

Article

End-User Needs of Fragmented Databases in Higher Education Data Analysis and Decision Making

Amanda Briggs^{1,*}  and Francesco Cafaro^{2,*} 

¹ Richard M. Fairbanks School of Public Health, Indiana University-Purdue University Indianapolis, Indianapolis, IN 46202, USA

² School of Informatics and Computing, Indiana University-Purdue University Indianapolis, Indianapolis, IN 46202, USA

* Correspondence: amanbrig@iu.edu (A.B.); fcafaro@iu.edu (F.C.)

Abstract: In higher education, a wealth of data is available to advisors, recruiters, marketers, and program directors. These large datasets can be accessed using an array of data analysis tools that may lead users to assume that data sources conflict with one another. As users identify new ways of accessing and analyzing these data, they deviate from existing work practices and sometimes create their own databases. This study investigated the needs of end users who are accessing these seemingly fragmented databases. Analysis of a survey completed by eighteen users and ten semi-structured interviews from five colleges and universities highlighted three recurring themes that affect work practices (access, understandability, and use), as well as a series of challenges and opportunities for the design of data gateways for higher education. We discuss a set of broadly applicable design recommendations and five design functionalities that the data gateways should support: training, collaboration, tracking, definitions and roadblocks, and time.

Keywords: human-computer interaction; higher education; user interface design; large datasets; data analysis; human-data interaction



Citation: Briggs, A.; Cafaro, F. End-User Needs of Fragmented Databases in Higher Education Data Analysis and Decision Making. *Informatics* **2021**, *8*, 42. <https://doi.org/10.3390/informatics8030042>

Academic Editor: Valentina Dagiene

Received: 18 May 2021
Accepted: 21 June 2021
Published: 24 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Higher education professionals encounter an abundance of data, from recruitment and enrollment data to current student and alumni data, and accessing and managing data have become incorporated into everyday work practices. End users (e.g., advisors, recruiters, marketers, and program directors) are often tasked with working as data analysts. While they typically do not receive any formal training, they are asked to consistently access and analyze data for decision makers within their units. They create reports that affect programs, recruitment, enrollment, marketing, student experience, and curriculum decisions. They are expected to produce “high-confidence results” that further the credibility of their unit and overall their campus and institution [1]. For end users, however, the magnitude of available data can be overwhelming. Large universities typically use myriad tools for accessing and analyzing data [2]—an example is illustrated in Figures 1 and 2. As a result, data sources often appear to be fragmented from each other, each containing disparate data that does not communicate with other systems. We want to acknowledge that this study is limited to five midwestern universities in the United States. Institutions in different regions, countries, or cultures may follow different work practices and generate different results.

In this paper, we report the results of an exploratory study that investigated how end users access and navigate these institutional databases, the challenges they encounter, the effect they have on work practices, and how database administrators work with end users. Our aim is not to provide novel techniques or methodologies to handle big data (as done, for example, in [3]); rather, we contribute to the field of human-computer interaction (in particular, Computer Supported Collaborative Work (CSCW)) with an exploratory analysis centered on the needs of the users of those databases. Specifically, we conducted a survey

with 18 users followed by 10 semi-structured interviews, from which we identified three recurring themes that affect work practices (access, understandability, and use). We report on a series of challenges and opportunities that emerged from current work practices and discuss a set of design recommendations and functionalities that should be supported when designing data gateways that facilitate access, understandability, and use of large datasets in higher education settings.

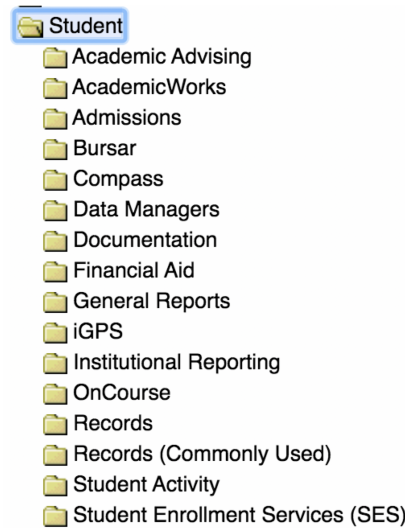


Figure 1. Views of one of ≥ 100 systems available for querying data at a midwestern university: an online folder menu for searching for data topics using an internal system.

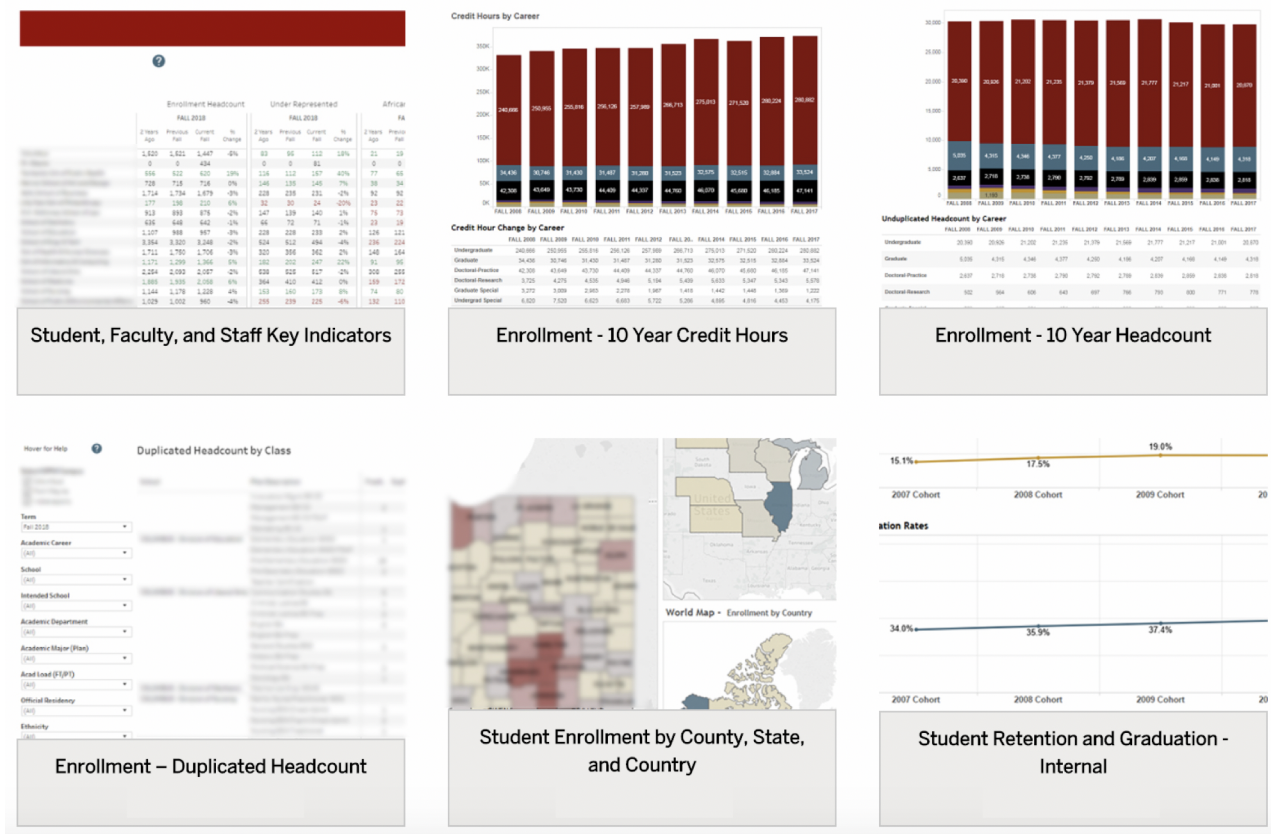


Figure 2. Another system available for querying data at a midwestern university: data visualizations available through Tableau.

