

## Linking Student Enrollment Records Across State Lines: Extending the Concept

National Center for Higher Education Management Systems (NCHEMS)

Supported by a Grant from the Lumina Foundation for Education

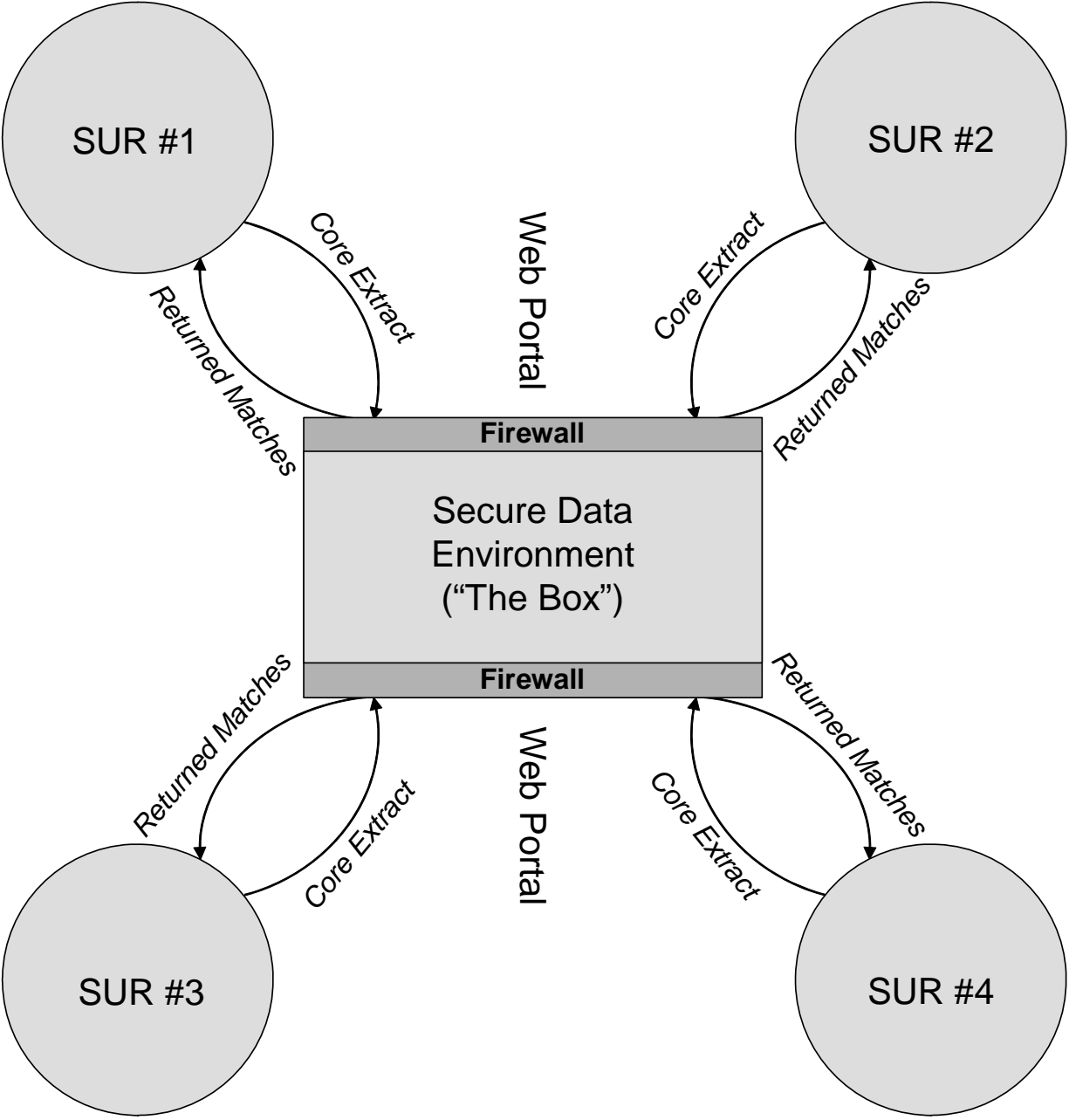
**Background.** Multi-lateral exchanges of data drawn from state-level student unit-record systems (SURs) hold considerable promise for enhancing policymakers' and researchers' understanding of how students flow through the increasingly complex postsecondary educational "pipeline." Because most students now attend multiple institutions in order to earn a degree and many cross state lines in doing so, constructing a comprehensive picture of longitudinal enrollment behavior requires drawing data from multiple sources and housing these data in a secure environment capable of supporting sophisticated data analyses. Supported by a grant from the Lumina Foundation for Education, NCHEMS examined the data contents of all extant state-level SURs and, as documented by a report published by Lumina in April of 2003 entitled *Following the Mobile Student*, determined that they contained the requisite data elements to support such an approach.

