

POLICY PERSPECTIVES

Data Collection Efforts at Postsecondary Institutions



Watson Scott Swail, Ed.D.
Maly Fung-Angarita, M.Ed.



The Educational Policy Institute, Inc. (EPI) is a non-partisan research organization dedicated to policy-based research on educational opportunity for all students. EPI is a collective association of researchers and policy analysts from around the world dedicated to the mission of enhancing our knowledge of critical barriers facing students and families throughout the educational pipeline.

The mission of EPI is to expand educational opportunity for low-income and other historically-underrepresented students through high-level research and analysis. By providing educational leaders and policymakers with the information required to make prudent programmatic and policy decisions, we believe that the doors of opportunity can be further opened for all students, resulting in an increase in the number of students prepared for, enrolled in, and completing postsecondary education.

For more information about the Educational Policy Institute, please visit our website:

www.educationalpolicy.org

Policy Perspectives does not necessarily represent the views of the Educational Policy Institute, its officers, or staff.

ABOUT THE AUTHORS

WATSON SCOTT SWAIL, ED.D., is the President and Chief Executive Officer of The Educational Policy Institute, a Virginia-based organization dedicated to policy-based research on educational opportunity for all students. Dr. Swail founded EPI in 2002 to meet a growing need for high-level research on educational opportunity issues. Since that time, Swail has served as the PI on over 70 studies on issues from early childhood reading to postsecondary outcomes for students.

Dr. Swail has a very broad understanding of education. A former middle school teacher, Swail became an expert in curriculum design and teacher professional development. He worked on several school reform initiatives before transitioning into the study of college access and success. Swail is the author of *Retaining Minority Students in Higher Education* (Jossey Bass, 2003), *Finding Superman: Debating the Future of Public Education in America* (Teachers College Press 2012), and *Stop Making Sense* (EPI Press, 2018). He also writes the regular *The Swail Letter on Education* at www.theswailletter.com and associated podcast.

In addition to his research and writing, Dr. Swail has taught educational policy and research at The George Washington University in Washington, DC, where he received his doctorate in educational policy. He earned his Master's of Science from Old Dominion University in Norfolk, Virginia, and Bachelor's of Education from the University of Manitoba, Winnipeg, Manitoba.

MALY FUNG-ANGARITA, M.ED., is a research associate at Westat and a former associate at the Educational Policy Institute. Ms. Fung-Angarita received her B.A. in International Affairs from Lafayette College in Easton, PA and her M.S.Ed in Higher Education Management from the University of Pennsylvania Graduate School of Education, where she served as a graduate assistant for the Office Academic and Student Affairs. There, she played an integral role in the creation of the school's first Student Government Association and Diversity Advisory Board.

Prior to joining EPI, Maly served as a street outreach counselor for youth in crisis at Seton Youth Shelter in Virginia Beach, Virginia. She taught middle school through the Urban Teaching Fellowship in New York City; ESL classes to recent immigrants in Queens, New York; and critical theory at the Institute for Recruitment of Teachers at the Phillips Academy in Andover, Massachusetts. Maly is a native of Venezuela and is fluent in Spanish and Cantonese.

Suggested Citation: Swail, Watson S., and Fung-Angarita, Maly (2018). "Data Collection Efforts at Postsecondary Institutions." *Policy Perspectives*. Virginia Beach, VA: Educational Policy Institute.

Contents

ACKNOWLEDGEMENTS	5
EXECUTIVE SUMMARY	6
INTRODUCTION	8
Background and Rationale	8
The Need for Data.....	10
RESEARCH METHODOLOGY	11
INSTITUTIONAL EFFORTS	13
STATE EFFORTS	21
<i>Florida</i>	22
<i>Texas</i>	22
<i>Virginia</i>	23
<i>North Carolina</i>	24
<i>California</i>	24
<i>New Mexico</i>	25
<i>Arkansas</i>	26
<i>Washington</i>	26
MULTI-STATE EFFORTS	28
<i>NCHEMS Multi-State Data Exchange</i>	28
<i>Achieving the Dream</i>	28
<i>Data Work Group</i>	30
<i>Other Organizational Efforts</i>	30
NATIONAL/REGIONAL EFFORTS	32
<i>Restructuring IPEDS</i>	32
<i>The National Student Clearinghouse</i>	33
<i>Feedback from the Field on a National Student Unit Record Database</i>	34
<i>Voluntary Systems of Accountability</i>	37
BARRIERS TO UNIT RECORD DATA COLLECTION SYSTEMS	39
DISCUSSION	40
REFERENCES	43
APPENDICES	46
Appendix A: Interviews with Expert Panel	47
Appendix B: FERPA	48
Appendix C: Open-Ended Survey Responses	50

Acknowledgements

This publication represents the final report of a study conducted by the Educational Policy Institute (EPI), in association with the AAAS Center for Advancing Science & Engineering Capacity, for the Alfred P. Sloan Foundation in 2008-12. The original report was never made public, and, although the information in this report is a trifle dated, we feel that the information herein remains relevant. We battled with the thought of updating certain data, but then felt that it wouldn't truly represent the report that it was. Thus, we decided to keep mostly as is, with a few exceptions for clarity.

Nevertheless, the information and findings of this study are interesting and potentially important. Our survey of over 800 higher education professionals, as well as in-depth interviews with targeted individuals, gives the study credence. Our hope is that this report will be useful for those in the academic and policy community.

This effort began at a January 2008 meeting hosted by The Alfred P. Sloan Foundation, the Association of Public and Land-grant Universities (ALPU), and the American Society for Engineering Education (ASEE), where the discussion focused on tracking student progress at engineering schools across the United States. Little was known about the outcomes of engineering students and the conclusions of the meeting resulted in this study. EPI, with proficiency in student retention and success, was one of the players in the original conversation and became the lead research agency on this project. Thus, the focus of the report is on the ability to track college students, with special focus on engineering students.

We would like to thank the staff of AAAS's Center for Advancing Science & Engineering Capacity, and Dr. Daryl Chubin in particular. I would be remiss in not acknowledging the effort of Ms. Maly Fung-Angarita, who indeed did some heavy lifting. As well, we thank former Sloan Foundation Senior Advisor Ted Greenwood, who helped make this study possible.

October 26, 2018

Watson Scott Swail, Ed.D.

President & Senior Research Scientist

Educational Policy Institute

Executive Summary

The issue of student retention and graduation from postsecondary institutions has grown in stature over the past decade. While the last 40 years of federal and state policies have focused largely on access to college, there is now a very real interest in not only getting students into college but also helping them earn baccalaureate and other degrees (Swail, Redd, and Perna, 2003). Aggregate data indicate gaps in college completion for students by socio-economic status, educational legacy, and race/ethnicity, but individual student experiences and outcomes cannot be determined from such data sources. In order for institutions, or the country as a whole, to impact student persistence and achievement (as is the focus of the current presidential administration), individual student-level data must be available.

The purpose of this paper is to identify and describe current and recent efforts at the institutional, state, regional, and national levels to collect individual student data for tracking progress in postsecondary education and to provide comprehensive information on promising and usable efforts to collect data. Our hope is that this information assists the Alfred P. Sloan Foundation and the larger education constituencies in envisioning, if not developing, a better way of collecting and assessing student data and informing public policy.

In collaboration with the AAAS Center for Advancing Science & Engineering Capacity with support from the Alfred P. Sloan Foundation, the Educational Policy Institute (EPI) conducted a multi-faceted exploration of the activities, challenges, and successes of data collection at the postsecondary level. Project activities included:

- Reviewing available literature and information on issues and projects related to student unit record (SUR) data collection.
- Interviewing 20 experts from 16 organizations regarding SUR systems and “tracking” student activities and movements. (Appendix A)
- Conducting a purposeful web-based survey of 819 individuals at institutions across the United States to gauge institutional data tracking ability and perceptions of current data issues and policy discussions.
- Corresponding with project managers at several universities and leading organizations working on SUR projects.

We identify and discuss in this report many examples of efforts to collect and track student-level data at the institution, state, and multi-state levels. We found that the predominant view is that a national student unit record system would be a useful, if not critical, tool for public policy. The greatest challenge, from our point of view, remains with interpretations of the Family Educational Rights and Privacy Act of 1974 (FERPA), which frequently prevent effective data sharing. Further, the majority of those interviewed agreed that a major issue in creating a single national student unit record system is establishing a common protocol for how an institution must report its data. There is also concern surrounding the accessibility of the data and accounting for variables that could complicate the data input such as name changes, student transfers between public and private institutions, community and four-year colleges, and across state borders.

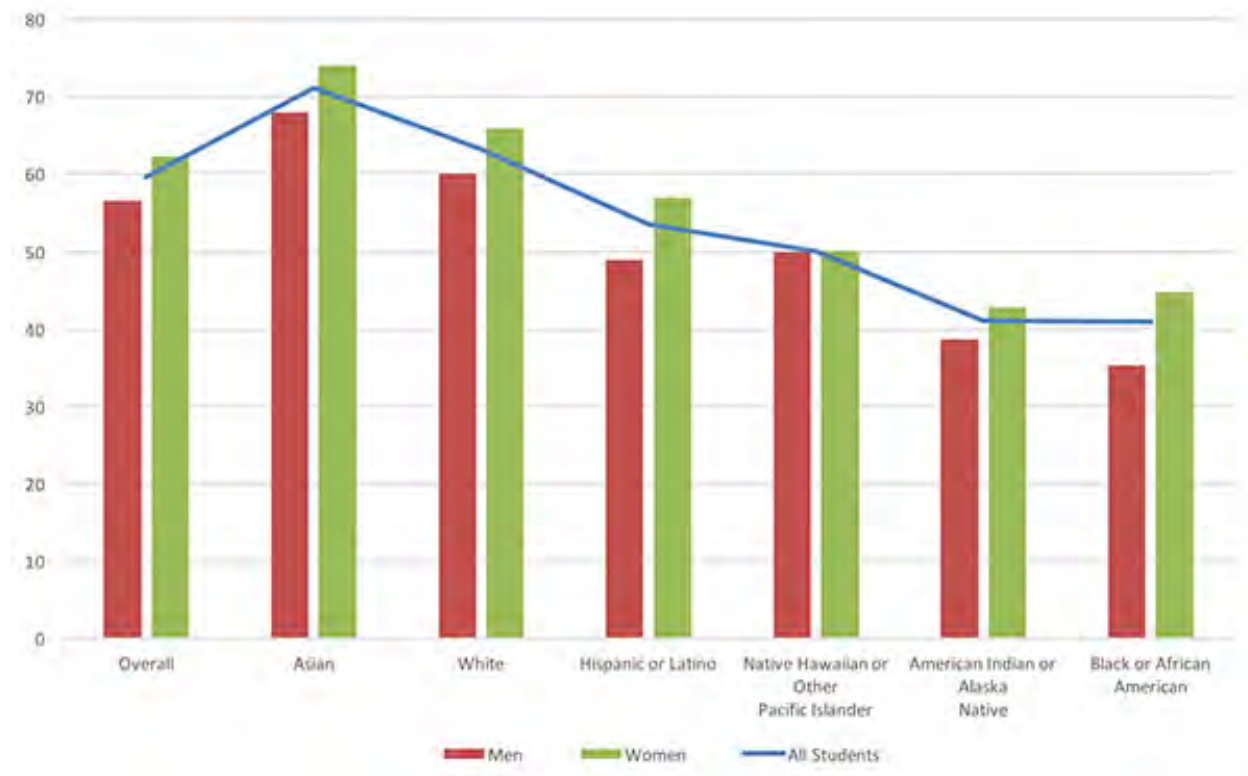
We determined that any national dataset would realistically reside within a voluntary structure. The current voluntary systems demonstrate what works and what does not. If we can build a system where membership is inclusive rather than exclusive and where institutions are noticed for not being involved, then it can fulfill the expectations and outcomes of a federally-mandated system. Current initiatives have paved the way for the development of such a system. The Data Quality Campaign has made significant progress in getting states to understand where their data systems reside with respect to other states and what “quality” really means. Together, the National Association of State Universities and Land-Grant Colleges (APLU) and the American Association of State Colleges and Universities (AASCU) created the Voluntary System of Accountability which has already established practices and agreements that can be used as a foundation for either a larger system or a new system. This paper details these efforts so that others may build upon them.

Introduction

Background and Rationale

The issue of student retention and graduation from postsecondary institutions has grown in stature over the past decade. While the last 40 years of federal and state policies have focused largely on access to college, there is now a very real interest in not only getting students into college, but also helping them earn baccalaureate degrees (Swail, Redd, and Perna, 2003). Cohort data from the freshman class of 2008 finds that, on average, 60 percent of first-time students graduate within six years of beginning their studies. As illustrated in Exhibit 1 below, men tend to have lower graduation rates than women, and Black, Hispanic, and other minority groups graduate at lower rates than White and Asian students.

Exhibit 1. Six-year graduation rates of degree-seeking students at four-year institutions, by gender and race/ethnicity (cohort year 2008)



SOURCE: U.S. Department of Education, National Center for Education Statistics, <https://nces.ed.gov/pubs2015/2015181.pdf>.

According to the National Science Board's 2008 *Science and Engineering Indicators* (2008), students in the agricultural, biological, physical and computer sciences, mathematics, and engineering persist and complete undergraduate programs at about the same rate as students in

the social and behavioral science and other non-STEM¹ students. As Exhibit 2 shows, six years after enrollment in a four-year college or university in 1995-96, around 60 percent of all four groupings have completed a bachelor's degree. Between 13 to 17 percent of students were still enrolled and may have eventually earned a bachelor's degree, and between 18 and 21 percent had not completed any degree and were no longer enrolled. These data are limited (and dated), however, since they do not include those students who may have transferred from one institution's STEM program to another institution. As well, it does not more fully break out "non S&E."

