.” [148] In the case of documents such as PDFs, this can be as simple as character recognition of the text within the document, and provision of this text for screen readers hoping to access the document. However, for documents that are difficult to extract text from, and images, alt-text is used. Alt-text offers a textual description of the image for users who may not be able to perceive the
image in question and may be read aloud by the computer. [148] Alt-text is traditionally implemented for users with vision impairments, but may also help users with cognitive impairments, or users with a slow internet connection. [107] In any case, it’s a classic example of universal design: assisting multiple populations with their perception of a media item.

Unfortunately, alt-text is a tricky idea to implement, as it is typically left up to the developer to decide how much contextual information about an image should be included. This can lead to alt-text of wildly varying quality. In lieu of traditional alt-text, the Equivalent Text Description developed by the R₂ D₂ center of the University of Wisconsin-Milwaukee offers three tiers of additional information for the user with varying levels of detail. [105] Officially, EqTDs offer “written descriptions of graphic elements” - useful for many different uses, such as those with a visual impairment or low vision, in addition to students whose native language is not that of the image, and those with cognitive limitations that impede their understanding of the image at hand. [106] To achieve these means, EqTDs are constructed with three types of description in mind: two working to fulfill accessibility goals, and one to fulfill usability goals. [105] In order to ensure the accessibility of graphic elements, and EqTD includes a “Brief Description,” which describes the “type and purpose of non-text information,” and “Essential Description,” which offers context and understanding about the image at hand. [105] An additional “Detailed Description” presents the user with a finer description of the elements of the graphic, including “layout, colors, logos…” and so forth. [105]
By focusing on these three descriptions, a system using EqTDs offers a quick overview, an understanding of the image context, and a detailed description of the image’s contents to users. As all of this information is to be displayed as text, this information can be read aloud and added to a pool of searchable information for the image. Any text offered by the user can also be embedded in future machine learning algorithms as a description of the image contents.

2) Accessibility and Audio

Captions are required for all audio, including that accompanied by a video, for the audio to be considered accessible. [148] Unlike alt-text and EqTDs, however, captions are relatively easy to implement. Captions merely requires writing in text the words spoken and sounds that occur in the audio and synchronizing the text with the audio to relay in “real time” what the audio is conveying for users who cannot perceive the audio. [148] Transcripts, text documents representing the audio without any timing information [133] are also beneficial, both for users who may not want to listen to the video or may not have the bandwidth to his technique is quite beneficial for those with hearing impairments, those accessing the content in another language, or just those who cannot access the audio at that point in time.

Adding a text representation of the audio to the pool of data about the context of that audio means bringing in all information conveyed in that audio file for search and analysis. While one may previously have known that an audio file was a 30 minute podcast was about a murder mystery in the 1990s from the description, with the transcript
of the audio, a user is able to potentially identify who is speaking, what the mystery was, how it ended, and all of the spoken information from the hosts about the story. This vastly improves the information surrounding this audio file on hand – searches can be conducted for specific jokes in an audio file, rather than trying to remember the overarching description and content of the file.

3) Accessibility and Video

For videos to be accessible, they require not only captions but an audio track that divulges context of the video playing in real time. [148] This audio description acts to inform users who are not able to view a video as to what is happening, rather than relying on the sounds of the video. Consider an action sequence in a film – it would be quite

![Figure 5-1: A screenshot of a TED talk video available on Youtube, with captions.](image)
difficult to know who was making what movement without some sort of additional
guidance as to what was happening on-screen!

Audio descriptions are not as widely supported as other accessibility provisions,
but they have seen an increase in support in recent years. [134] Audio descriptions
provide context for a video that is currently running, to a user who cannot perceive a
playing video. [149] Audio descriptions might describe the background of a set, what
characters are doing, and any visual gags shown onscreen in movies. From these, a user
can obtain an alternate description of the context of a scene - this is good as clarification
as to what one should be getting from a scene playing in a movie. While audio
descriptions are not required to have textual counterparts, speech recognition could
indeed be performed to extract this secondary layer of information about a video.

C. Improving the Searchability of Information by Including Accessibility Information

The impact of adding accessibility information to media on the meaning derived
from that media was investigated during this implementation of a fully accessible media
archive during the creation of the Fred Sammons Archives, a website celebrating the life
and times of prominent occupational therapist Fred Sammons, Ph.D. (Hon), OT, FAOTA.
[140] In developing the Fred Sammons Archive website, research goals were as follows:

- To ensure that everyone can access the archives, regardless of location or
  abilities,

- To create a high-quality and highly-searchable archive of data available
  for researchers and students.
As Sammons dedicated his life to ensuring an improved quality of life for those with disabilities, it made sense to ensure that measures were taken to ensure the accessibility of the site. By ensuring accessibility of the site, searchability was ensured as well, and the website because of a stronger tool for researchers interested in the life and times of Fred Sammons. [150]

When generating accessibility information, one must understand what a particular media format is conveying – and translate the message of the media into another format. Images become text; a video is augmented with audio - all to ensure that people with (and without) disabilities can understand why media is being included, and the message that is being conveyed. However, this process has side effects. Improving the semantic information surrounding media can also make using the media in follow-up data analysis steps easier. The more information surrounding a video or an image, the easier it is to perform something like a prediction task such as using the media.