To begin to investigate how such data might be linked and analyzed, NCHEMS convened a technical panel comprising researchers and individuals responsible for administering and using state-level SURs in November of 2003 (see Appendix for a list of meeting participants). Although the meeting raised and discussed a number of policy issues (addressed in a companion report), its principal purpose was to begin to outline methods for accomplishing the contemplated data exchanges and to discuss the many technical issues and options associated with this task. This brief report describes the outcomes of this discussion. A first section of the report summarizes the basic approach to the task supported by the panel. Subsequent sections then describe in a bit more detail specific features of the contemplated approach.

**Basic Concept.** Participants in the meeting quickly reached consensus that the best way to think about the many design issues associated with the development of a multi-state data-exchange capability would be to conceptualize it as a single centralized database containing data extracted from multiple "source" SUR databases. Consistent with the conclusions of a parallel interview-based study of policy leaders conducted by NCHEMS, however, members of the panel felt strongly that the actual implementation of such an approach should be undertaken incrementally, and be advanced under the rubric of "data exchange" rather than "building a centralized database." They also agreed strongly with policymakers that such data exchanges eventually ought to include data drawn from K-12 data systems and employment records.

The basic concept for accomplishing such an exchange is quite straightforward, as portrayed by Figure 1. Its basic elements are as follows:

Figure 1  
Basic Concept for Data Exchange



- Multiple SURs maintained independently by participating states. These will necessarily differ with respect to data contents, data structures, and operating environments.
- A secure data environment within which records can be matched to determine if the same student is present in multiple SURs (“The Box”). Selected student record data supplied by individual participating states and drawn from their SURs would enter this environment through a secure Web-based access portal and matching would take place behind a firewall. Alternatively, this environment might consist of a particular state’s SUR itself, so long as it has the requisite size and capabilities including secure Web-based access.
- A set of core data elements defined and coded in a standard format that constitute the substance of the records to be exchanged. Only a few such data elements will be needed to make this approach viable (see section on Data Elements below). These would describe the “dependent variable” of interest in such exchanges—in essence whether and how a given student is enrolled in or has completed a degree at an institution covered by another SUR database. “Independent variables” of interest, in contrast, will be much more extensive and varied but will remain within the individual SURs of participating states. Examples might be as varied as demographic descriptors, receipt of state scholarship funds, and participants in particular programs of interest to the state. After matched records are returned to participating states, they can be linked with any desired combination of “independent variables” of this kind within each state’s own SUR environment with no need for common data structures across states.
- An input protocol for participating states to submit core data elements. Under full-scale implementation, all participating states would submit records for all students with active records contained in their SURs on an at least an annual basis. Under more limited forms of implementation such as a bi-lateral exchange, states might only submit records for the students they want to match.
- An output report for participating states that returns the results of the matching process. Ideally, this would take the form of a unique unit record that contains the core data elements, generated for each student enrollment event involving an institution outside the original source state. For students who enrolled in three different institutions outside the original source state, for example, three such records would be generated and returned to the source state.

Such an approach has the advantage of keeping the amount of data actually exchanged to a minimum and of allowing each participating state to perform its own analyses on the augmented outcomes data generated by the matching process.