Exhibit 2. Persistence and outcome of postsecondary students beginning four-year colleges or universities in 1995: 2001 (percent)

Major in 1995	Number	Cumulative Persistence Outcome 2001				
		Bachelor's	Associate's or certificate	Still enrolled	No longer enrolled	Missing
All majors	1,369,400	58.0	6.6	14.4	20.8	0.3
Agricultural/biological sciences	115,300	60.8	4.0	16.7	18.2	0.3
Physical/math/computer sciences/engineering	153,600	59.4	7.3	14.1	19.2	0.1
Social and behavioral sciences	82,600	62.4	3.4	14.7	19.1	0.5
Non-S&E	599,000	57.7	7.6	13.2	21.2	0.2
Missing/undeclared	418,900	56.3	6.1	15.5	21.7	0.4

SOURCE: National Center for Education Statistics, 2001 Beginning Postsecondary Students Longitudinal Study, special tabulations (2007). Retrieved June 30, 2009, from <http://www.nsf.gov/statistics/seind08/c2/c2s.htm>

This information on graduation rates, though important, is not sufficient to determine how students make it through the pipeline or inform the process of increasing student retention. Unfortunately, this basically represents our knowledge of student persistence and degree attainment in postsecondary education, and very few analyses accurately describe, in detail, the pathways of students to and through undergraduate education. In order for institutions, or the country as a whole, to impact student persistence and achievement, more precise estimates of persistence and greater knowledge on *why* students do not graduate is needed. With the exception of the annual Integrated Postsecondary Education Data Survey (IPEDS) survey from the US Department of Education, there are few attempts to collect and disseminate useful data on student persistence and graduation. In addition, there are virtually no regular, *national* attempts to collect "student-unit data"² from institutions. (*It should be noted that the National Student Clearinghouse is currently (2018) filling this sector, which is discussed later. However, at the original time of this report we were primarily focused on federal and state issues.*)

While most, if not all, institutions collect student record data for internal purposes, very few share information with other institutions or state agencies. Currently, 40 states maintain student record databases, as do many individual institutions. However, these are neither published nor used on a broad scale. In particular, the lack of data from private institutions presents a

¹ Definitions of STEM vary and include a wide range of disciplines. In this paper the term STEM includes mathematics, physical and biological/agricultural sciences, engineering, engineering technology, and computer/information sciences. The terms STEM and Science and Engineering (S&E) are used interchangeably.

² Student-unit data are data collected about specific students on campus, and not a sample. Student-unit data provide the most accurate and useful information since they are collected at the smallest level of disaggregation.

challenge for researchers to obtain any type of unit record information on students. This is the case because private colleges and universities are autonomous entities and often have the ability to opt out of requests from the state governments beyond federal Title IV regulations. On the other hand, public institutions depend on state funding and therefore must comply with state requests for data submission. Although the percentage of students enrolled in private institutions is relatively small compared to their public counterparts, their omission from studies precludes a comprehensive understanding of the postsecondary student pipeline in the United States.

The Need for Data

There is little published data regarding student enrollment, retention, and graduation rates *within* college programs generally, and STEM fields specifically, which has made it difficult to conduct analysis on how programs and institutions can improve. While current data collection efforts provide information on student persistence by field of study, questions such as why and how many students leave their field or postsecondary education remain unanswered. In the case of STEM specifically, the majority of studies investigating persistence in science and engineering have focused on engineering students (e.g., see Chubin et al., 2005). This may be because enrollment and tracking of engineering majors is more commonly conducted than of students in other STEM majors. Even when students are tracked by a department however, there may be gaps in the data, which is further discussed in the section on Institutional Efforts.

In a related problem, institutions are very provincial in what they wish to collect and report in the public arena. Higher education in the US is highly competitive, such that institutions often resist providing data that can be used for comparison with peer institutions. In particular, private colleges and universities who highly uphold their institutional autonomy are hesitant to share their data. Certainly, institutions in the United States comply with IPEDS requests, as well as with the Common Data Set³ (CDS), which is jointly sponsored by US News and World Report, The College Board, and Petersons. Beyond these collection strategies however, there is little to increase our knowledge of student retention and persistence in higher education.

³ <http://www.commondataset.org/>

Research Methodology

The controversial topic of creating a national student unit record data system has recently resurfaced (2017) as members of congress introduced a bipartisan bill to overturn the ban on creating one such system.⁴ While proposals to repeal the ban have been circulated for years with little success, the recent senate and house bills have revived the conversation arguing that students and families need better information to help them decide on choosing which college to attend. Similarly, proponents of the bill affirm that a system that houses student-unit record data will greatly inform colleges and universities about the paths their students have taken after leaving their institution. They argue that understanding student trajectories can inform institutions about strategies that work to help students succeed. Meanwhile, major opponents of a student unit record system cite privacy concerns and contend that the sharing of this data is a violation of the Family Educational Rights and Privacy Act of 1974 (FERPA).⁵

To shed light on these issues, the Educational Policy Institute (EPI) conducted a broad study of data systems in higher education in a multi-pronged approach. First, EPI reviewed available literature and information on issues and projects related to student unit record (SUR) data collection.

Second, EPI interviewed 20 experts from 16 organizations (See Appendix A) in October and November 2008 regarding SUR systems and “tracking” student activities and movements. The topics of these discussions with experts included the National Student Clearinghouse (NSC), IPEDS, voluntary data systems, multi-state consortia, and the burden of creating a national SUR system. These experts were primarily selected from the original attendance list at a January 2008 meeting hosted by The Alfred P. Sloan Foundation, the National Association of State Universities and Land-Grant Colleges (AASCU), and the American Association of State Colleges and Universities (AACU) in Washington, DC. In addition, individuals were added based on the suggestions from the original group.

Third, EPI conducted a purposeful web-based survey of higher education professionals in 2009 in an effort to evaluate efforts at the institutional, state, regional, and national levels to collect individual student data for tracking their progress in postsecondary education. The purpose of the survey was to examine the extent to which different institutions and organizations track student progress throughout their post-secondary careers using student-level data. The survey also explored respondents’ perceptions of the feasibility of developing a national student unit-record (SUR) system and examined the major challenges and concerns in using data to track students. The responses captured in the survey underscore arguments that are as relevant to the conversation today and they were nearly 10 years ago.

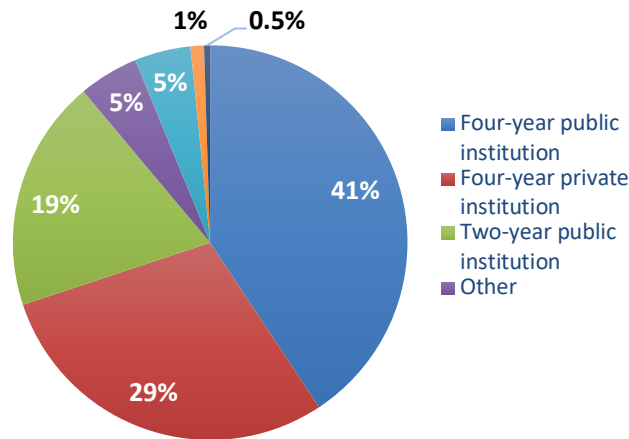
The web-based survey was distributed via email and recipients were asked to complete the survey on a voluntary basis. A total of 821 participants completed the survey and included representatives from four-year public institutions (41 percent), four-year private institutions (29

⁴ <http://www.chronicle.com/article/Here-s-How-a-Student-Unit/240165>

⁵ For more information on FERPA, see Appendix B.

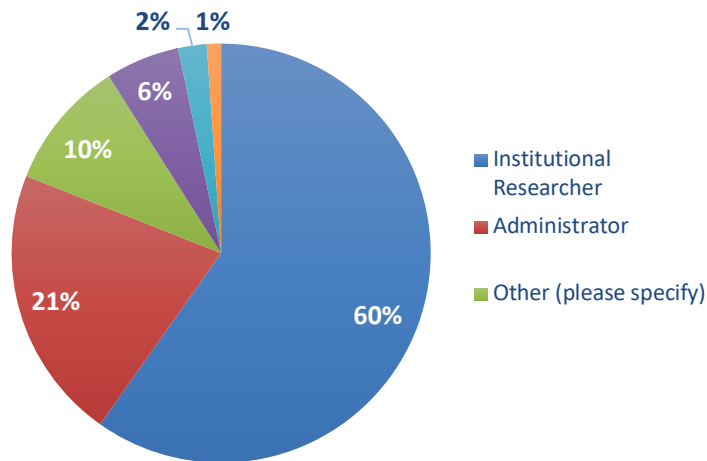
percent), two-year public institutions (19 percent), non-profit organizations (5 percent), proprietary or career colleges (1 percent), and two-year private institutions (0.5 percent).

Exhibit 3. Respondents' Institutions/Organizations



The vast majority of respondents consisted of institutional researchers (60 percent), followed by administrators (21 percent), educators (6 percent), non-institutional researchers (2 percent) and policymakers (1 percent). Approximately 10 percent of respondents identified as “Other” and included graduate students, registrar’s office staff members, faculty members, and data analysts.

Exhibit 4. Respondents' Roles



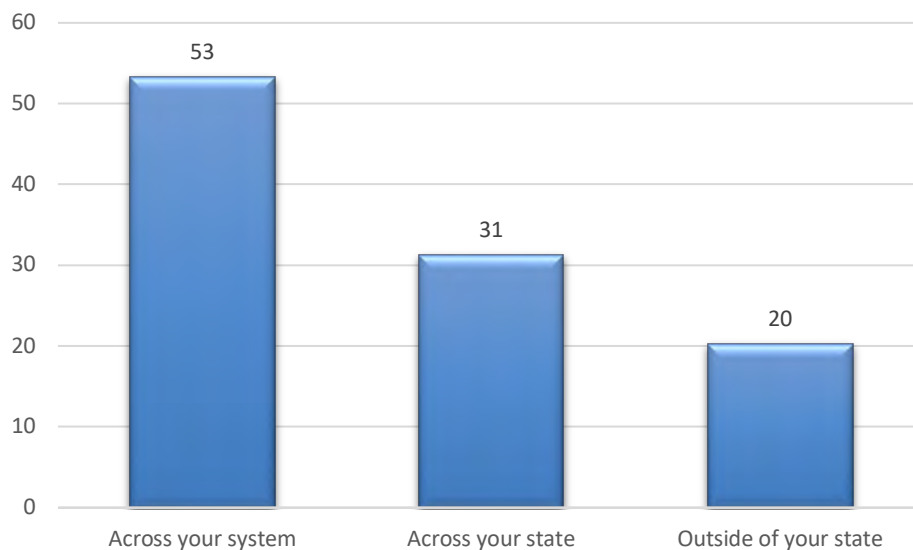
Institutional Efforts

This section presents current efforts by institutions to participate in data collection projects, whether individually or as members of a consortium. As previously described, EPI conducted an online survey of to gauge institutional data tracking ability as well as perceptions on current data issues and policy discussions. After the survey findings are discussed, we present findings from phone interviews with project managers at three universities engaged in collecting SUR data for students in the STEM fields. One private organization (American Society for Engineering Education, ASEE) provided written responses to questions regarding its projects. Finally, the Consortium for Student Retention Data Exchange website was consulted to attain information regarding institution-level data efforts.

Survey Findings Regarding Institutional Student Data Tracking

Fifty percent of the respondents indicated that their institution has the ability to track students across their university system. Of this number, over half of survey respondents (53 percent) indicated that their institution has the ability to track student progress across their educational system. Thirty-one percent have the ability to track student academic and educational progress across the state, and 20 percent of respondents indicated having the ability to track students outside of their state.

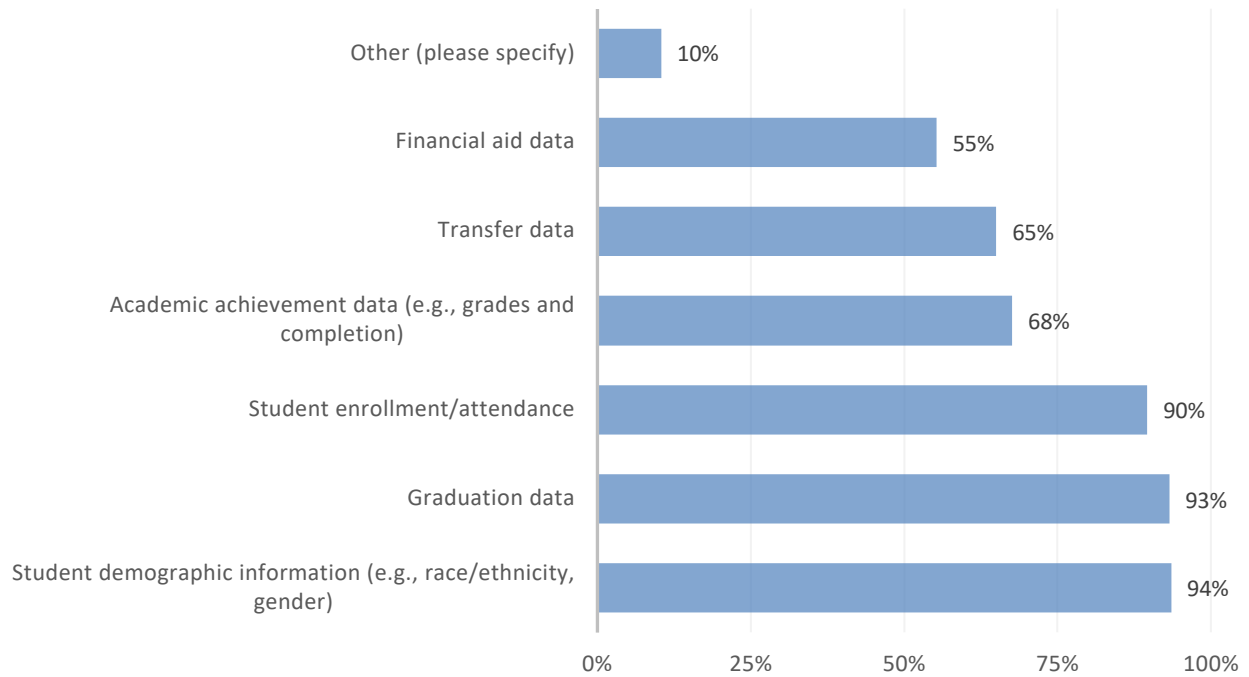
Exhibit 5. Ways that Institutions Can Track Students Educational Progress



The majority of respondents (87 percent) indicated their institution has the ability to track students by major. In contrast, 72 percent indicated that their institution has the ability to track majors across different disciplines to capture, for example, students who switch majors.

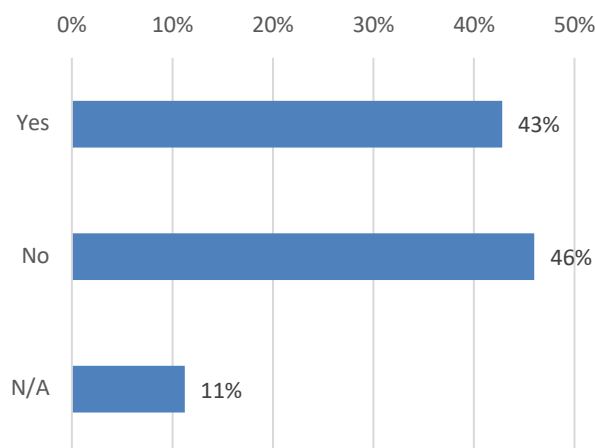
At the state level, nearly half of respondents (47 percent) indicated that their institution/organization participates in a larger, state-based student unit-record data system. Among the data available in the state-based SUR data system, respondents reported student demographic information (94 percent), graduation data (93 percent), student enrollment/attendance data (90 percent), student academic achievement data (68 percent), transfer data (65 percent), financial aid data (55 percent), and other (10 percent).