In the Fred Sammons Archive (FSA), media was categorized as one of five things: documents, images, video, audio, or objects. [141] Certain categories had special provisions – documents, for example, could be typed or handwritten. [141] [150] Objects were composed of several images showing angles of Sammons’ inventions. [141] [150] A summary of media available in the archives is available as Table 5-1, and a comparison of the media included in the archives and its resulting accessibility provisions is available as Table 5-2. The accessibility provisions taken were as follows.
Table 5.1: Media available in the Fred Sammons Archive. Items in bold are accessibility-related. [141]

<table>
<thead>
<tr>
<th>Media Type</th>
<th>Count</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documents</td>
<td>521</td>
<td>Documents in PDF Form</td>
</tr>
<tr>
<td>Photos</td>
<td>31</td>
<td>Photographs in PDF form</td>
</tr>
<tr>
<td>EqTDs</td>
<td>174</td>
<td>Text descriptions of photographs and some documents.</td>
</tr>
<tr>
<td>Fred’s Notes</td>
<td>119</td>
<td>Audio from Fred describing handwritten documents</td>
</tr>
<tr>
<td>Transcripts</td>
<td>127</td>
<td>Text transcription of documents.</td>
</tr>
<tr>
<td>Captions</td>
<td>372</td>
<td>Text transcriptions of video</td>
</tr>
<tr>
<td>Videos</td>
<td>372</td>
<td>Video interviews</td>
</tr>
<tr>
<td>Video Descriptions</td>
<td>373</td>
<td>Audio context for videos.</td>
</tr>
<tr>
<td>Objects</td>
<td>209</td>
<td>Images of objects. One object will have multiple views: each are counted.</td>
</tr>
</tbody>
</table>

Table 5.2: Media, related accessibility information, and searchability. [141]

<table>
<thead>
<tr>
<th>Primary Artifact</th>
<th>Accessibility Information</th>
<th>Searchable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documents</td>
<td>Text files and Fred’s notes (audio and text)</td>
<td>Yes</td>
</tr>
<tr>
<td>Photos</td>
<td>EqTDs</td>
<td>Yes</td>
</tr>
<tr>
<td>Videos</td>
<td>Captions and audio descriptions.</td>
<td>Yes (Captions)</td>
</tr>
<tr>
<td>Objects</td>
<td>EqTDs</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Equivalent text descriptions (EqTDs). EqTDs were constructed by students for all non-text items. Only brief and essential descriptions were added for items in the FSA – a quick overview for screen readers, an additional contextual detail as needed.[139]

Closed captioning. For all video files, closed captions were offered in two ways: both synchronized with the video, and available as a text file for download.[138] The commercial website Rev.com was used to generate these captions.[138]

Optical Character Recognition and PDFs. As there are a great many documents included in this archive, optical character recognition (OCR) was used to recognize text in these documents.[139] This allowed PDFs to be easily read by screen readers.

Figure 5-2: Example of image available from Fred Sammons Archive. “Photo of Fred standing next to a colleague.”[144]
Audio descriptions. Videos included in the archives numbered in the hundreds, and were repetitive in nature – interviews, where Dr. Roger O. Smith spoke with Sammons, occasionally lifting and using media – a notebook, an invention – in the

Figure 5-3: A screenshot illustrating the integration of accessibility options into an artifact view. [141] Items such as the audio introduction, general audio description, and video captions offer additional information about the media onscreen for users of the archives.
Rather than relying on audio descriptions for all items in this archive, audio introductions were created for each item, and designed one generic audio description to describe the set up of the interview. [139]

**Audio notes.** As some documents were too detailed for a simple EqTD, and handwritten (preventing the use of optical character recognition), a new type of accessibility information extraction from these documents was developed, referred to as “Fred’s Notes.” [139] This unique type of audio file required the dictation of handwritten items, while also providing any background he felt was appropriate for the items. This often led to entirely new information about items – context and stories that would not be

---

**Figure 5-3:** A screenshot of the front page of the Fred Sammons website. [143]
available at first glance. As a result, transcripts were obtained for all ‘Fred’s Notes.’” [139]

Besides these accessibility provisions, the website was tested to ensure it followed all provisions of the W3C Web Content Accessibility Guidelines, in addition to provisions by Access-ED for universal design standards. [139a]

D. Results

Final results can be categorized in two ways – the impact that the accessibility provisions had on research quality, the impact the accessibility provisions had on the universal design of the site, and the impact the accessibility provisions had on media context.