2. Background and Related Work

Many organizations work with large, messy datasets, and higher education is no exception. Researchers have sought to determine best practices for wrangling [4] these large datasets into manageable systems. As best practices have emerged across businesses, these practices are being adapted and implemented by higher education for the purposes of recruitment and marketing. “Our ability to collect and store data is growing faster than our ability to analyze it” [5]. These broad datasets also create an impression of fragmented databases that do not work together, creating additional work for users who must maintain their own “homebrew databases” [6] through a combination of existing databases and spreadsheets maintained by individual users. On top of this, many institutional higher education leaders do not believe their college or university is making good use of the data it has available [7].

2.1. Human-Computer Interaction

Human-computer interaction (HCI) examines how people interact with computer systems. HCI started in the late 1970s, when the advent of personal computers brought computational technologies to the masses [8]. The initial focus was on the “usability” of such system [8], which “stems from the designers’ desire to improve the users’ experience” [9]. HCI incorporated multi-disciplinary methods from psychology, computer science, human-factors engineering, cognitive science, ergonomics, design, communication studies, and more. Few examples of HCI work that may be relevant for understanding the scope of this paper include assessing the user’s performance and productivity with technologies in the workplace [10], using eye tracking to understand how users interact with websites [11], and conducting qualitative research to understand users’ needs and work practices [12]. Additionally, recent work in HCI has investigated topics related with human-data interaction [13], including how people interact with personal data [14], and on the design of interactive installations that allow users to explore large datasets [15,16].

2.2. Working with Data

Large datasets pose questions about “the constitution of knowledge, the processes of research, [and] how we should engage with information” [17]. The accumulation of these data offers questions for those who have to work with it daily: What do these data mean? What sort of analyses should be carried out on these data? [17]. These questions are not always simple to answer. As organizations continue to collect more data, industry leaders are questioning how to better utilize the knowledge that can be gleaned from these data. When top-performing companies “use analytics five times more than low performers” [18], it becomes imperative that organizations excel at leveraging the data they access. Higher education is faced with the need to keep up with organizational demands [18] and yet has “applications, platforms and databases that do not ‘talk’ with one another and are difficult to integrate” [7]. As higher education continues to transition to a recruitment-focused mindset [19], the need for skilled professionals who understand complex datasets as well as databases that are designed to “talk” to one another is imperative. As advisors, recruiters, and marketers spend increasing amounts of time “wrangling data,” they are taken away from the primary duties of their job. “Data wrangling” represents the amount of the workday spent determining what data are available, organizing data, and understanding how to combine multiple sources of data [4]. Through data mining and predictive modeling, universities can discover “hidden trends and patterns and make accuracy-based predictions through higher levels of analytical sophistication” [19]. These sorts of predictive analytics involve an extensive knowledge of mathematics and statistics that are typically employed by institutional researchers rather than school- or unit-level professionals tasked with identifying trends through detailed data analysis. Additionally, the task of wading through large datasets to begin creating meaningful visualizations and reports creates additional work for users.

2.3. Infrastructure, Work Habits and Data Decision Making

As the focus increases on attrition, graduation rates, and interventions in higher education, the data used need to be accurate and timely [20]. Examining recruitment market data and analysis across industries reveals that most analysis occurs by domain-specific experts rather than relying on large-scale models that can delineate “fine-grain trends” [21]. In higher education, at the school and unit levels, this sort of analysis is performed by a domain expert who examines multiple databases, rather than employing a data analysis model created for the specific marketing and recruitment needs of each school or unit. These ad hoc data analysts are trying to communicate clear needs and goals for the organization while producing “high-confidence results” that further the credibility of the unit and the institution. Additionally, “the work creates a strong need to preserve institutional memory, both by tracking the origins of past decisions and by allowing repeatability across analyses” [1]. Users are often employing a mix of institutional databases and homebrew databases—data repositories for a diversity of information needs, stakeholders, and work contexts—and they often have constraints of time, funding, and expertise [6]. These databases are often an amalgamation of results from other databases, paper forms, and event data. Database administrators have to enter and analyze information across multiple information systems that provide different functions to the administrator and the organization. Bergman et al. found that users consolidate project information when a system encourages it and “store and retrieve project-related information items in different folder hierarchies (documents, emails and favorites) when the design encourages such fragmentation” [22]. The creation of homebrew databases in higher education suggests that users are taking cues from the database design and creating work arounds for the fragmentation that exists across the multiple databases that they access regularly. Data fragmentation can lead to database users feeling disempowered. When users feel disempowered by data, as found by Bopp et al., they experience an erosion of autonomy, data drift, and data fragmentation [23]. When users encounter new data, they struggle to find a place to update them within existing systems that are not designed for expanding with new data points. In 2015, Pine and Mazmanian assessed the data being used by healthcare providers. “Data must be located, quality of data assessed, idiosyncrasies identified, and inevitably a range of nearly intractable problems are discovered in the real messy world of data” [24]. This experience can be translated to anyone working with fragmented databases that do not yet talk to each other. While the data systems and roadblocks encountered by data analysts have the potential to leave users feeling disempowered, they also have the potential to create collaborative work [25]. For nonprofits, the solution for systemic data fragmentation does not reside with the individual organizations, but in the policy fields surrounding the data itself [26]. Researchers recommended banding together to create systemic changes in data organization rather than individual organizations developing workarounds suited to their particular needs. In 2016, Volda found that non-profit managers desired quantitative data for decision making while qualitative data “helped people ‘connect emotionally’ with the organization” [27]. The data collected in these organizations were found to drive decision making, providing “credibility to organizational actions.” However, Volda posited that one of the myths of big data is that hard numbers provide greater clarity and accuracy. Without the human stories behind the data, organizations can experience a shift in their identity. As education becomes increasingly reliant on digital data, these data are used to understand and predict human behavior, particularly potential student behavior [28]. Data technologies are becoming a part of university policy, creating a culture of data-based decision making in education.

2.4. Information Design and Data Visualization

Preparing data for visualization involves a multi-step process that includes discovery, wrangling, profiling, modeling, and reporting [29]. The most complicated part of this process is often discovery and wrangling. “Visualizing the raw data is unfeasible and rarely reveals any insight. Therefore, the data is first analysed” [5]. Even after preparing data to

create visualizations, “reconstructing a repeatable workflow is difficult without a coherent linear history of the operations performed” [29]. As analysts struggle to create meaningful, replicable visualizations, they have to ensure they are communicating the patterns found in data. Useful information design is critical as “expert” analysts provide decision makers with reports and dashboards [30]. Reports and dashboards hold the ability to communicate powerful truths through visualizations for decision makers. The remarkable amount of information we are accruing, however, often fails to be translated into meaningful visualizations [31]. Kandel et al. argued that the improvement of data visualization systems can help data analysts more effectively wrangle data. These systems would include a combination of data verification, transformation, and visualization [32]. An improved visualization system could help domain experts spend less time wrangling data and more time working on their domain specialty. “Visual data exploration is especially useful when little is known about the data and the exploration goals are vague. Since the user is directly involved in the exploration process, shifting and adjusting the exploration goals might be done automatically through the interactive interface of the visualization software” [5]. This could prove especially useful for novice users of data visualizations who struggle to identify appropriate views, execute appropriate interactions, interpret visualizations and match their expectations of the system with the reality of the visualizations [33]. For those in higher education, especially advisors, recruiters and marketers, discovering the real-world problems they try to solve with data visualizations will improve the work environment for these ad hoc analysts. However, proving the profitability of a given data system remains the key driver to system-wide integration [34]. The depth of study around big data and data visualization may be applied to higher education; however, there is a lack of across-the-board standards, combined with limited research (exception are found in [27,28]) on how large datasets (and the way in which they are currently accessed) affects work practices and the user’s experience around data access in higher education.