Exhibit 6. Data Elements Available in the State Data



Outside of state-based data systems, nearly two-thirds (64 percent) of respondents reported using the [National Student Clearinghouse](#) to track students once they leave the institution. Only 14 percent of respondents use state income databases to track employment and income of graduates (via the Wage Record Information System, or WRIS). At the time of the survey, less than half of the participants (43 percent) believed that a national student unit-record system would be viable to implement. Another 46 percent did not believe this to be a possibility at the time and 11 percent of participants selected “Not Applicable” as their answer.

Exhibit 7. Survey Response Regarding Concept of a National Student Unit-Record System



Regarding the biggest challenges and concerns in tracking students in their systems, 430 respondents provided open-ended feedback, from which common themes emerged. For additional details and respondent quotes, please see Appendix C.

Of those that felt that a national unit-record system would not be viable, their responses cited privacy concerns, logistical burden of implementing such system, lack of institutional buy-in, compliance among institutions, and the added cost associated with reporting requirements. Some responses indicated that for such system to be useful, institutions would need to be able to access the data and use it for reporting and comparison purposes. In addition, some participants highlighted the need for better linkages with K-12 and employment data in order for the SUR system to produce useful data.

Among participants who indicated that the SUR system is viable, many suggested that while the technology and the political climate for such effort was favorable, there are other conditions that need to be in place first such as confidentiality and data security concerns, consistent data definitions and data layouts, and a standard ID for students, other than the use of Social Security numbers. Some participants expressed that, while a national system could be viable, it would be difficult to implement and would take time to come to develop. One participant pointed that, even if viable, it was uncertain whether the SUR system was desirable. Participants in favor of a SUR system pointed to the benefits of having access to this type of information. One respondent indicated that “the ability to track student progression through their academic career into their post-academic professional endeavors would be beneficial to all institutions.” Moreover, a national SUR system will generate useful information about student progress, effectiveness of institutional efforts and initiatives, and “demonstrate the ‘true’ state of education.”

The final question asked participants to describe their biggest challenges or concerns with tracking students in their system. The most frequently cited challenge was the ability to track students within an institution, across departmental programs, across their state, or post-graduation. Survey participants pointed to the lack information about where students come from (in terms of having access to better K12 data) and where they go when they dropout, transfer, or graduate. Data quality was another frequently cited challenge. Comments included questionable accuracy of data, errors and inconsistencies in data entry, and upkeep and maintenance of student records.

Other frequently cited challenges included privacy concerns, having appropriate resources (time, money and staff) to complete the required and desired data analyses and reports, and having access to data. Particularly, respondents expressed that large amounts of student data are reported to state systems and the National Student Clearinghouse with very little access to useful data in return. A handful of respondents mentioned the need for greater understanding of student success and data elements that capture students’ academic goals. Especially at community colleges, where student trajectories are not linear, institutions want to better understand at what point students achieve their goals and whether that achievement is defined as degree attainment, successful completion of specific coursework, or some other potential outcomes.

The results from this survey echo the sentiments across a wide spectrum of stakeholders. Certainly, the need and interest for greater transparency and accountability is growing. But many questions still remain, including on how to achieve such significant endeavor without compromising students' privacy and how to ensure that all institutions comply. For now, the introduction of the bills in the house of representatives and senate have reignited the discussion by getting the conversation started.

A summary of themes related to challenges and concerns included:

- Privacy and confidentiality issues, ranging from FERPA to data security to proper use of data;
- Logistical burden of collecting, analyzing, and reporting data – likely with no additional resources;
- Staff training and infrastructure;
- Defining and tracking student outcomes such as leaving, dropping out, graduating from another institution, changing majors within institution, etc.;
- Issues of data sharing and compatibility, whether between campus departments/units, from other colleges, from K-12 systems, state coordinating boards, etc.;
- Data quality and availability, such as using self-reported student level data in the absence of other data sources, data entry errors in existing data sources, duplicative student records, and lack of accountability for data reliability;
- General lack of necessary resources to support and maintain data collection and tracking mechanisms;
- Inability to use available data in meaningful ways due to data structure or fields collected.

Those who did not have significant challenges or concerns include:

- Overall success in querying data and conducting longitudinal analysis, particularly in developmental coursework and subsequent courses;
- Small size of student body makes tracking possible;
- Effective collaboration of institutional research offices in system schools;
- Availability of NSC and alumni databases to track students after they leave campus.

Feedback from Five Institutions on Current Data Collection Efforts and Projects

In a separate effort, EPI conducted phone interviews with project managers at three universities collecting SUR data for students in the STEM fields. One private organization (American Society for Engineering Education, ASEE) provided written responses to questions regarding its

projects. Finally, the Consortium for Student Retention Data Exchange website was consulted to attain information regarding data efforts. Information regarding these five sites/organizations is provided on the following pages.

Swarthmore College

The Sloan Foundation awarded a grant to Swarthmore College to examine the reasons for low retention of certain groups of students in science and engineering. Under the direction of Lynn Molter, Professor and Chair of the Engineering Department at Swarthmore, the project collected SUR data from a group of 22 institutions ranging from large public universities to small private liberal arts colleges.⁶ The data collected include information from college applications, anticipated major, AP scores, and demographics. Rather than following a cohort of students, the study followed every student in all 22 participating institutions on an annual basis for three years. Particular attention was paid to students in STEM fields to understand their migration to and from gateway science courses, the grades students obtain in those courses, and whether or not students enter a specific discipline having not previously anticipated it. This project is currently in its initial stage and follows a two-year pilot study where approximately 30,000 student records from the graduating classes of 2002 and 2006 of 11 institutions were collected.

University of Washington

In a six-year study funded by the Sloan Foundation, the University of Washington (UW) conducted a longitudinal study of undergraduate women pursuing degrees in science or engineering. The study was led by Professor Suzanne Brainard and results were published in the *Journal of Engineering Education*. One of the goals of this study was to obtain more accurate information on the retention of females pursuing science and engineering degrees at the institution. At the time of the study, very few institutions were tracking individual students throughout their postsecondary education (Brainard & Carlin, 1998).

At UW, students interested in majoring in engineering are only tracked once they have been accepted into one of the 10 engineering departments, which typically occurs during students' junior year and after completing pre-science (freshman year) and pre-engineering requirements (sophomore year). Tracking students after the junior year misses critical attrition points during the first two years of a student's undergraduate career and provides an inflated and inaccurate retention rate (Brainard & Carlin, 1998). By looking only at the graduation rates of those who are accepted into the engineering departments, institutions fail to account for students who switch out of science and engineering majors during their first two years, and those students who apply into the department but do not get accepted. This is further complicated by the fact that every school has its own guidelines and requirements as to when a student must declare a major, leading to variation in the timeline in which a student is tracked.

Thus, in an effort to better understand students' pathways to science and engineering, this study followed six cohorts of female students (beginning in 1991) pursuing science and engineering fields at the university and tracked their progress and major aspirations individually through an Annual Freshman Interest Survey, initial and follow-up interviews, and follow-up questionnaires for sophomores, juniors, and seniors. The Annual Freshman

⁶ Personal communication with Lynn Molter.

Interest Survey was mailed at the beginning of the year to all incoming female freshmen to determine how many were interested in majoring in science or engineering. After the survey was administered, initial and follow-up interviews were conducted in person or over the phone at different intervals throughout the year. The study utilized a tracking system to measure retention rates, monitor students' participation in activities, and analyze data each year. In addition to tracking progress, the study also examined intervention strategies offered through the Women in Engineering (WIE) Undergraduate Retention Program at UW, which included personal contact to each student throughout her academic career at the institution. The first two years were targeted in particular, as this time period is when students are most likely to change fields.

Through the surveys and interviews, the study identified a number of common factors (e.g., enjoyment of coursework, career activities, advising, and acceptance into the program) influencing students' decision to persist in science or engineering. The authors found that most women who switch out of science and engineering do so during their sophomore year. Among the most commonly reported reasons for not persisting in science and engineering was lack of interest and discouragement from low grades. According to the National Science Board (2007), approximately one-third of first-year students (30 percent) require remedial science and math classes because of deficits in academic preparation. Researchers have indicated that the major reason students majoring in STEM fields tend to struggle is because high schools haven't prepared them well enough for college. This may also be the case for the women studied at UW. Results from the study led to institutional and departmental changes (e.g., beginning the engineering curriculum in sophomore rather than junior year) to continue to improve the retention and completion rates of women in STEM fields.

In a conversation with Dr. Brainard,⁷ the issue of tracking students individually throughout their undergraduate studies was discussed. She argued that "it is incredibly expensive to do the individualized tracking, but that's really the only clean way that you're going to get retention rates and be able to see what's happening with persistence." Dr. Brainard added that at large institutions like UW, individual tracking does not occur. Another challenge in following students by major as they enroll at a specific institution is that admission forms do not require students to indicate an intended major, providing it only as an optional question. Currently, UW is among the 22 institutions participating in Dr. Lynn Molter's study at Swarthmore College.

Missouri University of Science & Technology

Missouri University of Science & Technology has a well-established system for tracking student progress from the moment students set foot on campus. Upon admission to the institution, students who express interest in engineering are assigned a Classification of Instructional Programs (CIP) code by the Registrar's Office as "freshman engineering."⁸ Students who are interested in a particular engineering program are coded as well. For example, a student with interest in electrical engineering is coded as freshman engineering with an emphasis area in

⁷ Personal communication with Suzanne Brainard.

⁸ Personal communication with Laura Stoll, Registrar at the Missouri University of Science and Technology.

electrical engineering. The data collected are used for reporting by different departments at the institution.

For over 20 years, the university has followed its engineering students through the Freshman Engineering Program where entering freshmen interested in studying engineering are admitted. This program provides advising to enhance academic success, through which students are paired up a member of the engineering faculty who serves as a source of information and advice during students' first year on campus. These advisors assist students in selecting courses, making academic decisions, and identifying campus resources. Ultimately, the Freshman Engineering Program provides students with information about different engineering fields, opportunities for professional presentations, visits to engineering departments, and meetings with the advisor, so that students can make informed decisions regarding their major and career path.

During the first two or three semesters on campus, students take a set of courses that are required by all engineering departments at the institution. In addition, students have the option to select a major in a particular engineering field prior to completing the freshmen engineering requirements. However, upon completion of freshmen engineering requirements, students formally apply for admission to one of the 15 different undergraduate engineering programs.

The Office of Enrollment Management coordinates the University Retention Committee, which uses collected data to develop strategies for increasing student retention. Since 2002, Missouri S&T has implemented several strategies to better understand and improve retention rates, including surveying first-time, full-time students who do not matriculate the following fall and conducting follow-up phone interviews. Thirteen percent of first-time, full-time, degree seeking students who enrolled in the fall of 2007 did not return the following fall. Of this group, 60 percent participated in a phone survey, and nearly 90 percent of respondents mentioned transferring to another institution. Other reasons for not returning to Missouri S&T included not wanting to pursue the engineering major, financial reasons, and other unavoidable circumstances (Missouri S&T, 2009). Most recently, Missouri S&T began surveying second-year students.

American Society for Engineering Education (ASEE)

The American Society for Engineering Education (ASEE) is conducting a pilot study that asks schools to provide data on engineering students over an eight-year period. Each year's cohort of first-time, full-time freshmen will be tracked for their retention from year to year, including the number who graduate in 4 to 8 years. Therefore, they will have data on eight different cohorts, from 2001 through 2008. Obviously, many of these students have not graduated yet, since the most recent cohorts are still in their first few years of college. ASEE will also track transfer student retention separately and examine retention by race and gender groups, when possible. Part of the goal of the pilot study is to see how much data schools can provide. ASEE will not track what happens to students when they leave engineering (i.e., move to another STEM field). In addition, ASEE reports that very few students transfer into engineering from other fields and that most transfers are from two- and four-year schools where the student was already on an engineering track.

Consortium for Student Retention Data Exchange

The Consortium for Student Retention Data Exchange (CSRDE) at the University of Oklahoma is a consortium of two-year and four-year institutions committed to achieving student success through data sharing. It was established in 1994 as a collaboration between Big Eight Data Exchange (BEDE), Big Ten, and Southern University Group (SUG) institutions. Faced with the need to improve student retention and graduation rates, leaders of the BEDE, Big Ten, and SUG institutions were interested in creating a mechanism to collect and compare retention data with peer institutions. The CSRDE had an initial membership of 163 institutions (Smith, 2002) and today has over 600 members.

The CSRDE collects data from member institutions through the annual Retention Survey and tracks the information by cohort. The cohorts at four-year institutions consist of first-time, full-time, baccalaureate degree-seeking freshman. At two-year institutions, the survey collects data from two different cohorts: first-time, full-time students and first-time, part-time students. In addition to the Retention Survey, four-year institutions have the option of participating in three additional surveys: STEM survey sections II and III, and the Community College Transfer Survey. Section II of the STEM survey collects information on institution-wide retention and graduation of first-time, full-time freshmen in STEM while section III tracks the retention and graduation of the same cohort by specific discipline. The Community College Transfer Survey follows the retention and graduation of cohorts (both full- and part-time) of community college transfers into four-year institutions who have earned 30 or more credit hours prior to transferring.

Each year, the CSRDE administers the survey to participating institutions and provides them with a set of reports comparing aggregate retention and graduation rates of institutions with similar characteristics. In addition to a Retention Report, institutions also receive a customized Peer Report containing benchmark data for a self-selected peer group of 5 to 20 institutions. Colleges and universities that contribute data on STEM and community college transfer students also receive a peer report with those particular data (CSRDE, 2010). Because the consortium collects data on multiple student cohorts over time, institutions can utilize reports for ongoing assessments. Moreover, reports are distributed within three months after institutions submit the Retention Survey.

Having timely access to data is essential to making adequate and informed policy decisions. Therefore, due dates for submitting surveys are the same for all institutions, and consortium members have ready access to data from peer institutions by requesting customized reports.

State Efforts

In 2002, Lumina Foundation funded a project by the National Center for Higher Education Management Systems (NCHEMS) which found that most states have data systems in place and gather much of the information that could be helpful when studying student patterns.

Following the growing attention on establishing a national data system using existing state databases, Ewell and Boeke (2007) conducted a second survey updating the previous 50 state inventory. This second study found 47 operational SUR databases across 40 states. Since then, two more states have established SUR databases (National Center for Higher Education Management Systems [NCHEMS], 2008). Approximately 81 percent of the nation's students are tracked through one of these SUR systems.