1) Research Quality

The Fred Sammons Archives’ research capability was greatly enhanced by the addition of accessibility data – far more information was available for the search engine to process. All text information could be added to a pool of searchable data for each artifact, upgrading the degree of search that was available as part of the Fred Sammons site. This meant the text generated by both Fred’s Notes and PDF optical character recognition, as well as that added by EqTD writing and caption generating, is added to the pool of searchable information and contextual information for each artifact. [150] Instead of relying on every mention of the words “button hook” in the title or description of an artifact, one could search primary sources themselves for mentions of this.
Furthermore, in some cases – such as the Fred’s Notes audio – the addition of accessibility information added new information that wouldn’t have been added to the database at all if not for efforts to add accessibility information. [150] This greatly improved the richness of information available to parse through. [150] Not only could a user parse keywords, titles, and descriptions, but also captions, transcripts, text files, and EqTDs. [141] [150]

2) Universal Design

Another side effect of adding this accessibility information was the impact it had on the universal design of the website overall. A user who wasn’t interested in watching or listening to a video could opt to read the transcript, available as a downloadable text file for every video. Users who, likewise, wanted to multi-task, could listen to Fred’s Notes for a handwritten item, instead of parsing the item themselves for the mention of a topic or person. Offering additional modes of accessing media for users helped ensure that more users could take advantage of this media at any given time.

3) Media Context

Finally, to see the impact these accessibility provisions had on media context, the content added by accessibility provisions was compared to the content without the accessibility provisions. Manually-added content for images was also compared to generated content, provided by the Microsoft CaptionBot service. [144] A sample of the data gathered is available as Appendix Four.
In total, a sample of 33 items was analyzed: 6 videos, 6 objects, 15 documents, and 6 pictures. Documents were a mix of typed and handwritten.

**Title and tags:** In most cases, the title and tags for an artifact offered some brief indication as to what the artifact contained, but no more than this. As an example, the Fred Sammons Archives contain a number of day planners [141], which are all titled as a “Fred Sammons Day Planner,” followed by a month and year. The keywords often echoed the information listed in the title – the day planners often had “daily,” “planner,” and “timer” listed as keywords. [141] Artifact 77, a photo titled “Photos Kamakura statue in Japan,” the keywords “photograph,” “statue,” “Japan,” and “Kamakura.” [141] However, at times information offered by the title was narrowed down by the keywords – such as in the case of Artifact 619, a video interview titled “Colorado OTs: Doris Schriver.” [141] Here, the keywords offered some additional information, such as “Bell telephone,” and “work hardening” – offering a little more information as to the content of the interview. [141]

**Description:** In most cases, the keywords offered as much or more information than the description – in the case of artifact 612, a video titled “Serving as an Expert and Consultant,” the description reads “Fred relates his contribution as a consultant.” [141] However, the keywords include “national research council” and “national academy of sciences” – providing context in the form of specific locations that Sammons contributed to as a consultant. [141] However, in 5 cases, the descriptions of the artifacts offered additional context not provided by the title and keywords. Artifact 629, a video titled
“Family Farming” with keywords such as “farm” and “family,” was described as “Fred describes pictures of himself working on his family’s farm.” [141] While the title and keywords nodded at the notion that the artifact dealt with a family farm, only with the description can a user see that Fred is, in fact, looking over family pictures.

**Metadata:** Metadata is a very structured way of providing additional information when compared to using a title or description – it usually relays specific information consistently, such as the era of the document or what categories (from a set number) the media fits into. In this case, metadata varied depending on the artifact type. For videos, metadata included the length of the video – which could be useful in assessing whether or not Fred had additional things to say on a topic. [141] Objects had no time information, properties or categories associated with them. [141] Images had a time frame associated with them, called an “era,” which allowed one to arrange artifacts in a sequence of categories in Fred’s Life. [141] Documents had the most metadata: document properties, an era, and categories. [141] For this reason, for all artifact types except objects, metadata offered additional information to the pool of searchable information.