3. Methodology

The research design primarily relied on qualitative interview data supplemented with open-ended survey questions to identify all data sources used (and data-driven decisions) by marketing, recruitment, and student support professionals. Because our research was exploratory in nature, and mostly focused on a pre-design phase [35] aimed at understanding users’ needs in this application scenario (data in higher education), we took a qualitative approach to our data analysis—as commonly done in CSCW literature [36].

3.1. Participants

Participants were initially recruited through the survey which was distributed by email based on publicly available directory information at their college or university of employment. Survey participants indicated if they were willing to be contacted for a follow-up interview. Additionally, database administrators were contacted to participate in an interview to understand their role and the future of these database systems. Survey and interview participants were employed at five U.S. colleges and universities: Franklin College, Indiana State University, Indiana University, Purdue University, and Rose-Hulman Institute of Technology.

A total of 28 users participated the study, with 18 participating in the survey and 10 providing interviews.

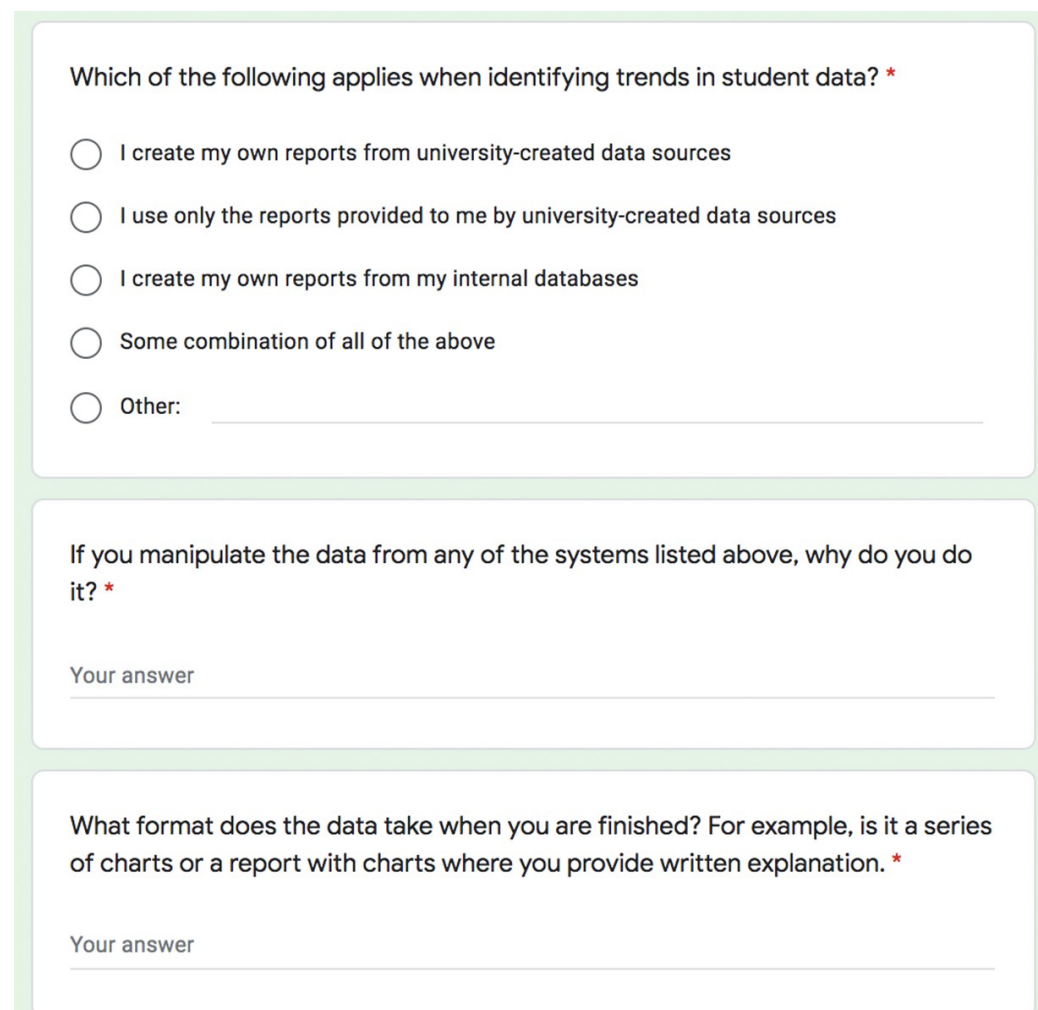
The 18 survey participants ranged in age from 18 to 55 and consisted of 5 men and 13 women. All participants held an associate’s degree or higher and worked full-time at their college or university. Participants self-identified their roles which included: recruitment, advising, marketing, program coordination (recruitment, advising, student success, etc.), graduate admissions operations and advising, student and administrative services, marketing and recruitment, programming, and data and IT.

The 10 interview participants ranged in age from 26 to 55 and consisted of five end users (those accessing the data for their daily work—marketers, recruiters, and advisors)

and five database administrators (those maintaining the databases that end users accessed—institutional analysts). The average age of the end users was 37.2 years old. Three interview participants identify as female and two as male. Their employment roles included academic advisors, program directors, and admissions managers. The database administrators included three women and two men, with an average age of 46.8 years old; all participants held a bachelor's degree or higher, with three of the five participants holding a doctoral degree. Database administrators' employment roles included a vice chancellor, directors, and analysts in the institutional research support office. The vice chancellor was considered an administrator because this role is responsible for the data.

3.2. Survey

A survey was distributed to database end users. Participants included end users at five midwestern colleges and universities. The participants were briefed about the nature of the study using a recruitment email. Before beginning the survey, the participants read and agreed to an informed consent document. Participants then completed an online Google Forms survey (see Figure 3) that included questions about the participants' use of databases at their respective institutions, data analysis and reporting, and the convenience of the data source, as well as demographic questions. At the end of the survey, participants were asked if they were willing to participate in a follow-up interview.



Which of the following applies when identifying trends in student data? *

- I create my own reports from university-created data sources
- I use only the reports provided to me by university-created data sources
- I create my own reports from my internal databases
- Some combination of all of the above
- Other: _____

If you manipulate the data from any of the systems listed above, why do you do it? *

Your answer _____

What format does the data take when you are finished? For example, is it a series of charts or a report with charts where you provide written explanation. *

Your answer _____

Figure 3. A screenshot of some of the survey questions administered to end users across five colleges and universities. * denotes required fields.

Database use questions included quantitative questions about the types of database systems used and how users identify trends in data. Following these questions, users were asked a qualitative question: “If you manipulate the data from any of the systems listed above, why do you do it?” Following this question, users were asked to identify what form their data analysis took post manipulation: “What format does the data take when you are finished? For example, is it a series of charts or a report with charts where you provide written explanation.” Further, users were asked about the time they spend analyzing and manipulating data, the ease of finding all of the data they need in one place, the datasets needed to report on data, and how the data they report on are used. Users were also queried about the challenges they face when collaborating: “What are the challenges you generally encounter when collaborating with others outside of your unit?” Finally, users were asked how they would improve functionality in the systems they use: “If you could improve any function in these systems, what would it be and which system(s) would it apply to?”

3.3. Semi-Structured Interviews

We conducted 10 semi-structured interviews with end users and database administrators. Interviews took about 20–45 min to complete (either in person or by phone). All interviews started with an overview of the study and an assurance of anonymity and confidentiality. End user interviews included a series of questions to understand which database systems users engage with and what sort of information they provide for the user. The interviews ended with a series of questions about what users like and dislike about the systems they use, what they would change if they could, and what they would like if they could have anything to make their work easier. Additionally, users were asked about their experience in data analysis and assessment. Database administrators’ interviews began with their history of working with the database system and what purpose the system serves. Participants were asked about what they feel the system does well and what could be improved, how user testing is incorporated, how feedback is elicited, and if systems can “talk” to one another.