Of all state databases, only a few are less than 10 years old (Ewell & Boeke, 2007). As a result, the majority of databases have enough data to longitudinally track several cohorts of students. Moreover, all existing SUR databases collect a core combination of data elements, including student demographics, enrollment, program of study, and degree awarded (Ewell & Boeke, 2007). Approximately half of the databases have worked to link postsecondary data to external sources, such as high school data systems. Twenty-three of the systems are linked to unemployment records to track the employment patterns of students (Ewell & Boeke, 2007).

States are actively seeking ways to create comprehensive SUR systems as interest in developing capable K-12 databases increased in the 2000s due to accountability requirements mandated by No Child Left Behind (NCLB). This growing attention by state officials facilitated the creation of the Data Quality Campaign (DQC), a national collaboration geared towards helping states improve student achievement through high-quality data. One of DQC's ten Essential Elements⁹ for a robust longitudinal data system includes the ability to match K-12 with postsecondary education student data. Every year, DQC distributes a survey to all 50 states to assess progress toward the 10 Essential Elements (DQC).

The Institute of Education Sciences (IES) also initiated efforts to assist states in the creation of data repositories. In May 2010, IES awarded grants to 20 state education departments¹⁰ that supported the design and implementation of statewide longitudinal data systems linking data from early childhood to career, including developing the ability to match teacher data to student outcomes (Institute of Education Sciences [IES], 2010). The grants are funded through the American Recovery and Reinvestment Act (ARRA) of 2009 and the total value of the three-year grants range from \$5.1 million to \$19.7 million. Additional grants have been provided as recent as 2015.¹¹

As states continue to enhance their data collection efforts, issues of student privacy and data quality remain a concern. Each state has interpreted Family and Educational Rights and Privacy

⁹ The 10 elements include: 1) Statewide Student Identifier; 2) Student-Level Enrollment Data; 3) Student-Level Test Data; 4) Information on Untested Students; 5) Statewide Teacher Identifier with a Teacher-Student Match; 6) Student-Level Course Completion (Transcript) Data; 7) Student-Level SAT, ACT, and Advanced Placement Exam Data; 8) Student-Level Graduation and Dropout Data; 9) Ability to Match Student-Level P-12 and Higher Education Data; and 10) A State Data Audit System.

¹⁰ For a list of all 20 states, see http://nces.ed.gov/programs/slds/fy09arra_announcement.asp

¹¹ https://nces.ed.gov/programs/slds/grant_information.asp.

Act (FERPA) differently and addressed challenges in their own way. Below, we describe the efforts of eight states at varying stages of development and their progress towards linking SUR data across educational and workforce systems.

Florida

Florida's data sharing system is perhaps the most comprehensive among all states (Mills, 2005). Legislation created Florida's Data Warehouse (FDW) system, but its success can be attributed to two decades of positive interactions between institutions and the state. Institutions give the state student-level data, and in turn, receive feedback in the form of state reports (Mills, 2005). State appropriations are allotted based on data-driven decision-making, which increases the willingness of institutions to cooperate. Florida's system began as a means of assuring that technical programs were compliant with state policy. The data collection was expanded to include independent and private institutions, and over time has grown into a comprehensive data warehouse serving kindergarten through higher education (Mills, 2005). Florida is one of four states to include all private institutions in its data system (Ewell & Boeke, 2007).

The FDW integrates existing data collected from multiple sources available in the state and provides a single repository of data for students in the K-20 (i.e., kindergarten through workforce) education system. It collects information on student enrollment and demographics, educational curriculum, educational institutions, and staff. According to the FDW website, data course or CIP codes are reported at the aggregate level, but not at the student level. In most cases, comprehensive data are available from 1995 onwards. Although social security numbers are used to connect this information about students, any identifiable information is removed to maintain confidentiality under these circumstances. This student-centered longitudinal design of the FDW allows state researchers to integrate data from varied levels of education and perform trend analyses.

FDW has assisted legislators and policymakers in data-driven decisions regarding education in Florida, making available valuable data such as high school course taking (e.g., Chemistry, Biology, Algebra I, etc.). Recent developments to the data warehouse include \$1.58 million from the Department of Education to incorporate financial data into its database (2010). The Florida Department of Education also created a portal linking FDW data to Florida Education and Training Placement Information Program (FETPIP) data, which tracks students who have graduated, exited, or completed a public education or training program within the state. The portal provides a single pathway to information from both databases, known as the Integrated Education Data Systems. As of 2005, Florida already had 5 out of 10 Essential Elements in place, according to the DQC. Currently, they have all 10 in place and 5 out of 10 state actions.

Texas

Texas is also one of the leading states in collecting and reporting education data. In 1999, the Texas Education Agency (TEA) developed a data warehouse for P-12 (i.e., pre-kindergarten through 12th grade) to provide integrated access to public education data. In 2001, the 77th Texas Legislature funded a multi-agency initiative to create an integrated data repository that would "provide a comprehensive view of public education in Texas" (Texas Education Agency [TEA], 2010, p. 5). The project, known as Texas P-16 Public Education Information Resource (TPEIR), is

managed by TEA and Texas Higher Education Coordinating Board (THECB), and integrates data stored in separate databases at each agency (TEA, 2010). The two data repositories were created to protect student privacy as mandated by FERPA. One database contains aggregated data approved for public release while the other contains confidential, student education data and is available only to authorized individuals. Texas believes, as a result of FERPA, that once SUR data are collected, it should not be re-distributed (Mills, 2005) and so they have developed a website to provide public access to integrated education data from multiple agencies (TEA, 2010).

In Texas, money for education is funded on a per-student basis and routed through the THECB. Therefore, THECB argues that student-level data are necessary, and all public institutions must report data to receive funding. Independent schools receive funding through the THECB and must submit data to the system to receive this funding. As a result, Texas has a comprehensive SUR system for private or independent school students, which is rare among states (Mills, 2005). THECB asserts that independent institutions do not mind submitting data to the system because of the benefits that they receive for sharing. Responsibility for data is assumed by the THECB, which limits the liability of independent schools who share data and schools benefit from receiving specific outcome information about their students, such as their employment patterns (Mills, 2005). DQC has rated Texas as having 9 out of 10 Essential Elements in place since 2005 and currently has completed 8 out of 10 action steps.

Virginia

Virginia's SUR database at the postsecondary level has been in place since 1992 and involves all public colleges and universities and nearly all private institutions. However, it is not currently linked to the K-12 system. At the K-12 level, Virginia has been collecting data through the Education Information Management System (EIMS) to meet federal and state accountability requirements. Virginia's Department of Education is also working towards achieving the 10 Essential Elements of the DQC (Neugent, 2006). Future developments of EIMS include adopting interoperability data transmission standards and implementing electronic transcripts to link student information from P-12 to community colleges and other higher education data systems (DQC, 2006; Neugent, 2006).

Similar to Texas, Virginia has a student directory with student names, enrollment, and institutions attended that is maintained by the State Council for Higher Education in Virginia (SCHEV). Community colleges may submit names of students to SCHEV that they believe have transferred to four-year institutions within the Commonwealth. SCHEV, in turn, provides these community colleges with data about the students who are enrolled in other institutions, which programs they are studying, and which degrees they receive (Mills, 2005). These data can be used by state researchers to determine persistence and completion for students who transfer within the Commonwealth of Virginia.

Grant funding is controlled by SCHEV and linked to data reporting. For this reason, independent schools willingly share SUR data with SCHEV. In turn, SCHEV provides services that are beneficial to independent institutions, such as submissions to IPEDS (Mills, 2005). Virginia has also fostered an environment of statewide data sharing between institutions and

agencies. This environment has enabled Virginia to take positive steps towards a comprehensive SUR system.

To maintain confidentiality and protect student information, the use of social security numbers as identifiers is limited by SCHEV to the fullest extent possible. Unless a record is required by law to contain social security numbers, they are not used (Mills, 2005). SCHEV intends to put a system into place that will give each student a unique identifying number which may be used in K-12 education and be linked to postsecondary records. Steps are also being taken to link postsecondary records with employment records through SCHEV and the state employment commission (Mills, 2005). As of 2005, Virginia had 6 out of 10 Essential Elements in place and currently have eight in place. They are rated as completing 5 out of 10 state actions.

North Carolina

According to the DQC, North Carolina has 8 out of 10 Essential Elements for establishing a longitudinal data system, including the ability to match student data between P-12 records, public colleges and universities, and the state's workforce data systems. Even though North Carolina's P-20 longitudinal data system is mandated by state policy, it currently has not created a data repository to integrate data from different agencies. In other words, though the data are collected, the communication between agencies is lacking. Progress on state actions have been slow, and only 4 out of 10 actions set by the DQC have been met. Therefore, North Carolina has not been sharing SUR data to the same extent as some other states.

The K-12 system, two-year public institutions, and four-year public institutions engage in very little communication with one another and are considered separate entities. For example, community colleges in North Carolina collect and share data with one another, but data from the University of North Carolina (UNC) system is not available to community colleges (Mills, 2005), because of UNC's interpretation of FERPA. Community colleges may receive aggregate data from four-year schools about transfer students, but no individual data are made available. Within the community college system, sizeable amounts of data are collected on demographics, courses studied, level of performance, degrees awarded, and credit hours (Mills, 2005).

While data sharing efforts in the North Carolina system are not as strong as they could be, efforts to link postsecondary data to employment records are very effective. The North Carolina Community College System has been working with the Department of Labor since 1991 to match student records and unemployment insurance wage records (Mills, 2005). This collaboration was approved by the state Attorney General and has yielded a strong relationship between the systems.

California

California lacks a comprehensive system for tracking student progress from K-16. Similar to North Carolina, California's education system is divided into four different segments: K-12 public schools, two-year public institutions, the California State University system, and the University of California system (Vernez, Krop, Vuollo, & Hansen, 2008). Each of these agencies has student data systems in place, yet given the lack of communication amongst them, the state is unable to determine the overall level of attrition at all stages of education, the number of

students taking college preparatory work in high school who pursue higher education, or which K-12 courses need attention based on college performance (Vernez, et. al, 2008).

In addition to these four databases, there are two other databases currently in use: the California Post-Secondary Education Commission, which partially integrates student data from all three higher education segments; and the California Partnership for Achieving Student Success (CALPSS). The latter is a voluntary regional system integrating student data from institutions in all four of the state's educational segments (Vernez et. al, 2008). All six of these data systems collect information on enrollment, student demographics, credits earned, GPA, degrees conferred, and financial aid (Vernez et al., 2008). While the details given within these categories vary in each database, the existence of large data collection systems at each level of education can enable California to develop a comprehensive K-20 longitudinal data system.

In 2002, the State Senate authorized the creation of the California Longitudinal Pupil Achievement Data System. CALPADS may be a good foundation for an integrated K-20 system. Its proposed structure enables the state to monitor student's academic progress over time as students move up in grades and/or change schools. It also allows for tracking student progress on the state's Student Testing and Reporting (STAR) tests, California High School Exit Exam, and the California English Language Development Test. California has now developed and assigned Statewide Student Identifiers (SSID) that will track all students throughout their K-12 education in the state without direct link to the student's names or other identifiable information. Although state officials expected the integrated K-20 system to be implemented in 2009-2010, progress has taken longer than anticipated, due in part to fiscal and political constraints. California has made substantial progress on getting the essential elements in place (i.e., from 1 of 10 in 2005 to 8 of 10 in 2010) but have not taken any of the actions set by DQC.

New Mexico

New Mexico has not cultivated an environment of data sharing. Over the years, New Mexico state officials have increasingly seen the need to develop a seamless P-20 system (New Mexico Higher Education Department (NMHED), 2009). Some of the state's major efforts to create an integrated education system include a joint initiative by the Higher Education Department and the Public Education Department (PED) to match high school graduation requirements with college placement requirements. This Alignment Task Force will establish agreements for learning outcomes at both the K-12 and postsecondary levels. Another multi-agency initiative is the Data Sharing Task Force, which aims to establish a student data system from P-16, including adult basic education. New Mexico's HED has taken the lead in this effort to initiate a dialogue between state agencies, including the Office of Education Accountability, PED, New Mexico State University, the Office of Workforce Training and Development, and the Children, Youth, and Families Department.

New Mexico also participates in several national efforts to improve its data sharing abilities, such as the DQC and was one of the initial five states to join Achieving the Dream (ATD; see National Efforts for more detail on ATD). The New Mexico Association of Community Colleges has been assigned as the leading organization in ATD to set a statewide policy agenda to improve community college student success. However, the state's interpretation of FERPA

regulations has placed a premium on maintaining privacy, and as a result, data sharing between community colleges and four-year institutions has been hindered (Mills, 2005). Community colleges have no access to records of transfer students and only receive aggregate institution-type level reports from HED (i.e., not information from individual institutions).

According to the DQC's most recent survey results, New Mexico has 9 out of 10 Essential Elements for a longitudinal data system, but it has only made minimal progress in state actions. The state has successfully developed governance structures and has built a data repository; however, it has not yet linked all K-16 records to workforce data systems.

Arkansas

The Arkansas SUR system established by the Arkansas Department of Higher Education serves multiple purposes including: resource allocation, IPEDS, Perkins, Southern Regional Education Board (SREB), tracking student retention and graduation, tracking students across institutions, and institutional coordination with the State Board. Additionally, the data collected provide legislatively mandated information to the General Assembly concerning enrollments, student semester credit hours, athletics, and state-supported scholarships. The Arkansas system includes all four- and two-year public institutions and, to a certain extent, private institutions. Recently, the state has begun to collect transcript data for institutions and has talked about incorporating private, for-profit institutions into their system (Ewell and Boeke, 2007).

Data collection within Arkansas occurs at institutions eight times per year (twice per term). In addition, workforce-oriented courses, such as English as a Second Language, General Educational Development (GED), and Adult Basic Education, are collected in the same database. These data, however, are not collected at the same level of detail. Arkansas has also begun to track career pathways; participating institutions, however, do not yet have access to these datasets (Ewell & Boeke, 2007).

Arkansas is one of the few states with multiple portals designed with specific users in mind. Public researchers, school leaders, teachers, parents, and students are all awarded different levels of access into databases depending on their role. The use of different portals allows the state to handle issues of student privacy while allowing teachers and parents to have access. Information to other stakeholders (e.g., researchers and policymakers) is available at the aggregate level only. According to the 2009-2010 DQC Survey Results, Arkansas meets all 10 Essential Elements of a longitudinal data system. The state has also been awarded the IES-ARRA longitudinal data systems grant in 2006, 2009, and 2010.