**Modular Design.** Members of the panel agreed that the best approach to realizing this concept would be through a modular design under which additional, more sophisticated, features could be added to the basic model outlined. The flexibility accorded by such an approach would allow states with quite varied capabilities to participate at different levels. It would also ensure consistency over time as the system evolves because basic procedures, data content, and data definitions could be maintained as additional modules allow overall system capabilities to grow. A first such enhancement might be one or more expanded sets of exchanged data elements to be added onto the proposed core. Another might allow the inclusion of term-level detail instead of annualized reporting. Similar, more sophisticated add-on modules might include a K-12 matching capability and a UI-wage record matching capability to enable more comprehensive longitudinal studies of the “educational pipeline” to be undertaken. Other enhancements might address reporting capabilities or tools to make matched data more accessible to institutions that want to use them in their own planning and decision-making. Such additional reporting capabilities might include standard high school or transfer institution feedback reports on subsequent student performance in a subsequent institution.

**Data Element Contents.** As noted, the actual number of data elements that need to be exchanged under this approach is quite limited because most analysis will take place within the SURs of each participating state after matched records have been returned. For the core set of data elements, the panel agreed that the

following should be included, reported on an annual or fall-term basis consistent with the reporting schedule for the Integrated Postsecondary Education Data System (IPEDS). States not able to generate this minimum set of data elements would be unable to participate in the exchange.

Record Identifiers. Three distinct record identifiers would be needed as key links for record matching:

- *State/Agency Code.* This would indicate the particular SUR database from which the data are drawn (which might be a single statewide SUR database but might equally be an SUR database associated with a particular multi-institutional system in a state that has several such systems).
- *Institution Code.* This would indicate the institution associated with each particular enrollment event identified. The federal Unit ID code used in IPEDS is recommended.
- *Student Identifier.* This would indicate the student associated with each particular enrollment event identified. Because most extant state SURs employ the student's Social Security Number (SSN) as a unique identifier, using some form of the SSN will be unavoidable. Using the SSN as a link element will also allow matching with non-SUR databases like the UI wage record. But the actual code contained in the record should be an encrypted version of the SSN to protect privacy (see Section on Privacy and Security below).

Basic Demographics. Because most demographics of interest for disaggregation will already reside in the SURs of participating states, these elements will be kept to a minimum. They are included largely for cross-checking and to provide an alternative route for matching records. The following elements are recommended:

- *Sex.* Standard IPEDS definitions and codes.
- *Race/Ethnicity.* Standard IPEDS definitions and codes,
- *Date of Birth.* In DD/MM/YYYY format.
- *Citizenship/Visa Status.* This code would be used to flag U.S. Citizens.

Enrollment Status Descriptors. These elements describe aspects of each enrollment event associated with a given student and a given institution contained in any of the participating SURs during the time period of interest. They include:

- *Degree-Seeking Status.* Standard IPEDS definitions and codes.
- *Enrolled Load.* This element is intended to reflect the intensity of the enrollment with respect to student time. At minimum, Full-time/Part-time Status could be used employing standard IPEDS definitions. Alternatively, the element could specify percentage increments of a standard full-time load (e.g. "less than 25%," "25%-50%," etc.) based on the number of credits actually enrolled for.
- *Program(s) of Study.* Standard IPEDS definition using the Classification of Instructional Programs (CIP) code. Provision should be made of up to three instances to reflect student enrollment in multiple programs of study at a given institution.
- *Credential(s) Earned.* Degree levels as defined by IPEDS. Provision should be made for up to three instances to reflect student enrollment in multiple programs of study at a given institution.
- *Field of Study for Each Credential.* Standard IPEDS definition using the Classification of Instructional Programs (CIP) code.