3.4. Analysis

The survey consisted of 18 open-ended questions and 8 demographic questions. Participants also identified the database systems they most frequently use to access and analyze data. All interviews were recorded (in total, 4.5 h of interviews) and transcribed. Because of the volume of information that we collected, we decided to follow a two-step process for our analysis.

In Step 1, a member of the research team with extensive experience in higher-education administration work used the results of the survey and interviews to identify the idea units [37] (i.e., we were not coding for keywords, but rather full sentences that explained the same concepts) from the transcripts. Then, a team of two researchers used thematic analysis to cluster those idea units around three macro-themes: access, understandability, and use [38].

In Step 2, a group of five researchers further analyzed the (many) idea units within each macro-theme to identify the challenges and opportunities related with access, understandability, and use. Specifically, the idea units related with access were organized into an affinity diagram, and then we repeated the process to create two affinity diagrams for understandability and use.

We concluded the analysis when the two researchers who were analyzing the data unanimously agreed that we had reached theoretical saturation [39], i.e., coding additional data points was not adding any new information to the analysis. Because the researchers agreed that we had reached theoretical saturation, we concluded the survey and interview process.

4. Results

In this section, we illustrate the design themes that we identified as related to Access, Understandability, and Use. Our description includes participants' remarks that exemplify challenges and opportunities for the design of data gateways for higher education. An overview of these themes is reported in Table 1.

Table 1. Overview of the design themes that we identified as related with Access, Understandability, and Use.

Macro-Theme	Design Theme	Description of Users' Needs/Challenges
Access	Data Source	Perception that data is not standardized across different systems
	Talking to Other Systems	The need for data systems to be better integrated
Understandability	Asking the Right Question	Not understanding where to obtain specific data needed for decision making
	Data definitions/coding	How the terminology surrounding commonly used data is defined
	Experience	The disparate level of experience between end users and database administrators
Use	Relationship building	Building relationships to facilitate collaboration
	Reporting/analysis	Report creation for decision making purposes
	Decision making	Aiding unit leadership in determining which decisions are worth investing resources
	Time	The time it takes to find data can be frustrating

4.1. Access

Access refers to how users retrieve data and how they regularly utilize these databases. Our participants mentioned the challenges related to accessing data, or their perception of being able to access to data, is essential to engaging with the available data systems. Participants mentioned that when they lack trust in a data source or understand how a data source relates to the data within the ecosystem of the university as a whole, this leads the user to feel less credibility in the data provided as a whole. In these campus- and university-wide systems, there is a concern that data systems are not communicating with each other, which causes an additional sense of limited access to data sources. In the affinity diagram, we clustered the idea units that we previously classified using the access macro-theme into two sub-themes: (1) data source; and (2) talking to other systems.

4.1.1. Data Source

For end users, there is a recurring perception that data are not standardized across different systems, creating conflicting data results within the same unit, for example because of the way data are generated and because of discrepancies and inconsistencies in data, reliability, and functionality of the systems. In the following, when we report participants' idea units; we use P to denote semi-structured interviews and S for the survey. Some end users have the perception that some systems contain wrong data. There is also a lack of understanding as to where the data originates for each system. This is particularly an issue when results are presented in different ways by different visualization tools, but all source data are coming from the same source.

"There doesn't seem to be an agreed upon consistency of where the data should come from when we're making consistent decisions." [P9]

As users are asked to create reports and provide information for decision making from the breadth of systems available to them, they question the accuracy of the systems. If the new data systems cannot provide data that are perceived to be accurate, the users cannot trust them moving forward as their personal, professional credibility relies on accurate data reporting.

"At times trying to get clarification as to what type of information is needed in order to decide which system to use to retrieve data." [S2]

“Some offices rely on internal databases (Excel) and have not trained on the campus-wide resources. That could mean we see wildly different information. It’s better when all parties pull from a consistent place.” [S1]

To improve functionality and usage, database administrators work to add new and unique data creating new opportunities for end users. In institutional research, administrators will create surveys to gather new data that help unit-, campus-, and university-level decision makers. Often, these new data will be integrated with existing data to create a broader and sometimes deeper understanding of various populations.

“A lot of the information that we get isn’t in any kind of database at all to speak of. We have a number of different surveys that we administer to faculty, staff and students and people like that. We will take that information and integrate it with information that we extract from SIS.” [P1]

Database administrators and end users both commented on the “shadow databases” that exist across units. These “shadow databases” are colloquially defined as the spreadsheets that a single user or single department maintains. These databases are not shared widely and are used to understand data trends or datasets that cannot be accessed otherwise. Both administrators and users understood the need for these databases, with one user mentioning that these databases can sometimes meet unit needs better than institutional resources.

“It seems like a lot of schools on campus have a ‘shadow database.’ A lot of it may be replicated in places like AdRx, but these databases match their needs a little better than what the university provides.” [P10]

A database administrator reiterated the occurrence of these “shadow databases,” adding that some end users do not see the benefit of sharing these data or collaborating with institutional resources.

4.1.2. Talking to Other Systems

Database administrators and end users cited the need for data systems to be better integrated to improve usability, data collaboration, and overall integration and accessibility of the data between systems.

“We could improve upon leveraging all the different data sources that are available on campus, like incorporating things coming from vendors like Academic Analytics and EAB.” [P5]

This sort of data collaboration would make reporting easier for end users, who then do not have to compare data across multiple systems. There are, however, limitations of data integration: if it is done by simply forcing users to switch to a single tool (rather than incorporating the users in the design process), it can result in users not engaging with specific systems anymore.

“I know the university wants us to use Salesforce, but it doesn’t communicate with other things. Sometimes it’s quicker to say this is going into Excel.” [P10]

4.2. Understandability

The theme Understandability refers to how users are often working from a viewpoint that is influenced by their understanding of data definitions and their level of university work experience. Database administrators mentioned that some users operate under colloquial definitions specific to their experience or their unit as to what a coding term means, rather than using the university-wide definition. This analysis led to three sub-themes: (1) asking the right question; (2) data definitions/coding; and (3) experience.

4.2.1. Asking the Right Question

For database administrators, a key to helping end users find the data they need is coaching them through how to ask the right questions. Most end users, however, feel that

a lack of experience or training hinders their knowledge, causing them to feel that they are not appropriately using or querying the data system. For example, being able to identify trends such as accounting students taking courses at another institution requires users to formulate questions that enable database administrators to research the answer. For end users, obtaining the “right” data can be an exercise in trial and error. Not knowing where to obtain the specific data or report that they need for decision making can be a challenge.

“And you don’t know if it’s going to be the right report when you get it, you know. There’s some questionable stuff, it’s just a trial in there.” [P6]

Database administrators see themselves as serving as translators between the person who is making a data request. However, knowledge of the campus culture is important in being able to meet users’ needs.

“You have to understand people’s questions in order to know what data to use and when to use it and things like that. Well, part of that is kind of understanding the zeitgeist of the university.” [P1]

The complexity of questions that pass through a database administrator’s office can be difficult to predict, making it essential that such questions can be easily translated in a technical, data-oriented language.

“I would never have anticipated that the dean of business was asking me if our students are transferring to another institution to take accounting courses. That’s just not something I would ever dreamed up in my wildest dream, you know what I mean?” [P5]

4.2.2. Data Definitions/Coding

Participants’ remarks highlighted how there is a lack of consistency in how the terminology surrounding commonly used data is defined. For instance, one end user mentioned that the consistency in data did not occur because different datasets were wrong, but also that most end users operate with varying definitions of what data to be pulling, which could be caused by a lack of institutional knowledge.