**Captions:** In this analysis, three types of captions were considered: captions generated using Microsoft’s CaptionBot service, EqTDs manually generated, and transcriptions of audio and video information. The latter always added more information about the context of the media and content involved in a bit of media, as the full content of the transcribed audio/video became text, which one can perform textual analysis on. This would allow for the assessment of information available in each video, search videos
for mentions of specific topics, and even analyze aspects of the spoken words such as places, people and things most frequently referred to. EqTDs typically generated more information, though this was restricted to the essential description and not the brief one (which sometimes acted as a glorified title). As an example, one can consider the essential description for artifact 78, titled “Orthoplast Splint Cut-Outs.” [141] Although the keywords offer additional information (the item has to do with Johnson & Johnson), as does the description (indicating that the artifact has a mixture of illustrations and descriptions), the essential description from the EqTD offers the most information as to the content of the item:

“The 11-page document contains patterns from Johnson & Johnson for a variety of splints, such as a toe splint, ulnar deviation splint, functional wrist splint, and a RIC tenodesis splint. Also included is a pamphlet from Johnson & Johnson with instructions for making a splint and for ordering splint materials and patterns. Lastly, there is a sheet with instructions for how to butt-bond orthoplast splint material to form longer or larger sheets.” [141]

This offers full contextual information for a user with disabilities and a lot of additional search terminology for content analysis.

Finally, the automated captions were often lacking – many items in the Fred Sammons Archives are hand drawn signs or diagrams, which CaptionBot cannot parse at this time. [142] Furthermore, that which could be parsed was often identified wrong – in the case of artifact #870, a poster from the 1992 AOTA conference, CaptionBot wrongfully identified it as “an envelope on the floor.” Interestingly, there was also a lot of variance in CaptionBot’s responses: signs were each identified as “a close up of a logo,” “a close up of a piece of paper,” and “a envelope on the floor.” Interestingly, in
one case, CaptionBot did offer information that was not otherwise offered by the captions manually generated: in the case of artifact 931, a photo of Fred and a colleague standing next to each other, CaptionBot identified that they were standing in front of a window.

**Audio Descriptions:** For this analysis, the content of audio descriptions was not analyzed, as the content of the video did not change much from interview to interview. However, it is worth noting that the final level of detail would come from an audio description in the event that scene information did change from video to video. The mood and meaning of audio, music and the like will change depending on the scene displayed by any accompanying video.

E. Discussion

In general, the addition of accessibility information improved the searchability of the Fred Sammons archive, universal design of the archives, and context available for media across the board. Meaning was extracted and converted to other formats, notably text, which is more comfortable for content analysis systems and search engines to parse. With text, one can perform textual analysis to understand the features of the text such words used and co-occurrences, which can help one better understand of the content of media. [143]

Providing additional context for media included in the Fred Sammons Archive could help the automatic caption generators of the future understand what proper captioning looks like. Machine learning algorithms could be trained on the image and video descriptions, and improve their feature recognition. This in turn could speed up the
analysis of images and video in the future, and allow for reliance on automation (at least partially) for generating accessibility data.

Furthermore, ensuring accessible media guarantees that is can be understood by any who may work with the data in the future - collectors, analysts, and businesspeople. This helps businesses ensure compliance with regulations that may require data to be accessible [69] and to advocate for accessibility across the process of data science.

F. Summary

Ensuring media accessibility can benefit many people – those accessing the media for data analysis purposes, future machine learning algorithms, and those interested in creating a rich pool of data around a media artifact for research reasons. As shown via the integration of accessible media practices with the Fred Sammons archives, integrating media accessibility can help develop a powerful research tool, utilizing various accessibility features in developing additional context and background for media that is not available via titles, keywords and metadata alone. In extending this work to developing a body of media context for data collection and analysis, a solution becomes evident for the problem of context generation for media integrated into data collection systems.
VI. DISCUSSION

The simple TTSS interface outlined here – and stated modifications to this interface – hide a lot of heavy lifting. Figure 6-1 illustrates this, considering the user interface as the “tip of the iceberg”: observations, media, complex data and the like all hidden below the surface. The interface outlined organizes, structures, and provides a stable hierarchy for the data gathered, along with allowing the data gathered to be filtered depending on the user’s needs. Measurements provide “proof” of a recorded result, reason behind why a question in a trichotomous assessment was answered in a specific way. Adding accessibility data additionally helps retain context of data, ensuring the overall accessibility of data in the long term.

Figure 6-1: A representation of the work carried out by a TTSS system. The TTSS interface remains visible, hiding large quantities of data from the user and only displaying what the user needs – in a sense, acting as the tip of an iceberg.
This approach works well for data collection - breaking down the data collection task via a hierarchical task analysis, and structuring it via a TTSS-styled taxonomy. In doing this, a replicable data collection protocol is generated. This hierarchy allows the generation of trichotomous scores at each level of a taxonomy, and thus view an overview score (a top-level score that reflects all questions) and drill down into sections to view scores in each of the sub-sections therein. This helps summarize data that has been collected, view the summaries, and thus understand this data as it is collected.