Washington

Washington has a sophisticated and extensive data tracking system from multiple agencies including K-12, community colleges, four-year public institutions, adult education, and the Employment Security Commission (Mills, 2005). In an effort to improve and expand its data sharing capabilities, Washington has begun working towards implementing a P-20 data system. Washington's goal is to become the nation's best longitudinal system with access to student records from K-12 into the workforce. To achieve this, state legislature created the Education Research and Data Center (ERDC) in 2007 to conduct analyses of education issues across the P-

20 spectrum (Pulley, 2010). Agencies included in the state's P-20 system include the Department of Learning, Office of Superintendent of Public Instruction, Higher Education Coordinating Board, Independent Colleges of Washington, State Board for Community and Technical Colleges, as well as the Employment Security Department and the Workforce Training and Education Coordinating Board.

Significant progress has been made by the ERDC since it was created. The development of data sharing agreements with P-20 partner agencies has allowed the linking of K-12 files and postsecondary records. Both the four-year public system and the community and technical college system have experience collecting student data including enrollment, admissions and completion information, courses taken, degree type, and primary field of study. Information from both systems has been incorporated into the ERDC for analysis. Washington's experience with data-tracking has allowed the state to study individuals in the traditional education pipeline, but an unfulfilled goal is to track less traditional students. Washington was among the 20 states to receive an IES-ARRA grant this year and has all 10 DQC Essential Elements in place, but have taken only 3 out of 10 state actions.

Multi-State Efforts

NCHEMS Multi-State Data Exchange

In a study of multi-state data exchanges, the National Center for Higher Education Management Systems (NCHEMS) concluded that it would be feasible to establish a multi-state unit record system (Ewell, Schild & Paulson, 2003). To demonstrate this feasibility, NCHEMS created a pilot study in 2003-2004 to link data between Ohio and Kentucky (NCHEMS, 2008). Within the study, 77,000 students in Ohio and 26,000 students in Kentucky were successfully matched and the frequency of cross-state enrollments was determined. The results suggested a continued value of data exchanges across states. Subsequently, NCHEMS invited two more states – West Virginia and Tennessee – to create a one-time multi-state data exchange.¹²

Guided by the experience from the Ohio-Kentucky study, specific parameters for the multi-state data exchange were established. First, only a few data elements on student demographics, institutions, and enrollment for six complete years were required to be exchanged. Due to privacy and FERPA-related concerns, it was decided that a third-party data administrator would collect, match, and redistribute data files to each participating state. Second, the Ohio-Kentucky study emphasized the need for a formal written agreement delineating specific responsibilities as well as limits on data access and usage. Because laws and regulations vary by state, a bilateral agreement between NCHEMS and each participating state was required. One last requirement was a third-party administrator to de-identify matched files before returning them to each state.

Once all the rules and governing structures were in place, the four-state data exchange successfully matched six years of enrollment history and nearly fifteen million records (NCHEMS, 2008). Like the Ohio-Kentucky results, graduation and retention rates were boosted by small percentage points by this matching process. The results of the data exchange project, however modest, confirmed the ability to exchange data from multiple-state SUR systems and the potential to obtain more accurate estimates of cohort retention and completion. In addition, the researchers pointed out that perhaps other multi-state groupings (e.g., states in New England) would possibly show higher value added. Implications from this study suggest that (a) third party organizations are essential, (b) data quality and privacy concerns remain issues to be resolved, and (c) even with agreements in place, efforts and progress may go slowly. Ultimately, inter-state exchanges may be futile if state budget constraints preclude any use of the data generated. In the following sections, some multi-state initiatives from non-federal entities are highlighted.

Achieving the Dream

In 2003, the Lumina Foundation for Education and eight national partners designed and launched *Achieving the Dream (ATD)*, a multi-year, national initiative to help community college students succeed, particularly students of color and low-income students (Brock et al., 2007). ATD focuses on encouraging schools to better understand and use student outcome data and to identify factors that facilitate or hinder progress (Brock et al., 2007; Jenkins & Kerrigan, 2009).

¹² Indiana was also invited to participate in the project, but eventually withdrew from the project.

Similarly, colleges were expected to use the information collected to drive institutional changes such as academic program review, strategic planning, and budgeting.

In order to improve student success, ATD encourages colleges to follow a five-step process to bring about policy and practice changes. The steps include: commitment to improving student outcomes; using data to identify and prioritize problems; engaging stakeholders in developing strategies; implementing, evaluating, and improving strategies; and institutionalizing effective policies and practices (Jenkins & Kerrigan, 2009). To support this process, ATD provides financial and technical support to colleges over a period of four years. During the first year, ATD offers planning grants of \$50,000 to colleges and implementation grants of \$400,000 following the initial stage (Brock et. al., 2007). The technical support includes outside consultants who assist colleges with performing data analysis, as well as interpreting and communicating findings to faculty and staff. Initially, 27 institutions in five states (Florida, New Mexico, North Carolina, Texas, and Virginia) joined ATD. Currently there are over 100 institutions in 22 states participating in the initiative.

ATD colleges are asked to submit student records to a centralized database created specifically for the initiative. A large portion of the data required for participating community colleges already exists in administrative records, including transcripts (Brock et. al., 2007). Colleges can also include optional data into their own institution's longitudinal database as long as they provide definitions and valid values for each data element. According to the *Instructions for Data Submission* prepared by JBL Associates (JBLA, 2009), colleges are required to submit two types of data files each year:

- Student General Record: student demographics, high school information, and application data are submitted for each first-time (i.e., entering in the fall), credential-seeking student once per year.
- Student Term Record: academic activities such as grade point average, credits attempted, credits earned, major, and CIP code are submitted for each term.

To protect students' privacy, colleges do not submit any identifying information (e.g., names, addresses, or phone numbers). While Social Security numbers are required in both student data files, ATD uses a Student ID Encryption Program that assigns a unique ID number to each student so he/she can be tracked over time. The encryption program generates two files for each of the databases submitted: one secured file for JBLA (i.e., the third party assigned to collect and report data for the initiative) and a file for the institution's records containing Social Security numbers and the encrypted student ID.

A report evaluating the early progress of ATD found that while some faculty and administrators interviewed agreed that longitudinal tracking was a powerful method of identifying gaps in student achievement, few colleges used these data to develop strategies for implementation. Interviewees also criticized the usefulness of the data they were required to share, expressing it was burdensome and duplicative of what their institution was already

doing (Brock et. al., 2007). However, a study by Jenkins and Kerrigan (2009), examining data usage by faculty and administrators at participating ATD colleges, found that those involved in the initiative used data on student outcomes more frequently and participated in discussions on improving student progress much more frequently than their counterparts who were not involved in ATD.

A critical component of the ATD initiative is public policy efforts, both at the national and state levels. Guided by a framework that addresses specific policy areas, each participating state's policy efforts are managed by a lead organization that serves as a network committed to defining a common state policy agenda to increase community college student success.

Data Work Group

Among other efforts to improve community college student success is the Data Work Group, an inter-state collaborative to improve the measurement of student progress between six ATD states – Connecticut, Florida, North Carolina, Ohio, Texas, and Virginia (Goldberger, 2008). The group feels that the current federal method for evaluating community college success is far too simplistic because only degree completion of full-time undergraduates within three years or less is examined (Goldberger, 2008). Many community colleges feel that these parameters overlook a large student constituency (i.e., part-time students), and that many students who begin their education at community colleges transfer to four-year schools and graduate there.

To better gauge the success of community colleges, the Data Work Group made changes to the definition of community college success (Goldberger, 2008) to include: degree completion within six years (i.e., instead of three), number of part-time students still seeking degrees, and counting students who were accepted at and transferred into four-year institutions. This way of evaluating outcomes is completed in addition to analysis of completion by first-time degree-seeking students. In addition, data from younger and older students are examined separately to seek out trends and allow institutions with different student populations to compare data accordingly. Community colleges within states also collaborate to provide outcome data for students who transferred within the state (Goldberger, 2008).

The Cross-State Data Work Group plans to continue to refine its approach to measuring community college performance and to expand its analysis to include more recent cohorts of entering students to track changes in system performance over time. They have also identified several other strategies, such as developing intermediate outcome benchmarks, analyzing the performance of different student subgroups, and assessing the benefits of various interventions to help increase success rates in order to help states use longitudinal data to improve community college outcomes (Goldberger, 2008).

Other Organizational Efforts

Other organizations embarking on these efforts include National Association for Independent Colleges and Universities (NAICU), which represents over 1,000 private institutions in the United States, and the National Governors Association. NAICU's University and College Accountability Network (U-CAN) was created in response to a call by Congress for the availability of better data (University and College Accountability, 2008) and provides

information on college or university admissions, enrollment, student demographics, graduation rates, campus size, tuition and fees, and other campus information that is helpful to students (University and College Accountability, n.d.). NAICU's U-CAN is an effort by private institutions to become more transparent.

In addition, the National Governors Association (NGA), while focused on secondary education, has taken steps to create a common method for determining states' high school graduation rates (Implementing graduation, 2006). All 50 state governors agreed to work toward creating a common method for calculating graduation rates, along with common methods for collecting data, documenting dropouts and completions, and reporting findings. The efforts made by the NGA can provide an example to postsecondary educational leaders when considering the creation of a national unit-record system.

Finally, the National College Access Network (NCAN) created a Web-Enabled Student Tracking (WEST) system, which allows college access programs to collect and report information about the progress of their students through an online database. NCAN has partnered with the NSC to develop a postsecondary module that will display students' enrollment status, which will aid college access programs who wish to track the success of their participants.

National/Regional Efforts

At the national level, the US Department of Education inserted a provision into the 2008 Higher Education Act Reauthorization that restricted the federal government from collecting student-based unit record data. This followed a multi-year dialogue about how to improve the Integrated Postsecondary Education Data System, better known as IPEDS, which collects massive amounts of data from every Title IV institution (i.e., institutions that are eligible to deliver federal financial aid). Data collected include student enrollment, retention, graduation, and even other areas such as institutional finances.

Still, there remains the opinion that there should be a national student unit data system operated by a third-party. The idea of a federally-controlled SUR data system has been met with mixed reviews in higher education. While some believe that it is a good idea, others adamantly believe that current efforts are sufficient or that the potential cost to student privacy outweighs the benefits. Lobbyists for various associations suggest that a SUR system would be too demanding on institutions; others cite concerns over student privacy and do not want the federal government deeper into the business of collecting personal information from students.

Restructuring IPEDS

The Integrated Postsecondary Education Data System (IPEDS) is the Institute for Educational Sciences' (IES) primary data collection system. More than 10,000 institutions report data to IPEDS each year. Presently, the IPEDS system collects institution-level data in the fall, winter, and spring of each year about enrollment data for full- and part-time students, level of enrollment (undergraduate, graduate, or professional), degree completion by level of degree, the number of students receiving financial aid, the type and amount of aid received, graduation rates within 150% of the nominal time to degree, and expenses including tuition, room, board, and student fees (Cunningham et al., 2005). All Title IV institutions are required to submit data to the IPEDS system, which is made available to researchers and students via the IPEDS website.

The lack of student-level data in IPEDS makes it difficult for National Center for Education Statistics (NCES) to study changes in student enrollment patterns, track students who concurrently enroll in multiple institutions or transfer or track the cost of higher education per student accounting for financial aid. Growing interest in the feasibility of a SUR system that could cross state borders intrigued the US Department of Education, which, through NCES, studied the feasibility of initiating a SUR system on the federal level, and whether or not the federal government should create such a system (Cunningham et al., 2005).

In place of current reporting elements of the IPEDS system, it was suggested that institutions submit individually identifying information for each student, enrollment information, degree completion statistics, and financial aid information. Once data were collected, identifying information could then be removed and the dataset could be used to advance research and understanding of postsecondary students. Gathering student data in lieu of institutional (i.e., aggregate) data has some advantages for higher education institutions, which would be able to use the information to determine what happens to students who leave their institutions. Understanding of the patterns of non-traditional students, who make up an increasingly large

part of the total student population, would also be improved under a more comprehensive system.

Cunningham et al. (2005) determined that the federal implementation of such a system would not be an unreasonable burden for most institutions, given that 40 states already have unit record systems in place. Many institutions already submit unit record data to fulfill Title IV requirements or submit data to private organizations. Adequate time investments and financial allocations are necessary for success, but also very viable options.

The authors assert that all student information should be reported into the system and that no students should be able to opt out on the basis of not utilizing financial aid (Cunningham et al., 2005). The argument used to support this is that all students benefit from a school's receipt of financial aid money, whether or not an individual student receives direct aid. The accuracy of information also depends on the completeness of the record, which would be compromised if students were given the option to withhold their information from the IPEDS system.

The National Student Clearinghouse

Perhaps one of the most comprehensive SUR systems presently in place in the US is the NSC. The NSC is a non-profit organization dedicated to reducing administrative tasks for institutions by verifying educational records (NSC, 2008). It began in 1993 as a means for loan lenders to track students to secure repayment. The mission of NSC, today, is to facilitate the exchange of information among institutions, loan guarantors, and third-party entities. To access NSC data, institutions must submit student rosters to NSC, which are then matched to a list of students receiving loans (provided by lenders). NSC matches this information and returns compiled information to the institutions and the lenders (Carey, 2008). The current cost to receive NSC services is 10 cents per enrolled student. Degree verification services are five cents per student and there is a minimum fee of \$400 to use NSC each year.¹³ Per our conversation with NSC, it was created and continues to act "to serve the institution," and prefers to look at cohorts of students rather than institution-wide reporting. In addition, it is not NSC's intention or interest to rank institutions and they do not favor crossing that line, in order to protect their users. In addition to information already collected, the APLU hopes to work with NSC to supplement their data with transcript records. This could be an additional benefit to institutions when looking at student achievement.

Information from NSC can help produce cohort graduation rates at two- and four-year institutions and has the ability to follow students beyond a particular school, which is critical for transfer students. NSC presently houses about 80 million student records, or approximately 91 percent of all students through 3,200 participating institutions (NSC, 2008). Hundreds of high schools across the US also use NSC services and may join for its *Student Tracker* service, which allows high school administrators to see which students pursued higher education information opportunities and the level of persistence to graduation (NSC, 2008). This information could be beneficial to school districts as they implement change and seek to prepare students for higher education. Recently, the use of NSC at the high school level has been specifically targeted. In 2009, Bill and Melinda Gates Foundation awarded a \$2.9 million grant to NSC to develop a

¹³ Based on 2010 information.

system to enable high schools in all states to track their students' academic postsecondary success and to provide accountability at the K-12 level (Lederman, 2009). While NSC is the largest student record system to date, the scope of services provided is limited as they specifically focus on enrollment, lending, and transfers. In addition, there are variations in how data are provided to NSC.