“Not everyone has the same access to and understanding of the reports.” [S13]

“So there’s a lot of institutional knowledge that if you don’t have, you’re totally at a disadvantage when you try to begin pulling reports.” [P9]

The same end user elaborated on wanting the university to address data definitions by building in more “roadblocks”—or internal checkpoints sanctioned by the university—to ensure everyone was at the same level when reporting data:

“I think building in a few more roadblocks to help with the validity of what we’re using these datasets for, that would really take it to the next level.” [P9]

For database administrators, many of the discrepancies in the data can be attributed to an inconsistency in the data definitions among end users.

“People can sort of operate from a vernacular definition of what ‘enrolled’ means, for example, or what ‘first generation’ means. But without being able to ask specifically what it means for the data owner, or the steward, can influence how it is interpreted at the back end when it is queried.” [P4]

“If we are working with someone in another department the terminology, and reporting procedures can cause confusion.” [S16]

One reason that database administrators cited for the discrepancies in the data was that there is a lack of documentation on specific definitions for terms such as “enrollment” for both administrators and end users. Because the definitions do not seem to be consistent among end users, this creates a perceived lack of consistency in the data. Database administrators spend time working with units and end users to understand their process and share with them how to use the system to their benefit. One administrator mentioned that an information technology group is creating a cookbook with data definitions to create a trail documentation, but the process has been slow.

4.2.3. Experience

Participants mentioned that many end users begin their roles with a lack of training for analyzing and reporting on data. Between database administrators and end users, the level of formal training was vastly different. Database administrators had a high level of training and experience with degrees in fields related to data or coding. For end users, everything is learned on the job or from colleagues with more experience.

"It's just users using the systems and becoming experts over time because they've used the systems. And not because they're brought in and given a clean training." [P9]

"Familiarity with the system and functionality. Most users only use the very basic functions." [S16]

End users worry that their lack of experience leads to concerns about the validity of their analysis and reporting.

"There's definitely error that comes with learning. That's the danger of not having that formal option available. And I get that the training is difficult, because everyone wants to do different things with the data. But there's the validity questions that come up when we're all just kind of wild westing this experience of pulling information." [P9]

End users are learning from colleagues, asking questions of administrators and taking as many trainings as they can to become proficient in pulling data. However, most of them have no formal training in data analysis.

"When I was in financial aid, I had taken a few of the free entry-level workshops from UITS like SQL, data retrieval, stuff like that. So, no formal [training] outside of a couple of workshops." [P6]

The level of training administrators and users needed for their respective roles varied greatly. Most administrators were trained extensively in data retrieval and analysis with all interviewed administrators having some level of formal training ranging from a bachelor's degree to a doctoral degree. End users often have little to no formal training when they started their roles. While administrators are performing more detailed analysis for large-scale, campus and university decision making, this lack of training can be viewed as contributing to the varying levels of understanding that end users have regarding data sources, analysis, and how to ask the right questions for better data retrieval. Administrators acknowledge that becoming an expert in navigating these data systems is often about who you know that can help to explain the systems and sources.

"A lot of what you need to know to navigate these systems is dependent on whom you know. ... And it's very easy for people to assume they know what they're doing when they don't actually. This is one of the reasons that numbers sometimes get mixed up and people are pulling wrong information. Well-meaning people who go in who look at some code and oh, yeah, that's what you need. And it's not what you need." [P1]

4.3. Use

Use refers to how university data systems are employed by users to build reports that aid with decision making in their unit. As reported by all end users interviewed, users were split between providing reports that were generated by the data systems and manipulating and combining the data in some way to create reports for their decision makers (school deans, vice chancellors, etc.). There was a level of decision making associated with the use of these reports within units that included the allocation of budgetary resources, student experience planning, and campus- and university-level reporting. Relationship building emerged as a necessity for both end users and database administrators to understand both the data available within the system and the data needs of different units. This macro-theme generated four sub-themes: relationship building, reporting/analysis, decision making, and time.

4.3.1. Relationship Building

Relationship building was a unique byproduct of these data systems. Both end users and database administrators need to build relationships with each other to facilitate collaboration. University- and campus-wide databases can serve as a means of access for end users and oftentimes are used by end users as a way to communicate with students regarding opportunities available to them. When end users lack access to what they need to complete their job, they use additional notes and spreadsheets to facilitate relationships internally with colleagues and in furthering the student experience.

“If somebody were to ask me, ‘Hey, how many prospectives contacted you this month?’ I’m able to go back and look at my notes. . . . I can see how many students applied and I can see maybe the interactions I’ve had, too.” [P7]

Relationship building created operational benefits for the unit, campus and university where both database administrators and end users worked. These operational benefits include collaboration, facilitating work through relationships, management, and maintenance. One database administrator mentioned how important building relationships with end users and other administrators was for creating a sense of collaboration and facilitating work. In addition, for database administrators, relationship building is a crucial way of understanding how end users are accessing and inputting data. Another database administrator elaborated on this idea, commenting that understanding the context of the request and building credibility with others helps them to do their job well.

“I think that building the relationship helps us understand context more, and it also helps us in terms of building kind of that, for lack of a better term, kind of credibility to which we’re all in this journey together.” [P5]

4.3.2. Reporting/Analysis

End users create reports for decision making purposes within their department or unit. The analysis that goes in to creating these reports involves understanding the availability of data then modifying the data before analysis can begin. When users create reports, they typically involve some sort of data on customers (students) and either existing data visualizations or self-created visualizations for their supervisors or unit-level leadership. Users at the unit level typically tend to modify data, or blend and combine data, to create a richer understanding of what is happening with student-level trends. Users who regularly report on data customize the reports they are providing.

“I customize my reports, make notes to them. I share these on a weekly basis, which include enrollment data, to school leadership.” [P10]

“Some is a merging of the different data elements/parameters that are not available in just one system and organizing to tell the story of our findings.” [S6]

These reports help users to understand if they are on track within their unit and allow their leadership the opportunity to understand the trends in admission and enrollment data.

“Mostly to track our admissions and our applicants, make sure we’re on par for at least trying to come close to or beat the last few years of admissions and headcount.” [P6]

While users are often pulling together trend reports for their leadership, database administrators are trying to understand the system-level trends that are occurring and reporting out with white papers and research briefs to support decision makers.

“We’re sharing what we’re doing every day out there. We have all kinds of stuff on our website, you know, all kinds of, you know, white papers and research briefs and things like that.” [P1]

Administrators are also creating more data visualizations to aid users in reporting. Some users are combining these visualizations with their own internal data. The availability of these visualizations for end users has continued to increase through the use of Tableau reports.

4.3.3. Decision Making

Administrators want end users to be informed decision makers when accessing data, and simultaneously end users are helping their unit leadership determine which decisions are worth investing resources.

“We’re usually looking at trend data in terms of, is it worth it?” [P9]

The data that users are accessing also help to determine how programs are structured and can result in changes to academic programs. However, some end users find that the qualitative information, or student stories, makes a bigger impact with decision makers than the quantitative data.

“We use that data when we speak with the faculty directors and co-directors about how the program is shaped, what we’re seeing as far as trends and enrollment. . . . But, you know, sometimes I think it’s probably more of the independent meetings with the qualitative information that gets through easier than the quantitative information.” [P6]

End users create reports for decisions making with a frequency as high as once a week for their decision makers. While they may not be the final decision makers, they want to provide as much information as possible for decision making about admissions, courses, and program design.

4.3.4. Time

For end users, the time it takes to find data can be frustrating. Most users are accessing data on a weekly basis which creates multiple hours each week spent retrieving and analyzing data for reporting. Sometimes what are perceived as simple requests within a unit turn into a full day of work for users who have to access different datasets and create comparisons.