As mentioned, the flow of this approach can be impeded when a system requires measurements or the integration of other complex data. However, since the assessment is computerized, there’s no reason not to automate the data collection procedure as well. And so, to make the process efficient, the collection of quantitative data can be integrated via sensors, and offer it as a decision support tool for users – allowing them to view measurement data in the system, and proceed as they desire. The TTSS structure also organizes the incoming data, relating it to specific questions so data can be drilled into for details, just like items in an assessment.

With the capture of media in a system for data collection, one runs into the problem of extracting meaning from media. There are various automated ways to do this, but none are particularly mature just yet – so manual context collection is preferred. To structure this, adding accessibility data for each media item added to the system is suggested. In doing this, one ensures that individuals with disabilities are able to understand all of the data being collected by the system and provide the system with
information as to the meaning of media collected. Meaning is added when accessibility data is collected, and thus this is a perfect solution for the integration of media context parallel to the collection of observational data.

A. Implications of Work

The most significant implication this work has is on the existing protocols in place for data collection for non-data scientists. Data collection can be an intimidating task – understanding what data to collect, and why it should be collected, can be daunting. With the system outlined, not only can be create reusable data protocols thanks to exhaustive and categorized question sets, but:

- Questions are easy to answer, due to their trichotomous nature, and
- The assessment’s structure can help in educating users in regards to the different aspects of data collection to consider for a particular topic.

As an example, consider the AccessTools assessment mentioned before – data collection is separated by building feature, and details explore what constitutes ‘accessibility’ for a particular building feature. AccessTools’ trichotomous questions also make the assessment easy to understand for users who are not data scientists – the assessment is simply a series of questions that can be answered with a ‘yes,’ ‘no,’ or ‘maybe.’ This helps to educate a user who may be learning about built environment accessibility. In addition, a user who is more advanced can skip around and answer items related to specific building features in any order they wish – so TTSS assessments don’t let down power users in their efforts to be usable for beginners.
However, another implication exists, due to the unique presentation approach this work has – using the *presentation* aspect of the assessment. The presentation of this system can be used to aggregate data about things – separating the things into a series of topics, or categories, and at the most detailed point offering measures and media for that particular topic. Precautions are taken to ensure that the measures and media are accessible. This is a format that could be useful for a number of things, but in particular e-commerce might find it useful. Products, easily organized into categories and sub-categories, could be structured in a TTSS-inspired manner – at the most detailed level, presenting measures (reviews, colors, and data) about products being sold. Since we build in accessibility for these measures, this would help e-commerce sites ensure that their presentations are accessible, which would be of benefit to them.

B. Strengths of System Design

This system has a number of strengths in its current state – the most prominent being its flexibility. One taxonomy is designed to work for multiple data collection cases; in the case of AccessTools, the taxonomy is designed to work for multiple buildings which might have a plethora of features. Features can be designated as non-applicable in the event a particular building does not have them.

Additionally, all decisions are ultimately left up to the user. This provides the user with full agency, and helps them feel in control of their data set. Measurements are integrated, but as decision support – they are folded into the assessment, which is used to offer information as to why the measurements might be helpful (i.e. a ramp question
asking about the slope of a ramp) so a user can understand. It’s worth noting that automated completion might be of benefit to users, and can be a potential future endeavor.

Finally, our system prioritizes usability and accessibility – especially in using accessibility data to provide data context for all users, in addition to computers. Prioritizing accessibility is a great way to ensure that a design is universal, and everyone can use your work. By putting this idea first, we can ensure that our work is appropriate for all of those who may be interested in it.

C. Weaknesses of System Design

Some might consider the need to design an entire assessment prior to data collection a weakness of this system – or at the very least, a feature that goes against the idea of exploratory data analysis. After all, one might not be able to assess all that data that will be needed to solve a problem. Constructing an assessment is useful for structuring data that one can predict, but what about those measurements one cannot predict, that only become known as being needed after the fact?

The solution here is merely to reassess a taxonomy regularly and ensure new information doesn’t need to be added. Taxonomies should act as a structure for data and as a data collection protocol: something that is agreed is necessary in data science. Unstructured data can make the process of eliciting meaning from it more difficult in the long run. The act of breaking up a problem into multiple sub-problems is celebrated in data science, and can help with the parallelization of data
collection and analysis across a large team. [125] Additionally, while many do not follow a cut-and-dry method for collecting and analyzing data in their companies, structured thinking can help one assess where human intuition is necessary to proceed, versus the use of a computational tool. [125]

Another criticism that might be lobbed is that even though it’s more pleasant for people using the system, one should rely less on human judgment when it comes to making decisions and let the computer take charge. The creation of a system for presentation isn’t as beneficial for individuals as the creation of a system that analyzes data for the user and comes to a conclusion on their behalf. This is, after all, where the beauty of using computers for decision-making comes into play – computers can catch details that people cannot!