Feedback from the Field on a National Student Unit Record Database

Expert Panel on IPEDS

In discussions with our expert panel, the prevailing sense was that IPEDS did not incorporate enough information to be sufficient or helpful as a comprehensive SUR system. IPEDS data were criticized by some respondents for its lack of comprehensiveness of cohorts, its inability to follow students across institutions, the comparability of data between institutions, and its narrow focus of first- and full-time students. For example, the executive director of research at a regional community college in Illinois stated that, "in community colleges, there are very few first-time, full-time freshman" [making IPEDS less relevant to these schools]. In addition, transfer rates from these schools are important to examine, as many schools express that this can be an indicator of success. The respondent offered, by example, that their three-year graduation rate was under 10 percent but that their transfer rate was 40 percent. A recent survey done by the college suggested that 70 percent of their transfer students attended a college or university after completing a minimum of 30 credits.

A few respondents noted that while the current IPEDS system may not be as effective or complete as researchers would like, they felt that it is important to consider whether a national system would yield enough significance for the drawbacks to be warranted. For example, there were concerns regarding the use of social security numbers as record identifiers, that individually identifiable data should not be held by the federal government or higher education institutions (rather, it belonged in the care of a neutral third party broker), whether higher education institutions would be forced to submit data, the potential use of the data to rank institutions, and the burden that such a system poses on institutions. Several respondents indicated that while many institutions already have the capability to report SUR data, there are significant start-up costs for institutions such as software updates, staff training, or hiring new staff. This can be especially burdensome to small institutions that are economically strained. One respondent stated that if an institution were to choose between investing in research regarding SUR systems and investing in a program, "the institution will invest in the program."

Some individuals believed that if SUR data were collected in a similar manner to current IPEDS, it would not benefit educational researchers and policymakers. An institutional research from Colorado indicated that analysis of graduation rates by discipline "does not make sense at the graduate level," let alone the undergraduate level because of high variability between institutions at the undergraduate level. He went on to note that at his university, more than half of the freshmen enter without a major, making it very difficult to develop a "starting cohort."

There was some positive feedback regarding a restructuring of IPEDS. Some respondents believed that the standard formatting on a large scale could make the system more user-friendly

and that despite criticism to the current IPEDS system, it has worked well in various aspects of data collection within higher education. Some suggestions provided were that a federal system of data collection could be linked to funding of *all* institution types, to increase the number of private institutions which provide data. In addition, as IPEDS has increased the number of reporting requirements for institutions anyways, some interviewees felt that institutions may naturally move toward SUR data collection.

Expert Panel on NSC

During our interviews, respondents had varied opinions about NSC and the role that it could play in the creation of a national SUR system. While some believed that the design and reputation of NSC make it a good choice for housing a national record system, others were concerned about its business-minded model. Those in favor of considering NSC as a potential holder of student records at the national level asserted that NSC has “maintained confidentiality clearances since they began their operation and... have worked within the laws of FERPA.” Institutional representatives particularly favor NSC because of its allegiance to institutions, not the federal government or other policy organizations.

Other individuals did not believe that NSC was an appropriate forum for housing national SUR data, as it was designed to consolidate the efforts being made between institutions and financial lenders. For this reason, some find its purpose and mission to be too narrow to accommodate a completely national system. Specifically, NSC data neither distinguishes what degree a student graduates with nor whether the student is full- or part-time; leading some to believe it is unlikely that NSC will be able to provide institutional information for other purposes. In addition, no data are collected regarding program of study, course completion, or progression. One respondent noted that, although NSC is a “way of linking systems without threatening transparency or accountability,” the system could never provide “the basis for a public system of accountability.”

Finally, researchers who were not directly affiliated with a postsecondary institution noted that NSC is often unresponsive to many of the needs of policy-based organizations. Due to the high costs of NSC, the system is often unfeasible to utilize for smaller studies or organizations. Some believe that NSC is money-driven, and that its “business design” would make it unsuitable for housing a system for the public good.

Survey Respondents on Viability of a SUR System

The survey described in the Institutional Efforts section also asked respondents about the current and future viability of a national SUR system. A divided opinion emerged on the issue of creating a national SUR system: 34 percent stated they thought such a system was viable right now, whereas 37 percent of respondents felt that it was not viable currently (11 percent responded that the question was “Not Applicable”). Those who responded that it was not currently viable were also asked whether a national system would *ever* be viable. Of those queried, 58 percent reported that it would never be viable.

As a follow-up item to the closed-ended viability questions, respondents were asked to provide open-ended responses detailing their opinions. Themes are provided below for each of the three

categories of why respondents felt a national SUR system was currently viable, not currently viable but potentially viable in the future, and not viable ever. Additional comments are provided in Appendix C.

Among respondents who felt that a national SUR system was currently viable, common themes included:

- The technology exists to make this possible (e.g., records are already electronic and states are already beginning to do this).
- It is already used at the state level (i.e., information collected at the state level can be “fed into a national database”).
- A form of this already exists through the NSC (e.g., NSC could be expanded or parallel system could be created).
- If there were a system, it would be beneficial to institutions, researchers, and students, as it would allow for better student tracking (e.g., given the mobility of students, this would provide a way of assessing student success when they have moved to a different institution) and is essential to serving students and promoting access and completion of higher education.
- It would help with Federal accountability and funding.

Additionally, some of the respondents – despite indicating that a national system was already viable, were concerned that it was not a good idea to move forward yet and/or listed foreseeable obstacles. Examples of these comments include:

- Viability is not enough; even though it is viable, it is not “cost effective,” “desirable,” and may not be “politically acceptable.”
- Problems with privacy issues.
- Data quality issues (e.g., lacking standardization or “data formats and definitions”).
- Issues with the political will to fund a system.
- Issues with cooperation (i.e., “buy-in from state education departments,” local governments, institutions, and the public is needed and if a mandatory system would be required to make this possible).

Of the respondents who indicated that a national SUR system was not currently viable and answered the question regarding whether a national SUR system would *ever* be viable, 43 percent said it would be viable in the future. Respondents provided areas that needed to first be addressed to make it possible, many of which mirrored the responses provided about obstacles by the respondents who indicated the system was already viable (see above). The areas to be addressed included:

- Financial resources for the startup and maintenance of a system.
- Cultural change regarding data collection and sharing as a means to support the educational enterprise and not for penalizing institutions.
- Logistical burdens placed on campuses – additional resources and/or a mandatory system might be needed.
- Technology needs to be improved.
- Data quality needs to be improved.
- Privacy concerns must be addressed.

Of the respondents who indicated that a national SUR system was not currently viable and answered the question regarding whether a national SUR system would *ever* be viable, 58 percent said it would not ever be viable. Interestingly, the obstacles mentioned echoed those mentioned by respondents who said it would be viable; the difference is that this group sees the obstacles as insurmountable. These themes included:

- Lack of data quality or ability to standardize and ensure quality.
- High costs and additional staffing requirements.
- Low perceived return on investment.
- Lack of support from members of the higher education community, particularly private institutions.
- Issues with privacy and fear of government misuse of data.
- Basic lack of need for the system, given NCS and existing longitudinal data bases.

Voluntary Systems of Accountability

Many individuals and organizations believe that the next step toward reliable SUR data is not through mandatory reporting, but through the creation of multi-state consortiums and voluntary participation. The goal for many groups is to create a voluntary system that is so influential that all higher education institutions want to participate. Regional accrediting bodies could mandate participation for accredited institutions or associations with similar goals and create a program-wide data sharing system. Multi-state consortiums could also be formed to track students regionally. An obstacle to this approach is that states interpret the FERPA laws differently, which hinders the sharing of student data. Many of our interviewees suggested that FERPA laws are clarified before a national voluntary system could be fully implemented. In addition, state attorneys often advise against these multi-state data systems due to legal privacy issues. One of our respondents stated that “every single data request goes through the Office of General Counsel at colleges and universities.” These legal counsels often “err on the conservative side for fear of being sued,” making it difficult for voluntary data collection systems to receive sufficient information.

The APLU and AASCU have asserted the position that as colleges and universities are held accountable by an increasing number of constituencies, it is wise to establish a voluntary system for longitudinal data collection (McPherson & Shulenberger, 2006), thereby calling for a Voluntary System of Accountability (VSA). Approximately 70 percent of four-year institutions in the United States are represented through APLU and AASCU, which means that the voluntary system can have a significant impact on higher education (Miller, 2008). In June 2008, there were 250 institutions participating in the voluntary system. APLU and AASCU predict that the number of participants will increase with time. George Mehaffy, vice president of AASCU, believes that taking steps towards voluntary accountability is a proactive way for higher education institutions to make their own decisions without the assistance of the federal government (Miller, 2008).

The primary product of the VSA is called *College Portrait*, an online self-reporting tool that allows for timely feedback and increase student learning by assessing courses and skills, providing the public with a range of consumer information about participating institutions.¹⁴ APLU presidents and provosts agree that there are three groups of stakeholders to which they should hold themselves accountable: prospective and current students and their parents, college and university faculty and staff, and policymakers and public and private funders of postsecondary education. College and university faculty and staff can benefit from voluntary accountability data if it is used to show which practices are the most educationally effective. The ability to compare information between institutions can be an educational tool for improving best practices. APLU believes that if an institution's focus is on student learning, the institution should have the ability to compare its levels of student engagement with other postsecondary systems and relative to their own goals and predictions. Over time, these data can provide a base for analyzing an institution's success or shortcomings.

¹⁴ <http://www.voluntarystem.org/>.

Barriers to Unit Record Data Collection Systems

As has been highlighted throughout this report, barriers and complexities to the collection of SUR data exist at all levels. At the institutional level, these include the burden of data entry and tracking, the decision of when to begin tracking a student, and how to handle transfers and major changes. As well, there exists the issue of who should take responsibility for tracking students – the departments in which students major or the institution as a whole? At different institutions, different offices (e.g., admissions, institutional research, registrar, academic departments) implement tracking systems and have different guidelines by which they collect data. In addition, reporting requirements may increase the workload of current staff or require the addition of new positions, presenting financial challenges for some institutions, particularly if this were to be federally mandated. It should be noted that opinions regarding burden tend to vary, based on our survey and discussions with the expert panel. Some noted that many institutions already maintain a SUR system, so the creation of a national system would not be an additional burden. Others believed that the time involved in maintaining SUR data could be a huge institutional burden and potentially not yield information significant to an individual institution.

At the state level, the primary barriers include collecting data that varies across institutions and handling student transfers across state lines. To cope with the latter challenge, there are efforts to create multi-state exchanges, but the issue of consistency in data across institutions and states and student privacy remains a concern. The same concerns apply at the national level. Overlaying all levels is the cost of SUR systems and how to incent institutions to collect data and comply with submitting it at the state and national level. Generally speaking, the expert panel expressed concerns that institutions would struggle with how to report data and that concerns institutions would not want all data public or to be used for rankings.

This report has outlined the findings of the NCHEMS studies and the work of the DQC efforts, which collectively illustrate that states are now in a much better position to look at longitudinal data tracking of unique students throughout the system. In fact, some states provide exemplary models for collecting student data and can provide a foundation for other states to also begin this process. In addition, multi-state exchanges have shown preliminary successes in their approach and could possibly be taken to scale. This is in the context of a cumbersome system, with a largely absent federal support for these efforts. In addition, FERPA has caused states, institutions, and school districts to significantly limit the data made available for research and other uses. The restrictions placed on student data sharing vary across states, along with their interpretations of the law. National action in this area specifically may be required to clarify FERPA in such a way that, at the least, this variability is eliminated.

Discussion

The discussion of student data systems can be cast as tradeoffs in policy and practice. In simple form, we need to know what happens to students, from middle school, through high school, and to and through higher education. Our national longitudinal studies, sponsored by the US Department of Education and including the Educational Longitudinal Study (ELS), the Beginning Postsecondary Student (BPS) Study, and Baccalaureate and Beyond (B&B), do an excellent job of providing us with time capsule snapshots of what happens to a particular cohort of students. But these surveys are administered years apart, which may be too long to inform public policy. A systematic effort to yield much better numbers on an annual or biennial basis could help public policy and institutional practice.

The current IPEDS reporting provides graduation rates but is limited by the restriction of certain subsets of students, its use of aggregated data from institutions, and the accuracy of data entry and reporting. This last item has consistently been perceived as a severe limitation of IPEDS. IPEDS will continue to improve, but any movement toward student unit record data is a non-starter.

Through our review of literature, interviews with key stakeholders, and a national survey, we find that the predominant view is that a national student unit record system would be a useful – if not critical – tool for public policy. Some respondents and reports noted the limitation of not being able to track students from level-to-level, school-to-school, and state-to-state. With a growing population that is more transient and, as Cliff Adelman noted back in 1999, tends to “swirl” between institutions, we cannot simply assume that the majority of students who are traditional in their college going are representative enough of the larger college population.¹⁵

Community colleges have always been averse to the current and simplistic use of graduation rates as a measure of institutional efficiency and excellence. While these schools have relatively low graduation rates, they also tend to have significant transfer rates, which, in most of our current systems, are not used effectively as output measures. Policymakers focus on graduation rates, not transfer rates.

Private schools, both proprietary and non-profit, have also pushed back on large, student unit record data systems. This objection typically pivots on claims of institutional burden, and for some small private schools, this argument can be sincere. There can be an extra burden for a small staff in providing additional data for reporting purposes. But for most colleges and universities, the institutional burden argument is limited at best. If public institutions have figured out largely how to do this under current systems, often at the bequest of the state department of higher education, and if others voluntarily participate in pilot projects, it cannot be too onerous. Indeed, we would posit that institutional burden is more an argument for the sake of argument, and much less an issue in the real world. Still, the institutional burden argument is a major reason why national systems have not gone forward and have resulted in

¹⁵ See Adelman, Clifford (1999). *Answers in the Tool Box: Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education. (<https://www2.ed.gov/pubs/Toolbox/Title.html>).

the federal government backing away from their “HUGE IPEDS” proposal just a few years back.

As one of the expert panel members suggested, getting people to agree on what variables to use sounds simple enough, but remains a very difficult task. Defining a “student” gets quite complex (full-time, part-time, first-time, transfer, degree seeking, etc...), and that is perhaps one of the simpler variables. Other variables to be defined, such as retention, persistence, and even education units (e.g., quarters, semesters, academic years) are unnervingly complex. Nevertheless, it can be done, as some of the programs in this report have proven.

Overall, the majority of those interviewed agreed that the major issue in creating a single national student unit record system is establishing a common protocol of how an institution must report its data and developing technical definitions of the data elements to ensure uniformity. There is also concern surrounding the accessibility of the data, especially if an institution could be ranked in national publications featuring open and transparent institutional data. Finally, a national system would have to account for variables that could complicate the data input such as name changes, student transfers between public and private institutions, community and four-year colleges, and across state borders. There is a huge transfer rate between community colleges to four-year schools, and many institutions are now becoming interested in tracking students from primary school through postsecondary school and then into the job market. This interest has led to further exploration of a national student unit data system to track P-20 student data.