“What should be like, ‘Oh yeah, I can just get you that number,’ becomes like, well, maybe in a day or two, after I stare at things for hours, I can get that answer to you.” [P9]

These time-consuming activities feel frustrating to users who want datasets to be simpler to access without as much cross comparison. While one user says that all data can be technically accessed in one place, there are multiple datasets that need to be pulled to understand the trends and to perform analysis. Database administrators also find this sort of data querying to be time consuming and would love to provide reports immediately because end users need data quickly.

“People want things now, you know, and that’s not always possible, give how complicated data is.” [P1]

5. Discussion

At a coarse level, the three macro-themes (access, understandability, and use) led to a set of design recommendations that we discuss in the following. At a more fine-grain level, the challenges and opportunities that we identified within each macro-theme inspired a series of design functionalities that should be considered in data systems for higher education and other industries regularly accessing large amounts of data.

5.1. Design Recommendations

5.1.1. Access

According to Hick’s Law [40], the amount of time it takes a user to make a decision increases with the amount of choices available to the user. We recommend creating a space for users to access to all data systems in one location, aligning with industry best practices [41]. This gateway will provide a brief overview of each system, which will avoid overwhelming users with too many options, and it should also include a textual or visual description of the common data sources that are shared by different data systems (to limit the perception of dealing with fragmented or incomplete data). We recommend simplifying complex data retrieval steps into smaller steps for easier access to the data source to make

data retrieval choices quick and simple for users. More work is created for users the longer it takes users to interpret choices and make a decision.

5.2. Understandability

Both users and database administrators referenced the disparate definitions for key terms throughout use at the university. This could make data retrieval challenging if there were not a shared understanding of terminology. We recommend creating a standardization of data definitions enabling users to operate similarly as they do on other data systems. A standardized data retrieval structure will simplify the learning process for users and allow those with disparate levels of experience to work at similar levels. We also recommend simplifying the user experience by making all data retrieval systems operate similarly to simplify the process for end users. By visually representing similar codes near each other, users will group elements together if they are sharing an area with a similar boundary. This will aid users in contextually understanding data definitions.

5.2.1. Use

Users are engaged in both challenges and opportunities related to their regular usage of data systems. According to Parkinson's Law [42], all work will expand to fill the time allotted to it. Visualizing time and allowing other users to share their time will enable users to understand how long a task should take to complete. By creating savable, replicable, and shareable workflows, users will reduce the amount of time spent retrieving data to build reports, while creating consistent work across units, which will build credibility. We recommend using this space to apply the principles of collaborative design to teamwork to improve relationship building. Users working together to solve problems creates a shared understanding of both the problem and the solution.

5.3. Design Functionalities for Data Gateways

Data wrangling [4] remains a drawback for users. In line with existing industry practices [41], the design functionalities that we discuss in the following refer to a web-based gateway: the web-based implementation fits well the design space because database managers, end-users, and decision makers typically work at a desk in an office space, such as the one in Figure 4. In the following, we discuss five design functionalities that such a gateway should support for addressing the challenges and opportunities that we identified. The reader should notice that the focus of this section is not to present a specific, novel web gateway; rather, we discuss design functionalities that should be incorporated in existing solutions. In this regard, we want to highlight that, currently, there are no commercial data access management systems that incorporate these design functionalities.



Figure 4. The user workspace of the workers that we interviewed is a fairly typical office space; thus, our focus is on providing guidelines for designing desktop-based interfaces.

5.3.1. Highlighting Multiple Sources and Data Visualization

During the interviews, users mentioned that they sometimes struggle to identify the appropriate data source when beginning a project. By alleviating this pain point on the front end of acquiring data, users should be given equal footing when compared with their more seasoned peers in understanding how each system should be used and what data are available per system. The suggested uses for each data source should be provided by the university, enabling users to gain a system-wide understanding of data sources, as well as serving as an institution-level sanction on the data. The gateway should be a visual interface accessible through a website that allows users to discover and report on data [29]. This system should include a combination of verification, transformation, and visualization, as recommended by Kandel et al. [34]. The interface should allow for qualitative data to be imported to address the users' desire to tell a complete story for their decision makers, allowing them to "connect emotionally" with the customers they are trying to serve [27]. This will allow novice users to access data in a controlled manner that will create the roadblocks desired by database administrators. The gateway should eliminate the frustration expressed by users over credibility (or lack thereof) that occurs with slight variances in data. Users should be able to point back to a replicable pathway as to how data were obtained, transformed, and visualized. Users should also be able to footnote and source data as well, creating another layer of accountability for decision making and collaboration. This would reduce the "invisible work" [43] that end-users need to do to identify data sources and tools.

5.3.2. Training

Within the gateway, users should find a series of training modules that can guide them through how to use the systems available. The training needs to be university-sanctioned and approved to assist users to wrangle, analyze, and visualize data to elaborate on the "fine-grain trends" [32] found in the school- and unit-level data.

5.3.3. Collaboration

Interview participants were concerned about the difference in data pulled within units. To facilitate groups of users working together (within a unit, campus, or university), we

recommend implementing a system that tracks and collects reports pulled by users for access to other users. User’s profiles can allow them to access previous reports, as well the steps they used to access those reports (see Figure 5). The profiles may be available for viewing by other users within their unit for ease of replication. Creating a system that allows users to collaborate and develop shared metrics alleviates the problem created by fragmentation in organizations when the data do not seem to be “connected to each other in any systematic way” [17] This design functionality removes the onus off of individual communication and allows users the opportunity to discover how other users are accessing an utilizing the data on their own. This is different from the typical collaboration features implemented in commercial tools, which typically focus on enabling multiple people to simultaneously edit data and reports.

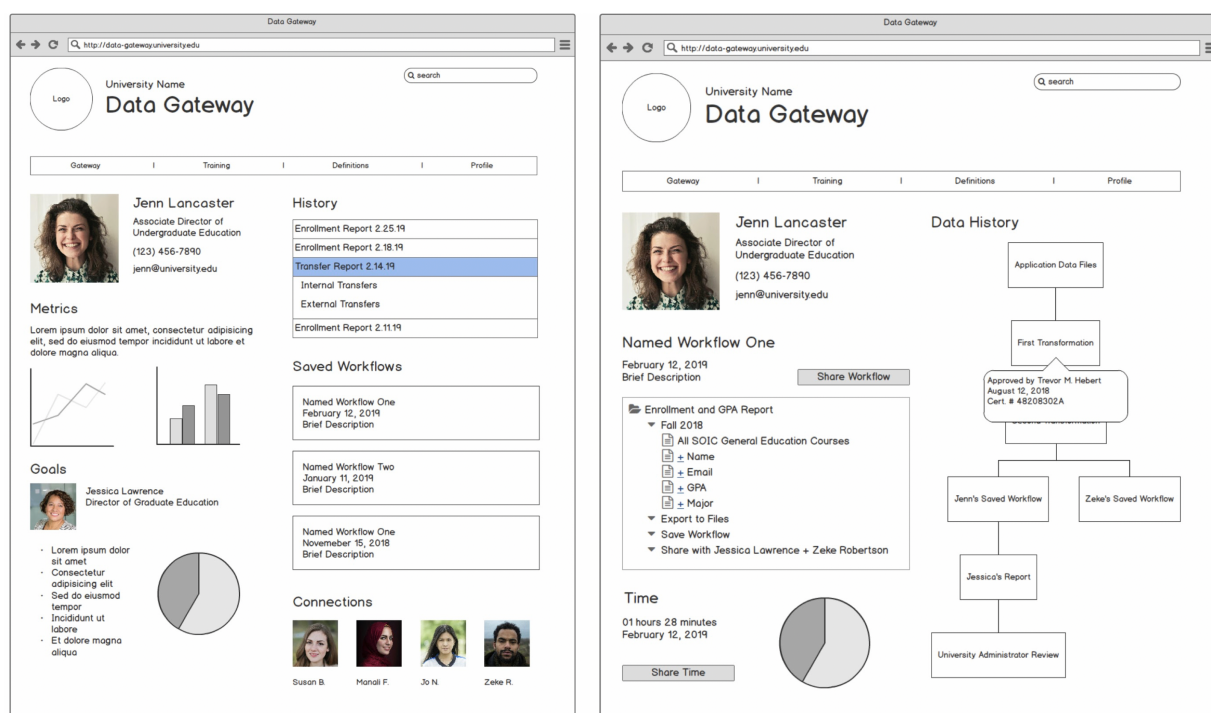


Figure 5. (Left) Rough framework for a profile (including a list of previous reports that other users can view and replicate and other functionalities to facilitate connections and collaborations across units). (Right) Workflows, Time, and the Data Ecosystem which includes a digital ledger system to track time and also visualize the workflow in the larger university ecosystem.