While this is undoubtedly true, the opposite is also true – people can often catch details that computers miss. There is a limit to the number of decisions that can be automated. [127] For this reason, decision support tools – ones that offer a user data to help them come to a decision – are quite popular in particular fields. One such field is the medical field - clinical decisions range from diagnoses to management decisions, and in all cases, a precise amount of data should be offered to the decision-maker – enough to assist, not enough to bombard or overwhelm. [126] This often requires decision support tools to focus on three different types of decision support: information management tools, tools for helping one focus attention, and tools for offering recommendations for a specific patient. [126] In all cases, however, the emphasis is put on the tool acting as
decision support – assisting the doctor in making their decisions, and not overriding the expertise of the doctor.

Finally, some may indicate that a weakness of this system is the act of adding accessibility data – as noted, while some of the addition of accessibility information can be automated, a lot of it must be manually added. This is especially problematic in the case of automated media collection, where media is collected over time automatically – such as scraping Twitter for images. Depending on how many images are scraped a day, a system just wouldn’t be able to analyze every picture, without a significant amount of cost, time and effort.

This is true – adding accessibility data for collected media can be a process, and an overwhelming amount of work when it comes to media collected automatically. However, much work is being done in the fields of automatic caption generation and extraction of content meaning from media. [34] [35] [36] In time, automatic and manual attempts will be able to be combined, and thus reduce the amount of work required for ensuring accessible media when collecting data.

D. Future Work

There are three major places this work could be extended. The first is the integration of the media accessibility data into AccessTools – in doing this, one could have a complete mobile system for TTSS assessments. AccessTools as is allows for new taxonomies to be loaded via the web interface, so integrating a field or two for
accessibility metadata would allow the application to that advantage of these recommendations.

The second place is the extension of the automated data capture. Right now, data can be captured from smartphone sensors, as in the case of AccessTools using said sensors to help a user make decisions related to accessibility. However, in the future, this capability could be extended, via capturing data from sensor networks or drones and using middleware to translate captured data into TTSS-appropriate answers. In this way, the entirety of the assessment completion process could potentially be automated based on data capture. This would help facilitate the shift from human-powered data capture, to a tool that captures “big data” – automatically sorted and human-readable. Multiple days of data (such as the aforementioned weather example) could then be plotted and analyzed.

Finally, in the future, the integration of additional assessments into xFACT would be fantastic to see. At the time of this dissertation being written, interest has already been noted by several parties, and the AccessTools project, in particular, had just received a $600,000 grant. With burgeoning interest and the potential for improving existing data collection and presentation methods that these approaches have, there’s no doubt that assessments that build on this work will continue to be used for data collection and presentation in the future.
VII. CONCLUSION

In this work, the creation of a data collection and presentation system that allows for the usable human collection of observational data is discussed, which extends Smith’s work (1993) in developing a functional assessment for people with disabilities. [52] The TTSS assessment illustrated by Smith can be used for purposes beyond assessments for occupational therapy, which is illustrated by walking through the development of two new assessments using this framework. This TTSS assessment structure can additionally be implemented in an iOS application. The assessment app integrates automated data collection via sensors and media capture, in order to improve the efficiency of the original system. Finally, faced with the problem of integrating context with data, the novel idea of integrating accessibility data in order to provide context is considered. This approach does improve the overall searchability (and thus, the potential for analysis) of the media added to a website, and is suitable for retaining the context of data added to a system.

A. Contributions to the Field

In short, this work’s contributions to the field in this work are as follows:

• Evaluation of using a TTSS framework, previously used in occupational therapy, for general data collection by an individual
• Development of a system for aggregating data from the environment via smartphone sensors for built accessibility determination.
• Analyzed the differences between standard metadata, generated accessibility data and ‘full’ accessibility data for media.

• Developed a modular system for displaying accessibility data for media in a web archive.

B. Broader Impact

The work carried out in this dissertation has the potential to improve the data collection efforts of everyday people who need to obtain data for their work, or are interested in taking advantage of data analysis to benefit themselves. As an example, consider a school teacher – they may be curious about whether or not a new method of teaching a sub-area of algebra is working for her class. With the tactics outlined here:

• They can create a TTSS assessment for observing a student’s understanding of algebraic concepts

• They can integrate data from her grade book as the solutions for some of the assessment questions

• They can use OCR’d tests as primary sources in her data collection, allowing that data to be accessed by her computer as well.