The issue of working across states is still a perplexing one. Florida and several other states have an excellent handle on tracking students from Kindergarten to undergraduate and graduate school within state, but once the state border is crossed, the data tend to fall apart. Clearly, an interstate system would be desirable.

The greatest challenge, from our point of view, remains with FERPA. This law has come to serve as a useful tool of legal counsel at school districts and postsecondary institutions who wish not to participate in research projects or large-scale studies under the auspices of “protecting” students. In fact, the current voluntary and similar systems discussed in this publication have built in work-arounds to protect against the misuse of student data. But FERPA, like affirmative action issues, has become the legal shield against federal and/or state requests. The problem with FERPA has always been the legal interpretation. To be clear, FERPA, as law, never stated that data were prohibited for research reasons. In fact, it supported the use of data for research, as long as the identities could be protected. FERPA also never said that social security numbers could not be used, and the research and education communities largely agree that other student identifiers should be used in lieu of SSN. In the end, certain stakeholders and policymakers rely on FERPA to impede policy-based research on the basis of some pseudo-political philosophy of fairness and protectionism that fails to exist.

At some point in time, we must move beyond the politics and the legal barriers to collect better student data. The National Student Clearinghouse has been mentioned as a potential home for a national dataset, as it currently houses data on approximately 91 percent of postsecondary

students. The Achilles heel of the NSC is that their dataset is extremely limited to only a handful of variables. The Clearinghouse would only be truly effective for this purpose if the number of variables collected on students could be increased such that we had a better idea of student status and transfer. This stated, with regard student unit record database, the NSC system is far beyond anything else available at this time.

It is unlikely that a mandatory system will be put in place by the federal government in the foreseeable future, and the HEA reauthorization of 2008 seemingly closed the door on this possibility. Thus, it seems that any national dataset, however one defines national, will reside in a voluntary system. We have enough experience from the current voluntary systems to know what works and what doesn't. If we can build a system where membership is inclusive rather than exclusive and where institutions are noticed for lack of involvement, then, perhaps, it can fulfill the expectations and outcomes of a federally-mandated system.

Such a system would require states and institutions to come to agreement on variables, timetables, and technical issues. While there isn't an "easy button" for this type of work, we have salvation in the fact that current initiatives have paved the way. The Data Quality Campaign has made significant progress in getting states to understand where their data systems reside with respect to other states and what "quality" really means. The APLU and ASCCU Voluntary System of Accountability has already established practices and agreements that can be used as a foundation for either a larger system or a new system.

References

- Achieving the Dream. (n.d.). Retrieved on April 5, 2009 from <http://www.achievingthedream.org/default.html>.
- Brainard, S. G. & Carlin, L. (1998). A longitudinal study of undergraduate women in engineering and science. *Journal of Engineering Education*, 87(4), pp. 365-375.
- Brock, T., Jenkins, D., Miller, J., Gooden, S., Martin, K., MacGregor, C., et al. (May, 2007). Building a culture of evidence for community college student success: Early progress in the Achieving The Dream initiative. Retrieved on April 5, 2010 from <http://www.mdrc.org/publications/452/full.pdf>
- Carey, K. (2008, May 22). Demise of a good idea. *Inside Higher Ed*. Retrieved on August 11, 2008 from <http://www.insidehighered.com/views/2008/05/22/carey>.
- Chubin, D.E., May, G.S., & Babco, E. (2005) Diversifying the engineering workforce. *Journal of Engineering Education*, 94, pp. 73-86.
- Consortium for Student Retention Data Exchange. (2010). Retrieved March 29, 2010 from <http://csrde.ou.edu/web/index.html>
- Cunningham, A.F., Milam, J., & Statham, C. (2005). Feasibility of a student record system within the integrated postsecondary education data system. U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: US Government Printing Office.
- The Education Data Warehouse and Data-Driven Decision-Making. (2010). ARM Snapshots Series. Florida Department of Education, *Volume 1*(2). Retrieved from <http://www.fldoe.org/arm/pdf/ARMSnapshot2.pdf>
- Data Quality Campaign. (n.d.). Retrieved June 29, 2009, from <http://www.dataqualitycampaign.org/>.
- Data Quality Campaign. (2006). Virginia case study: Building a student-level longitudinal data system. Retrieved on April 7, 2010 from <http://dataqualitycampaign.org/resources/details/321>
- EdSource. (n.d.). CALPADS represents a major shift in California's approach. Retrieved from http://www.edsource.org/iss_datasys_calpads.html
- Education Research & Data Center. (ERDC; n.d.a). Retrieved April 6, 2010 from <http://www.ERDC.wa.gov/>
- Education Research & Data Center. (n.d.b). *Washington's application: Grants for statewide longitudinal data system under ARRA*. Project Narrative. Retrieved April 6, 2010 from: <http://www.ERDC.wa.gov/arraslds2009/>
- Ewell, P. & Boeke, M. (2007). Critical connections: Linking states' unit record systems to track student progress. Indianapolis, IN: Lumina Foundation for Education.
- Ewell, P.T., Schild, P.R., & Paulson, K. (2003). Following the mobile student: Can we develop the capacity for a comprehensive database to assess student progression? Indianapolis, IN: Lumina Foundation for Education.
- Florida Department of Education PK20 Data Warehouse. (n.d.). Retrieved August 11, 2008, from <http://edwapp.doe.state.fl.us/>.
- Florida's Statewide Longitudinal Data System. (2010). ARM Snapshots Series. Florida Department of Education, *Volume 1*(1). Retrieved from <http://www.fldoe.org/arm/pdf/ARMSnapshot1.pdf>
- Goldberger, S. (2008). Test drive: Six states pilot better ways to measure and compare community college performance. Boston: Jobs for the Future.
- Implementing graduation counts: State progress to date (2006). Washington, D.C.: NGA Center for Best Practices.
- Institute of Education Sciences. (2010). 20 States win grants for longitudinal data systems. In Statewide longitudinal data systems grant program. Retrieved on July 7, 2010 from http://nces.ed.gov/programs/slds/fy09arra_announcement.asp
- Integrated Postsecondary Education Data System. (2008). Retrieved August 18, 2008, from <http://nces.ed.gov/IPEDS/>.
- JBL Associates. (2009). Instructions for data submissions. Bethesda, MD: JBL Associates, Inc.

- Jenkins, D. & Kerrigan, M.R. (2009) *Faculty and administrator data use at Achieving the Dream colleges: A summary of survey findings*. New York, NY: Community College Research Center and MDRC.
- Lederman, D. (2009, February 13). A national (but not federal) student database? *Inside Higher Ed*. Retrieved on June 27, 2009 from http://www.studentclearinghouse.org/about/news/pdfs/InsideHigherEd_021309.pdf.
- McPherson, P. & Shulenberger, D. (2006). *Toward a voluntary system of accountability program (VSA) for public universities and colleges*. Washington, D.C.: National Association of State Universities and Land-Grant Colleges.
- Miller, M.A. (2008). The Voluntary System of Accountability: Origins and purposes. *Change*, 40, 8-13.
- Mills, J. (2005). *State data systems and privacy concerns: Strategies for balancing public interests*. Boston: Jobs and Future.
- Missouri University of Science & Technology. (n.d.). Freshman Engineering Program. In *Future students: Undergraduate degree programs*. Retrieved on May 19, 2010 from <http://futurestudents.mst.edu/degrees/undergraduate/fe.html>.
- Missouri University of Science & Technology. (2009). Retention committee 2008-2009 Annual report. In *Office of Enrollment Management*. Retrieved from <http://enrollment.mst.edu/committees.html>.
- National Center for Higher Education Management Systems. (2008). Tracking postsecondary students across state lines: Results of a pilot multi-state data exchange initiative. Retrieved April 9, 2010 from <http://www.nchems.org/c2sp/>.
- National Science Board. (2008). Science and Engineering Indicators 2008. *volume number 1 (NSB 08-01); volume number 2 (NSB 08-01A)*.
- National Science Board. (2007). National Action Plan: Addressing the critical needs of the US Science, Technology, Engineering and Mathematic Education System.
- National Student Clearinghouse. (n.d.). Retrieved August 27, 2008, from <http://www.studentclearinghouse.org/about/default.htm>.
- Neugent, L.W. (2006). Educational Information Management System (EIMS): Report to the Virginia Board of Education. Richmond, VA: Department of Education. Retrieved April 8, 2010 from http://www.doe.virginia.gov/VDOE/VA_Board/Meetings/2006/ItemC-apr-pln.pdf
- New Mexico Higher Education Department. (2009, September 18). *P20 policy programs Pre-K to college*. Retrieved April 8, 2010 from <http://hed.state.nm.us/content.asp?CustComKey=435112&CategoryKey=435119&pn=Page&DoName=hed.state.nm.us>
- New Mexico's state profile. (n.d.). Retrieved April 8, 2010 from Achieving the Dream website: http://www.achievingthedream.org/_images/_index03/NewMexico_State_Profile.pdf
- Pulley, J. (2010). *Off the charts: Columns of data provide support for student achievement*. Focus [Magazine]. Indianapolis, IN: Lumina Foundation for Education. Retrieved from <http://www.luminafoundation.org/publications/>
- Romberger, K. (2008). Proposed changes to FERPA regulations. *Education Law*. Retrieved June 26, 2009, from <http://educationlaw.foxrothschild.com/2008/03/articles/ferpa-1/proposed-changes-to-ferpa-regulations/>.
- Smith, T.Y. (2002). The consortium for student retention data exchange: A case study of consortium development [AIR 2002 Forum Paper]. Retrieved from ERIC database (ED472660).
- Student database poses privacy concerns but some say it would provide better data on higher ed trends. (2005). *University Business*, May 2005. Retrieved August 18, 2008, from <http://www.universitybusiness.com/ViewArticle.aspx?articleid=394>.
- Swail, W., Redd, K., & Perna, L. (2003). *Retaining Minority Students in Higher Education*. ASHE-ERIC Reader. San Francisco, CA: Jossey-Bass.
- Texas Education Agency. (2010). *Texas P-16 public education information resource*. Retrieved May 19, 2010 from http://www.texaseducationinfo.org/tpeir/about_data.aspx
- The Voluntary System of Accountability. (VSA) overview of College Portrait. (2008). Retrieved August 22, 2008 from <http://www.voluntarysystem.org/index.cfm>
- University and College Accountability Network. (n.d.) Retrieved October 13, 2008 from <http://www.ucan-network.org/>

- United States Department of Education. (2008). *Family Educational Rights: Final Rule* (Federal Register No. 34 CFR Part 99). Washington, D.C.: Government Printing Office.
- Vernez, G., Krop, C., Vuollo, M. & Hansen, J.S. (2008). *Toward a K-20 unit record data system in California*. Santa Monica, CA: Rand Corporation.
- Winnick, S.Y., Palmer, S.R., Coleman, A.L. (2006). *State longitudinal data systems and student privacy protections under the Family Educational Rights and Privacy Act*. Austin, TX: Data Quality Campaign.

Appendices

Appendix A: Interviews with Expert Panel

To support this project, an expert panel was established. The panel consisted of 20 individuals from 16 organizations, who were primarily selected from the original attendance list at the January 2008 APLU meeting in Washington, DC. In addition, individuals were added based on the suggestions from the original group. During October and November 2008, The Educational Policy Institute spoke by telephone the expert panel. Topics of discussion included the issues related to student unit record systems and “tracking” student activities and movements. Discussions included the National Student Clearinghouse, IPEDS, voluntary data systems, multi-state consortia, and the burden of creating a national student unit record system.

Expert panel members included the following:

Trudy Bers	Oakton Community College
Marianne Boeke	NCHEMS
Stephen DesJardins	University of Michigan Center for Higher and Postsec educ.
Peter Ewell	NCHEMS
Michael Gibbons	American Society for Engineering Education
Christine Keller	APLU and Voluntary System for Accountability
William Kelly	American Society for Engineering Education
Charlotte Kuh	National Research Council
Paul Lingenfelter	SHEEO
Hans L'Orange	SHEEO
David Longanecker	WICHE
Todd Massa	State Council for Higher Education in Virginia
Lou McClelland	University of Colorado at Boulder
Patrick Mulvey	American Institute of Physics
Karen Paulson	NCHEMS
Rich Reeves	National Student Clearinghouse
David Schulenberger	APLU
Susan Traiman	Business Roundtable
John Vaughn	Association of American Universities
Tom Weko	NCES

Appendix B: FERPA

The Family Educational Rights and Privacy Act (FERPA) is a federal law that protects the educational records of students enrolled in any school which received US Department of Education funding. With the exception of directory information, educational institutions must have permission from eligible students (students who are over 18 years of age) or parents before releasing records. Exceptions to this policy include the release of student records to other institutions for transfer purposes, for audit or study purposes, or in the case of health and safety emergencies (US Department of Education, 2008). Records may only be shared if the information does not identify the student, or if the circumstance for the disclosure meets one of the statutes established within the law (Winnick, Palmer, & Coleman, 2006). Different interpretations of the law however have enabled some states to collect student records. In the absence of a consistent interpretation of the FERPA act, any unified move towards a national data repository has been stifled.

The ability to maintain student privacy is a concern that has been raised by many organizations which object to a nationally maintained SUR system. For example, the issues of a student's right or ability to withhold information, the re-disclosure of information gathered, and the liability of institutions once data has been submitted (Cunningham, Milam & Statham, 2005; Student Database, 2005) have all been discussed. Currently, students who receive federal financial aid are already tracked through unit record systems because they have provided consent by completing and signing the Free Application for Federal Student Aid (FAFSA). While some argue that additional data on students who do not receive financial aid is unnecessary, Cunningham et al. (2005) disagree, asserting that data about unaided students is important for comparing to aided students. Additionally, some individuals and organizations believe that non-aided students benefit from financial aid in the form of subsidies and deferred tax revenues (Cunningham et al., 2005). Therefore, student-level data should be gathered on non-aided students in addition to aided students to give a comprehensive picture.

Some institutions have expressed concern about the legal liability of releasing information to the NCES if that information is then used improperly or disclosed to an unauthorized third party (Cunningham et al., 2005). The fear of legal liability for institutions is unfounded because submitted data become the property of NCES. It therefore becomes the responsibility of NCES to ensure that all re-disclosures of information are handled in a manner that is legal through FERPA. It is possible that other federal entities would like to use the data for reasons that are not educationally based but would only be allowed if permitted by law or if Congress created laws to make this possible (Cunningham et al., 2005).