5.3.4. Tracking

Users and administrators can track the modifications and transformations that occur with data by enabling a ledger system. Tracking the data origination and modifications builds trust among users. For example, a private, permission-only blockchain that users and database administrators can access could create a clear history of the data and how it has changed hands. What the blockchain would offer that a traditional spreadsheet or tracking software cannot is that it creates a single, accessible, incorruptible tracking system that can provide immediate results on data history. It provides a fool-proof way of understanding and visualizing the data source and its transformations while building credibility for those who access and report on data. The digital ledger could be included in the gateway interface as users access the history of the data from their origination to the eventual database they are pulling them from, as well as any use by other users in their unit. The ledger enables users to build a sense of trust based on concrete assignments. This serves as a system of checks and balances within units as data are accessed and the data

access history is clear for all users. Database administrators can follow the breadcrumbs associated with data questions from end users.

5.3.5. Definitions/Roadblocks

A shared vocabulary was considered an essential foundation for all who access data. As higher education integrates more data technologies, working to predict potential student behavior [21], it is necessary to build in roadblocks now so that all users are acquainted with the university's standard of data acquisition and reporting. As the demands for data continue to grow, it is essential that users have a sense of standardized procedures built into the expectations for their work practices. While roadblocks are often negatively viewed in user experience design, taking away from a seamless experience [44], these sorts of roadblocks could ensure that users are all following a standardized method of acquiring and interpreting data that has been laid out and approved by database administrators and university decision makers. We propose building into the gateway a feed of recent activity on the landing page. Within this feed, the system will send notifications to users based on recent history with suggestions for better ways of pulling data or quicker ways to find data. Additionally, database administrators can provide a "most used" feature that will either point to or provide access to the most used data requested by users and units.

5.3.6. History, Workflows, and Visualizing Time

One of the most common frustrations of users was the amount of time required to access and analyze data. While some tasks can currently be automated, others require combining data sources or accessing multiple data systems requiring additional time. When preparing data to create visualizations, "reconstructing a repeatable workflow is difficult without a coherent linear history of the operations performed" [29]. We propose incorporating a history of workflow task actions that can be referenced through a visualization of actions and steps. Users should be able to access these workflows, share them with colleagues, and place them in the larger ecosystem of the university data retrieval. Visualizing the time spent on a task can be educational for users' supervisors and database administrators, allowing both groups the opportunity to optimize work for end users. Here, users can quantify the otherwise "invisible work" [43] performed when querying data systems. Additionally, sharing best practices or found shortcuts will facilitate a sense of collaboration. By being able to replicate a colleague's work or understand how data were acquired helps individual users in units understand how other users acquire and use data as well as creating replicable workflows that can reduce the time spent on a given data-related task. Placing the workflow visualization in the larger ecosystem of data use at the university highlights the role end users are playing in data storytelling and provide a context for the work that they do. A visualization links a user's actions to other users' actions, as well as the workflow that leads to users being able to access the data they are using. Each step of this process is visualized and available for users to see the full data universe. This context will also help users to build consensus among end users about the types of data being retrieved and the time users spend gathering data [17].

5.4. Limitations

This study is limited to five midwestern universities in the United States. Institutions in different regions, countries, or cultures may follow different work practices and generate different results. Universities were not chosen based on similarities in admission rates, budget, attrition rate, or graduation rate, so each institution may access and store data in varying manners. Respondents to the study came from a variety of backgrounds, thus making some of the results non-generalizable for larger groups. For instance, further study may need to be performed to gain a deeper understanding of how the end user needs of advisors are impacted versus those of marketers. Additionally, our focus on the user interface design means that there may be additional implementation challenges in the back end that should be investigated in future work (maybe with software engineers).

6. Conclusions

The goal of this research was to identify challenges and opportunities that occur in typical usage scenarios and craft a set of design recommendations and design functionalities that can be implemented. A survey and semi-structured interviews with database administrators and end users highlighted three macro-themes: access, understandability, and use. These themes were used to identify challenge and opportunities for designing such systems. We discussed three broad design recommendations and five finer-grained design functionalities to address challenges and opportunities when designing data access systems for higher education. While this study included research on collaboration while performing data work, it did not include any research on the benefits of collaboration during training, which could facilitate a greater understanding of data and better relationship building overall. This study is limited to five midwestern universities in the United States. Institutions in different regions, countries, or cultures may follow different work practices and generate different results.

It is worth noting, however, that these design recommendations can be incorporated into other data gathering work practices—including discovery, acquisition, retrieval, transformation, and reporting—in higher education and extended into wider business practices, particularly those that collect an abundance of data, but lack the resources to consistently access, analyze, and report on the data. In particular, these industries can include health, construction, transportation, finance, government, and retail. Within health, the data created not only from patient records but also the Internet of Healthy Things can be overwhelming to sort through and categorize as the industry continues to grow. Construction and transportation use logs that could benefit from not only a blockchain-style distributed ledger but also a standardized training system and a way to collaborate both internally and externally. Both the finance industry and many state and local governments are already incorporating smart contracts, but both could profit from creating a tracking system that places the data in the larger ecosystem that each industry resides within. As higher education seeks to create a better customer experience for students, retail is focused on the same goal for both in-person and e-commerce shopping. Within these different realms, it is imperative to harness the wealth of data accessed by users.

Author Contributions: Conceptualization, A.B. and F.C.; methodology, A.B. and F.C.; software, A.B.; validation, A.B. and F.C.; formal analysis, A.B. and F.C.; investigation, A.B.; resources, A.B.; data curation, A.B.; writing—original draft preparation, A.B.; writing—review and editing, A.B. and F.C.; visualization, A.B.; supervision, F.C.; and project administration, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Indiana University (protocol code 1808111834 and date of approval 25 September 2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Keim, D.A.; Mansmann, F.; Schneidewind, J.; Ziegler, H. Challenges in visual data analysis. In Proceedings of the Tenth International Conference on Information Visualisation (IV'06), London, UK, 5–7 July 2006; pp. 9–16.
2. Daniel, B.K. Big data in higher education: The big picture. In *Big Data and Learning Analytics in Higher Education*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 19–28.
3. Khan, S.I.; Hoque, A.S.M.L. Towards development of national health data warehouse for knowledge discovery. In *Intelligent Systems Technologies and Applications*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 413–421.
4. Rattenbury, T.; Hellerstein, J.M.; Heer, J.; Kandel, S.; Carreras, C. *Principles of Data Wrangling: Practical Techniques for Data Preparation*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2017.
5. Fisher, D.; DeLine, R.; Czerwinski, M.; Drucker, S. Interactions with big data analytics. *Interactions* **2012**, *19*, 50–59. [[CrossRef](#)]