The work here can also be applied to those looking to devise ways of understanding employee morale, planning for upcoming projects, and things of this sort. By putting usability first, all people are able to take advantage of this method of data collection and presentation.
VIII. REFERENCES


IX. PUBLICATIONS


APPENDIX 1: EXPLORING THE WORKFLOW OF THE xFACT TAXONOMY EDITOR

The ‘General’ tab allows an individual to set metadata about their taxonomy, including the title, version number, editor name, and the last date the taxonomy was edited.

The ‘Response Sets’ tab allows a user to create response sets. Several styles are available to choose from. Users can rename the options as they choose, and all styles conform to the TTSS method.
The ‘Outline’ tab allows a user to add new items to their taxonomy, indent them, and assign response sets to each item. The prompts tab (not shown) allows the user to create prompts of useful information for specific questions in the outline.

Finally, the assessment is scored after all editing has been completed by clicking the ‘Score Taxonomy’ button, and answering questions onscreen.
APPENDIX 2: DEVELOPMENT OF A TAXONOMY FOR OBSERVING WEATHER SEVERITY

Note that this taxonomy is in no way supposed to be exhaustive, it exists as an example of a potential data collection protocol made with xFACT.

I. PRECIPITATION

a. Rain
   i. Reduced Visibility
   ii. Travel Impact
      1. Impacts User While Driving
      2. Impacts User While Riding a Bicycle or Scooter
      3. Impacts User While Walking
   iii. Damage to Property

b. Sleet
   i. Reduced Visibility
   ii. Travel Impact
      1. Impacts User While Driving
      2. Impacts User While Riding a Bicycle or Scooter
      3. Impacts User While Walking
   iii. Damage to Property

c. Snow or Ice
   i. Reduced Visibility
   ii. Travel Impact
1. Impacts User While Driving

2. Impacts User While Riding a Bicycle or Scooter

3. Impacts User While Walking

iii. Damage to Property

d. Hail

   i. Reduced Visibility

   ii. Travel Impact

      1. Impacts User While Driving

      2. Impacts User While Riding a Bicycle or Scooter

      3. Impacts User While Walking

   iii. Damage to Property

II. EXTREME TEMPERATURE

a. Outside

   i. Outside, In Sun’s Path

      1. Poor personal comfort

      2. Danger to at-risk populations

      3. Danger to all populations

   ii. Outside, In Shade

      1. Poor personal comfort

      2. Danger to at-risk populations

      3. Danger to all populations

b. Indoors
i. Indoors, with Personal Temperature Adjustment
   1. Poor personal comfort
   2. Danger to at-risk populations
   3. Danger to all populations

ii. Indoors, without Personal Temperature Adjustment
   1. Poor personal comfort
   2. Danger to at-risk populations
   3. Danger to all populations

III. HIGH WINDS
   a. Instability
      i. Instability when walking
      ii. Instability when driving

   b. Dangerous Conditions
      i. Debris
      ii. Damage to Property
      iii. Damage to Buildings

IV. RISK OF DANGEROUS CONDITIONS
   a. Flood
      i. Water Impeding Travel
      ii. Water Impeding Comfort At Home

   b. Fire
      i. Fire Impeding Travel
ii. Fire Impeding Comfort At Home
APPENDIX 3: DEVELOPMENT OF A TAXONOMY FOR ASSESSING VPAT COMPLIANCE

Please note that all questions in this taxonomy are inspired by the VPAT. [111] Additionally, a change in answer set exists between Parts I and II, and II and II – this is a good example of where a prompt may be useful. (I/III use positive statements as proof of access, II uses positive statements to point out problems).

I. COMPATIBILITY WITH ALTERNATIVE HARDWARE

A. Keyboard and Mouse Usage

1) Keyboard is Compatible with Software

2) Focus Moves As Input Focus Changes

3) Focus Is Exposed For AT

4) UI Element Information is Available
   i. Identity
   ii. Operation
   iii. State

B. Touchscreen Usage

1) Focus Moves As Input Focus Changes

2) Focus Is Exposed For AT

3) UI Element Information is Available
   i. Identity
   ii. Operation
   iii. State
C. Assistive Technology Usage

1) Focus Moves As Input Focus Changes

2) Focus Is Exposed For AT

3) UI Element Information is Available
   i. Identity
   ii. Operation
   iii. State

II. INTERFERENCE WITH EXISTING ACCESSIBILITY FEATURES

A. Interference with other software features.

1) Disrupts existing software features.
   i. Overrides User Selected Contrast
   ii. Overrides User Selected Colors

2) Disables existing software features.