Identity theft and the possibility of the system being hacked are also concerns surrounding the creation of large and comprehensive database. Student records would be connected with social security numbers, which increases this fear. While many institutions are choosing to use identification methods that do not involve social security numbers, social security numbers are needed to accurately link records across institutions and state lines (Cunningham et al., 2005). While NCES acknowledges the concerns that individuals and institutions have about the safety of data, it is also legislatively protected and well equipped to keep the data secure. Any wrongful disclosure of information within the system is "a Class E felony punishable by up to

five years in jail and a \$250,000 fine” (Cunningham et al., 2005, p. 36). NCES already handles large datasets that contain identifiable information, and no hackers have succeeded in breaking through NCES firewalls.

The final rule released in December 2008 failed to offer clarification of what a “rational basis” was or by whom the decision to share private information should be made (Romberger, 2008). For the time being, this is left to individual institutions.

Appendix C: Open-Ended Survey Responses

Institutional Level

Regarding the biggest challenges and concerns in tracking students in their systems, 430 respondents provided open-ended feedback, through which some common themes emerged. Representative quotes are provided here for the themes.

- Privacy and confidentiality issues (i.e., FERPA compliance, issues of using social security numbers, “ensuring appropriate levels of data security...” and “that student’s personally identifiable information is secure,” fears that data would be misused, and concerns that “government can sometimes unethically use information”).
 - In addition, there were concerns over students’ willingness to share data (e.g., “A lot of students do not identify with the school and thus do not have a vested interest in helping the school acquire this needed data”).
- A number of concerns were mentioned regarding reporting burden:
 - “The human resources needed to study the data, analyze, and report it” and to ensure that “data [are] reliable and valid” is of particular concern.
 - “Leadership will need to value these activities and require them as part of job duties. Until that happens, there will be no tracking of consequence.”
 - Lacking infrastructure of institutional research departments (e.g., “I am a one-person office, and I am responsible for all federal, state, and outside survey requests, as well as for providing data and research to faculty and staff throughout our campus.”).
 - The work involved in submitting data to the state, writing reports, and maintaining up-to-date records, especially given the economy.
- Staff Training (e.g., “making sure that everyone is aware of the capabilities of the system,” “educating people about how best to use such data,” and “training everyone to use the same definitions consistently, accurately, and in a timely fashion”).
- Timeliness of reporting (e.g., “receiving institutional data submissions in a timely manner” and “submitting student records... is relatively straightforward, but finding the time to process the results isn’t easy.”)
- Tracking students who leave, drop-out, or graduate from one institution presents a challenge. Institutions often lack the ability to follow students after they leave their institution; but a greater challenge is determining whether the students are “dropping out completely, attending another school, or working.”
 - Due to concurrent enrollment at multiple institutions, it is difficult to determine the student’s “home” institution, making retention calculations difficult.
- In a related concern, respondents reported issues in tracking student majors (e.g., “most students don’t declare a major until the junior year or later, so it’s hard to compare progress across majors,” “students’ major changes” or “internal transfers” [between departments] ,” NSC “does not indicate what program, major, or degree the student is enrolled in,” the “official major of a given student may not reflect the student’s intended major,” and students with “multiple majors”).
- The issue of data sharing creates an obstacle. Part of the difficulty of tracking students across institutions is the diversity of school systems (e.g., “Right now, there are too many silos of information,” “sharing data among institutions is a struggle” especially due to the “lack of cooperation between schools,” and many institutions are “restricted to only the institutional level data not system level”). This is further complicated by:

- Obtaining pre-college data (e.g., “The lack of readily available state data on individual high school and school district data in electronic format complicates getting a full picture and adequate context for student achievement at our institution”).
- Lacking accessibility of data (e.g., “writing appropriate queries and getting data from the state system” and “linkages between newly relevant data...”).
- Decentralized systems at the state and institutional levels.
 - In some cases, there is “no state-level SUR database.”
 - In other cases the state’s governing higher education office “does not make data available to individual institutions.”
 - There is also a lack of communication and data sharing across systems (e.g., between two- and four-year systems).
 - At the institutional level, some colleges have the ability to “track student retention across the institution without being able to track retention in individual departments.” Moreover, there is “incomplete coordination among separate tracking systems within individual school[s], such that... some data are only in an admissions database, others only in a class deans’ database, and others in the registrar’s database.”
 - Other options for obtaining student data include the NSC, “but that has limitations.”
- One of the challenges reported by many respondents related to data quality and obtaining accurate, consistent, and up-to-date information on students. In the absence of data sources, institutions often rely on student self-reported data (e.g., contact information, educational background, etc.). This is further complicated by:
 - Lacking consistent definitions of data elements or “standardization of variables.”
 - Difficulties and limitations of data input (e.g., “human error/omission at the data entry level leads to inaccurate data at the institutional level,” and “the [lacking] awareness of the importance of data accuracy by those who are tasked with collecting/entering the data on the front lines”).
 - Determining “who is in charge of the data and how much should be collected.”
 - Concerns over duplicate student records due to use of different unique identifiers (i.e., as a result of privacy concerns over the use of SSNs), and incomplete records/missing data (e.g., changes in names, IDs, marital status, etc.)
- Lacking resources also present a challenge (e.g., “the cost of software to keep the system upgraded and viable,” “the cost of starting and maintaining such a system,” lacking technology/data software, “not having a data warehouse,” and “antiquated student information systems” that entail manual data entry and reporting).
- Problems arising from the nature of the student body (e.g., tracking adult learners with “inconsistent attendance patterns,” foreign students who are not in NSC, knowing “who’s left for military service or missions,” and “non-credit students and the short-term nature of the classes/students”).
- Even with adequate data, the institutions struggle to analyze and understand the data in a meaningful way.
 - Institutions want to know “why applicants don’t come, why students leave before graduating, and what they do after leaving (graduated or not).”
 - Tracking data is not “in a form that is useful statistical modeling.”
 - “The manual nature of [the data often] prevents [it] from truly being useful.”

- The “lack of feedback received from the data submitted” makes the data entry process unfruitful to institutions.
- “Lacking the context of new and innovative strategies to help determine what worked and what didn’t” when students are not achieving desirable outcomes.

Additionally, some respondents indicated that they do not have any challenges or concerns.

Below are the quotations from some of these individuals:

- “We are fairly successful in querying data and providing longitudinal analyses, particularly for student success in developmental coursework and subsequent success in college-level developmental Math and English courses.”
- “Our student record system is semester based. We have created longitudinal databases to overcome it.”
- “We are small--1200--so tracking in house is not a problem.”
- “The IR [institutional research] offices in our system work very well together to provide useful system data as needed. We use [NSC] to track student movement outside of our campus.”
- “We have complete internal records and systems to track and study any and all topics of interest and importance. We graduate nearly all of our enrolled students, over 90%. And those we do not graduate, we know about very well as to both why they leave and what they do afterwards. We have more than adequate information on alumni outcomes.”
- “We have robust system perfectly able to track all of the 93% of our students who graduate. We have NO problems either with graduation rates or with studying everything we need to know from prospective students, through the admissions process, through enrollment and graduation. Alumni databases and surveys give us good outcomes data for our graduates.”

National/Regional Level

As a follow-up item to the closed-ended viability questions, respondents were asked to provide an open-ended response supporting their opinion on viability. Responses were provided by 382 of the surveyed individuals. For those who believed it was currently viable, there were themes that emerged regarding their beliefs about a national system. On the positive side of why and how a national system is currently viable, respondents felt that:

- The technology exists to make this possible (e.g., records are already electronic and states are already beginning to do this).
- It is already used at the state level (i.e., information collected at the state level can be “fed into a national database”).
- A form of this already exists through the NSC (e.g., NSC could be expanded or parallel system could be created).
- If there were a system, it would be beneficial to institutions, researchers, and students, as it would allow for better student tracking (e.g., given the mobility of students, this would provide a way of assessing student success when they have moved to a different institution) and is essential to serving students and promoting access and completion of higher education.
- It would help with Federal accountability and funding.

Though they believed it is viable, some still were concerned that it was not a good idea to move forward or could foresee obstacles in its creation. Below are the themes of these obstacles or arguments against forming a national SUR system, even though the respondents replied that it was currently viable:

- Viability is not enough; even though it is viable, it is not “cost effective,” “desirable,” and may not be “politically acceptable.”
- There would be problems with privacy issues (e.g., private colleges block efforts based on the issue of student privacy, the care by which these records are entered, and determining a student identifier that is not the Social Security Number).
- Data quality issues (e.g., lacking standardization or “data formats and definitions” in practice make it “difficult to compare across institution types given the difference in institutional missions across various segments,” there will be missing data issues if students refuse to have their records released, and data are not always correct at the time of reporting).
- Issues with the political will to fund this system.
- Issues with cooperation (i.e., “buy-in from state education departments,” local governments, institutions, and the public is needed and it a mandatory system would be required to make this possible).

Of the 294 respondents who answered the question regarding whether a national SUR system will *ever* be viable, 125 (i.e., 43 percent) said it would be viable in the future. Respondents provided areas that needed to first be addressed to make this possible, many of which mirrored the responses provided about obstacles in creating a current system (see above). However, these respondents felt these were surmountable. The areas to be addressed are as follows:

- Funding/Cost is currently an obstacle, but may not always be (e.g., in the current “economic recession... schools would find it difficult to cover the needed costs to do this” and there needs to be balance between the benefits of creating such a system and its cost; however, with funding, this was seen as a possibility in the future).
- Support for the effort does not exist pervasively yet, but may in the future (e.g., in the future, if the public, institutions, states, and policymakers support this, it will be possible; the emphasis will need to be on data sharing and not “punishment” of institutions based on their data; it is possible that a higher education crisis or change in leadership of associations may make support for this possible).
 - In addition, respondents provided insights into the notions which work against this movement; for example, the idea that making a system separate from NSC would be redundant. Also, as long as a SUR system is not seen as useful to the institutions, support will lack (i.e., the institutional “response is conditioned upon the ability to ‘disclose’ data back to institutions; that was a major issue last

time [the database was seen as a one-way street where institutions provided data but not receive data back in a meaningful way]”).

- Burden continues to serve as an obstacle (e.g., reporting burden on the institution will need to be overcome; or a mandatory system will need to be put in place).
- Technology needs to be improved before a SUR system is viable (e.g., “As longitudinal data systems mature... [and become] more widespread, it will be viable and politically feasible”).
- Data quality needs to be improved before a SUR system is viable (including consistency in data collection, the need for “better quality assurance measures...to ensure consistent, reliable data reporting from all sectors,” and the need for a software vendor to help develop a tool).
- Individual states and institutions need to establish systems first, and a national system can be developed from there.
- Issues with privacy need to be addressed first, but may become of less concern with time (e.g., current generations are already “much less concerned about privacy than were earlier generations, so it is possible that...a national [SUR] database might be politically acceptable [in the future];” the data elements collected and how it is shared would determine whether privacy is a concern; “Security and data ownership issues must first be worked out” and “will eventually be worked out”).

Of the 294 respondents who answered the question regarding whether a national SUR system will ever be viable, 169 (i.e., 58 percent) said it would not ever be viable. Interestingly, the obstacles mentioned were the same as those who said it would be viable; the difference is that this group sees the obstacles as insurmountable. Below are the themes of responses and quotes from the survey responses:

- Lack of data quality or ability to standardize and ensure quality will keep it from ever being viable:
 - It will be “too difficult to mandate consistent data standards ... [or get] the data entered in a timely fashion.”
 - There will be too many “opportunities for misunderstanding of requirements, delay in changing to new specifications, or simple error” and these complexities could not “be reconciled at a national level.” Similarly, another respondent stated that, “We have spent an inordinate amount of time during this last IPEDS report cycle responding to edit exceptions and queries that produced no real improvement in the data. This has not been a trivial expense. Given the real cost of a national system one has to ask how much it would contribute to our understanding and that understanding to a change in behavior versus other less expensive means of knowing.”
 - “The only way that a national unit-record system would work would be if the organization collecting the data...were to understand that the system was a reporting tool (not an audit tool) and that a certain amount of error and uncertainty/incompatibility will occur; given our experience with our state system, I believe that a [national] system would collapse under its own weight.”
 - A sub-issue is that of differences across states and institutions and the challenge in creating a national system which can fit “a wide variety of institutional types” and can account for differences in coding of variables.

- Cost of and staffing for a national SUR database:
 - Institutions report that they “already [have] problems with faculty's concerns about hiring staff for multiple reporting requirements instead of more faculty.”
 - It is felt that this would result in an “unnecessary administrative burden on institutions yielding little value to a student's learning experience while driving up cost to student to access educational experience. “
 - A couple of respondents offered an alternative (and what they say as, less expensive) option: “If we as a country need to find out what students are doing, then longitudinal sampling is less intrusive and infinitely cheaper (i.e., beefed-up versions of NELS, BPS, etc.)”
- Return on investment is perceived to be low, therefore making a national SUR database not viable (e.g., “States are already in progress with their systems and have invested large sums. I'm not sure how much thought has gone into cross platform... We submit lots of data at the state level and don't get access or good reporting back for our own use. I'm not seeing the ROI [return on investment] on a national system at this point”).
- Lacking support or backing from members of the higher education community, particularly private institutions (e.g., “as a private institution we would not be willing to give student level data on a national level any more than we would on the state level. We are private--and our data is private”) and policymakers (i.e., lacking “political will”) is seen as precluding a national SUR system from being instituted.
- In addition, issues with privacy and “fear of misuse of data” keep the public from supporting additional government control of individual student data. While it “is technically possible [to do this] (see social security administration and credit score companies)... [it] does not mean it should be [done]. It is not clear that the benefits outweigh the risks and costs.”
 - The main cause for concern is that the Social Security Numbers would need to be used as the identifier to link students across institutions and databases; the “use of which would jeopardize identity security.”
- Finally, there were sentiments that a national SUR system is unnecessary (e.g., “Clearinghouse does a nice job already” and “basic data (identification, demographics, enrollment, [and] graduation) is already largely available in the Clearinghouse.”)
 - There were a couple of negative sentiments about even talking about a national SUR system because it seen as taking focus away “from critical issues, such as learning” and is something that is seen as impossible to do:
 - “The nation, including AAAS and Sloan simply HAS to QUIT thinking [about] and exploring this idea...they simply cannot approach the time and personnel on campuses everywhere that will be required to produce this imagined system.”
 - “It is simply TOO MUCH work for colleges & universities to be asked to do. We are eliminating staff, cutting back academic programs and majors, and still we cannot balance the operating budget given the great losses in our endowment. We have NO resources, personnel or capital, to dedicate to this idea. We are already required to report more data to external agencies than AAAS and Sloan/EPI really know of and appreciate. At this time especially, we simply cannot be asked to engage in a national unit-record enrollment base. It does have some, a few, benefits, but these are far, far out-weighed by what would be the inevitable costs required of all colleges and universities to produce it.”



www.educationalpolicy.org

Copyright 2018