6. Voida, A.; Harmon, E.; Al-Ani, B. Homebrew databases: Complexities of everyday information management in nonprofit organizations. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; pp. 915–924.
7. Green, K. The Babel problem with big data in higher ed. Retrieved March 2018, 11, 2019.
8. Carroll, J.M. Human computer interaction-brief intro. In *The Encyclopedia of Human-Computer Interaction*, 2nd ed.; 2013. Available online: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed> (accessed on 23 June 2021).
9. Shneiderman, B.; Plaisant, C.; Cohen, M.S.; Jacobs, S.; Elmqvist, N.; Diakopoulos, N. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*; Pearson: London, UK, 2016.
10. Carneiro, D.; Pimenta, A.; Gonçalves, S.; Neves, J.; Novais, P. Monitoring and improving performance in human-computer interaction. *Concurr. Comput. Pract. Exp.* **2016**, *28*, 1291–1309. [CrossRef]
11. Zdziebko, T.; Sulikowski, P. Monitoring human website interactions for online stores. In *New Contributions in Information Systems and Technologies*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 375–384.
12. Blandford, A.; Furniss, D.; Makri, S. Qualitative HCI research: Going behind the scenes. *Synth. Lect. Hum. Centered Inform.* **2016**, *9*, 1–115. [CrossRef]
13. Victorelli, E.Z.; Dos Reis, J.C.; Hornung, H.; Prado, A.B. Understanding human-data interaction: Literature review and recommendations for design. *Int. J. Hum. Comput. Stud.* **2020**, *134*, 13–32. [CrossRef]
14. Mortier, R.; Haddadi, H.; Henderson, T.; McAuley, D.; Crowcroft, J. Human-Data Interaction: The Human Face of the Data-Driven Society. 2014. Available online: <https://arxiv.org/abs/1412.6159> (accessed on 23 June 2021).
15. Trajkova, M.; Alhakamy, A.; Cafaro, F.; Mallappa, R.; Kankara, S.R. Move your body: Engaging museum visitors with human-data interaction. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, Hawaii, 25–30 April 2020; pp. 1–13.
16. Elmqvist, N. Embodied human-data interaction. In Proceedings of the ACM CHI 2011 Workshop Embodied Interaction: Theory and Practice in HCI, Vancouver, BC, Canada, 7–12 May 2011; Volume 1, pp. 104–107.
17. Boyd, D.; Crawford, K. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Inform. Commun. Soc.* **2012**, *15*, 662–679. [CrossRef]
18. LaValle, S.; Lesser, E.; Shockley, R.; Hopkins, M.S.; Kruschwitz, N. Big data, analytics and the path from insights to value. *MIT Sloan Manag. Rev.* **2011**, *52*, 21–32.
19. Luan, J. Data Mining and Its Applications in Higher Education. *New Dir. Inst. Res.* **2002**, *113*, 17–36. [CrossRef]
20. Picciano, A.G. The evolution of big data and learning analytics in American higher education. *J. Asynchronous Learn. Netw.* **2012**, *16*, 9–20. [CrossRef]
21. Zhu, C.; Zhu, H.; Xiong, H.; Ding, P.; Xie, F. Recruitment market trend analysis with sequential latent variable models. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 383–392.
22. Bergman, O.; Beyth-Marom, R.; Nachmias, R. The project fragmentation problem in personal information management. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montreal, Canada, 24–27 April 2006; pp. 271–274.
23. Bopp, C.; Harmon, E.; Voida, A. Disempowered by data: Nonprofits, social enterprises, and the consequences of data-driven work. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 3608–3619.
24. Pine, K.; Mazmanian, M. Emerging Insights on Building Infrastructure for Data-Driven Transparency and Accountability of Organizations. *iConf. 2015 Proc.* 2015. Available online: <https://www.ideals.illinois.edu/handle/2142/73454> (accessed on 23 June 2021).
25. Le Dantec, C.A.; Edwards, W.K. Across boundaries of influence and accountability: The multiple scales of public sector information systems. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 113–122.
26. Benjamin, L.M.; Voida, A.; Bopp, C. Policy fields, data systems, and the performance of nonprofit human service organizations. *Hum. Serv. Organ. Manag. Leadersh. Gov.* **2018**, *42*, 185–204. [CrossRef]
27. Verma, N.; Voida, A. On being actionable: Mythologies of business intelligence and disconnects in drill downs. In Proceedings of the 19th International Conference on Supporting Group Work, San Francisco, CA, USA, 27 February–2 March 2016; pp. 325–334.
28. Williamson, B. Digital education governance: data visualization, predictive analytics, and ‘real-time’ policy instruments. *J. Educ. Policy* **2016**, *31*, 123–141. [CrossRef]
29. Kandel, S.; Paepcke, A.; Hellerstein, J.M.; Heer, J. Enterprise data analysis and visualization: An interview study. *IEEE Trans. Vis. Comput. Graph.* **2012**, *18*, 2917–2926. [CrossRef]
30. Dur, B.I.U. Data visualization and infographics in visual communication design education at the age of information. *J. Arts Humanit.* **2014**, *3*, 39–50.
31. Few, S. *Information Dashboard Design: The Effective Visual Communication of Data*; O’reilly: Sebastopol, CA, USA, 2006; Volume 2.
32. Chul Kwon, B.; Fisher, B.; Yi, J.S. Visual analytic roadblocks for novice investigators. In Proceedings of the 2011 IEEE Conference on Visual Analytics Science and Technology (VAST), Providence, RI, USA, 23–28 October 2011; pp. 3–11.

33. Sedlmair, M.; Isenberg, P.; Baur, D.; Butz, A. Evaluating information visualization in large companies: Challenges, experiences and recommendations. In Proceedings of the 3rd BELIV'10 Workshop: BEYond Time and Errors: Novel Evaluation Methods for Information Visualization, Atlanta, GA, USA, 10–11 April 2010; pp. 79–86.
34. Kandel, S.; Heer, J.; Plaisant, C.; Kennedy, J.; Van Ham, F.; Riche, N.H.; Weaver, C.; Lee, B.; Brodbeck, D.; Buono, P. Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Inf. Vis.* **2011**, *10*, 271–288. [[CrossRef](#)]
35. Roggema, R. Research by Design: Proposition for a Methodological Approach. *Urban Sci.* **2017**, *1*. [[CrossRef](#)]
36. Fiesler, C.; Brubaker, J.R.; Forte, A.; Guha, S.; McDonald, N.; Muller, M. Qualitative Methods for CSCW: Challenges and Opportunities. In Proceedings of the Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing; Austin, TX, USA, 3–13 November 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 455–460. [[CrossRef](#)]
37. Anderson, R.C.; Reynolds, R.E.; Schallert, D.L.; Goetz, E.T. Frameworks for comprehending discourse. *Am. Educ. Res. J.* **1977**, *14*, 367–381. [[CrossRef](#)]
38. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [[CrossRef](#)]
39. Sandelowski, M. Sample size in qualitative research. *Res. Nurs. Health* **1995**, *18*, 179–183. [[CrossRef](#)]
40. Roberts, R.D.; Beh, H.C.; Stankov, L. Hick's law, competing-task performance, and intelligence. *Intelligence* **1988**, *12*, 111–130. [[CrossRef](#)]
41. Pickett, R.A.; Hamre, W.B. Building portals for higher education. *New Dir. Inst. Res.* **2002**, *2002*, 37–56. [[CrossRef](#)]
42. Parkinson, C.N.; Osborn, R.C. *Parkinson's Law, and Other Studies in Administration*; Houghton Mifflin: Boston, MA, USA, 1957; Volume 24.
43. Hochschild, A.R. *Invisible Labor: Hidden Work in the Contemporary World*; University of California Press: Berkeley, CA, USA, 2016.
44. Beauregard, R.; Corriveau, P. User experience quality: A conceptual framework for goal setting and measurement. In Proceedings of the International Conference on Digital Human Modeling, Beijing, China, 22–27 July 2007; Springer: Berlin/Heidelberg, Germany, 2007; pp. 325–332.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.