B. Interference with operating system features.

1) Disrupts existing software features.
   i. Overrides User Selected Contrast
   ii. Overrides User Selected Colors

2) Disables existing software features.

III. SOFTWARE INFORMATION/CONTROLS

A. Visual Elements

1) Use of Color
   i. Color Doesn’t Solely Differentiate Items
ii. Variety of Color Options Present

iii. Variety of Contrast Options Present

2) Images

i. Bitmap image meaning is consistent

ii. Bitmap information has text counterpart

3) Text

i. Text information is available via OS
   a. Content
   b. Text Input Caret Location
   c. Text Attributes

B. Animation

1) Text Flashing and Blinking

2) Alternate View Exists for Animated Content

C. Electronic Forms

1) Allow keyboard-only entry
   i. Information can be accessed
   ii. Fields can be accessed
   iii. Form can be completed successfully
   iv. Form can be successfully submitted

2) Allow entry via AT
   i. Information can be accessed
   ii. Fields can be accessed
iii. Form can be completed successfully

iv. Form can be successfully submitted
APPENDIX 4: A SAMPLE OF IMAGE DATA FROM THE FRED SAMMONS ARCHIVES

All data is from the Fred Sammons Archives, [141] except for generated captions, which are made using CaptionBot, a free service from Microsoft. [142]

<table>
<thead>
<tr>
<th>Artifact</th>
<th>Title</th>
<th>Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>931</td>
<td>Photograph of Fred Sammons and Colleague at RIC</td>
<td>photograph, fred, sammons, colleague, ric</td>
</tr>
<tr>
<td>34</td>
<td>Original Button King Logo</td>
<td>Button King, Logo, Original</td>
</tr>
<tr>
<td>442</td>
<td>Fred with Adaptive Device</td>
<td>fred, young, adaptive, device, photo</td>
</tr>
<tr>
<td>77</td>
<td>Photos Kamakura statue in Japan</td>
<td>photograph, Fred, statue, Japan, Kamakura</td>
</tr>
<tr>
<td>931</td>
<td>Photograph of Fred Sammons and Colleague at RIC</td>
<td>photograph, fred, sammons, colleague, ric</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Artifact</th>
<th>Description</th>
<th>Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>931</td>
<td>Photo with Fred with a fellow counselor at RIC</td>
<td>Era: Post-Bissel</td>
</tr>
<tr>
<td>34</td>
<td>Original Logo for Button King Enterprises</td>
<td>Era: Early Company</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Categories: creation</td>
</tr>
<tr>
<td>262</td>
<td>Fred Sammons working with a patient at RIC</td>
<td>Era: Early Career</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Categories: public relations</td>
</tr>
<tr>
<td>442</td>
<td>Young Fred using adaptive device</td>
<td>Categories: clinical practice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Keywords: fred, young, adaptive, device, photo</td>
</tr>
<tr>
<td>77</td>
<td>Photograph of the Kamakura statue Fred visited when visiting Japan</td>
<td>Era: Maturing Company</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Categories: marketing</td>
</tr>
<tr>
<td>931</td>
<td>Photo with Fred with a fellow counselor at RIC</td>
<td>Era: Post-Bissel</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Categories: None listed</td>
</tr>
<tr>
<td>Artifact</td>
<td>Generated Caption</td>
<td>A11y Information</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>931</td>
<td>a person standing in front of a window.</td>
<td>Photo of Fred standing next to a colleague. The picture was taken at the Rehabilitation Institute of Chicago and the man standing next to Fred. is either a doctor or a vocational counselor.</td>
</tr>
<tr>
<td>34</td>
<td>a close up of a sign.</td>
<td>This photo shows two red butting hooks and a logo that states &quot;the finest buttoning aides&quot;. Button hooks are one of Fred Sammons' first products that he manufactured himself during the early stages of his business.</td>
</tr>
<tr>
<td>262</td>
<td>a group of people standing around a table and he seems .</td>
<td>This is a photo of Fred Sammons working with an amputee at the Rehabilitation Institute of Chicago.</td>
</tr>
<tr>
<td>442</td>
<td>a man standing in front of a mirror and he seems .</td>
<td>This photo shows Fred Sammons holding one of the first products he developed, the Button-hook.</td>
</tr>
<tr>
<td>77</td>
<td>a group of people standing in front of a window.</td>
<td>This a picture of Kamakura State and Temple in Yokohoma, Japan. This is where Fred Sammons worked to develop an international market for his products. Circa 1978.</td>
</tr>
<tr>
<td>931</td>
<td>a person standing in front of a window.</td>
<td>The picture was taken at the Rehabilitation Institute of Chicago and the man standing next to Fred. [sic] is either a doctor or a vocational counselor.</td>
</tr>
</tbody>
</table>