For a first enhanced data element module, members of the panel suggested the following additional Enrollment Status descriptors:

- *Residency Status*. This element is intended to reflect residency for tuition purposes.
- *Student Level*. Standard IPEDS definitions reflecting current level of study (e.g. 1<sup>st</sup> Year, 2<sup>nd</sup> Year, 3<sup>rd</sup> Year, 4<sup>th</sup> Year, Post-Baccalaureate, Masters, Doctoral).
- *Credits Attempted (Term)*. The actual number of credits enrolled for during the term [for those reporting term-detail data].
- *Credit Type (Term)*. The credit system used in calculating the above (e.g. semester, quarter, clock-hour) [for those reporting term-detail data].
- *Distance Education Flag*. This element is intended to reflect students enrolled in programs that are entirely delivered at a distance through correspondence, broadcast video or audio, or over the Web.

The panel also suggested a second enhanced data element module containing annualized information on financial aid status. This would include the following:

- *FAFSA*. This element would simply flag whether or not a student had a FAFSA application on file.
- *Need Level*. Standard FAFSA definition.
- *Dependent Status*. Standard FAFSA definition.
- *Source(s) of Funds*. This element would reflect which sources of financial aid were tapped by the student during the year of enrollment. Categories and codes are not specified but should include a) federal need-based grants, b) federal loans (all types), c) other federal scholarship support (e.g. NSF), d) state need-based grants, e) state merit-based (or other) support, f) institutional aid (if available).

**Reporting.** As noted, the primary output of the matching process would be unit-record data files reflecting each instance of enrollment in an out-of-state institution detected among participating states. In addition, aggregate reports would be developed that summarize this information for each institution in each participating state. These would show enrollments and program completions by field of study, load, and selected demographics. Additional reporting modules could be developed to reflect outcomes obtained from such external data sources as the UI wage record. All such reports should be generated on a schedule appropriate to inform the budgeting/appropriations cycle in each participating state.

**Security/Privacy.** Each participating state presumably already has guidelines about security and privacy that would govern procedures within its own SUR. Security procedures for operations within “The Box” and appropriate firewalls would be established that are at least as strong as those now in place for established multi-database warehouses like the one maintained by the Florida Office of K-20 Education Information and Accountability. Records would be transmitted in and out of “The Box” through secure Web portals that are accessible only to registered participants in the exchange.

The panel agreed that all participating states should sign a Memorandum of Understanding (MOU) [or, alternatively, a performance contract with the entity responsible for conducting the matching procedure] that specifies limits on the use of data and establishes guidelines for conducting the exchange consistent with “best practice” regarding FERPA and other applicable federal and state laws addressing the privacy of records. These guidelines will emphasize that unit-record data are being merged only for the purposes of

legitimate research to improve educational policy and delivery, and that no information about individuals will be disclosed. The panel also emphasized that participating states would benefit greatly from a standard “toolkit” on how to address FERPA and other privacy/confidentiality issues containing examples of the requisite public statements, consent forms, and other documentation. All aggregate reports generated through the exchange should also automatically block the generation of any cell in a report with fewer than a specified number of cases (from three to five).

Members of the panel agreed that there was no immediate alternative to using SSNs as unique identifiers for matching records, but believed that SSNs should be encrypted for transmission to and from “The Box.” This could be accomplished via a standard encryption routine or through a random assignment of SSNs to specially-created record numbers with a key file kept in a secure location. Back-up methods for matching records based on combinations of other data elements or (less satisfactorily) on the basis of additional “Directory-Type” information (e.g. Name, Address, Date of Birth) associated with each record should also be explored.

Finally, since the matching process is in each case a one-time event, the base data inside “The Box” might be purged after the matched records have been returned to each participating state. This would provide maximum security for such records, but would prevent any possibility of additional analysis or matching unless updated records were again supplied. On balance, the panel felt that preserving data for a five to six year period was preferable, though the provision of deleting all of a given state’s records after matching might induce reluctant states to participate.

**Organization and Governance.** While the panel spent relatively little time on this matter, its consensus was that a third-party entity should conduct the match on behalf of all participating states (although one state among those participating might in fact fulfill this function). This entity should clearly serve as the “agent” of the participating states and not have independent authority to establish data standards and policies. These would instead be established by mutual consent and be governed by a committee of participants. Similar arrangements should be used to establish policies regarding data access, how the quality of data is assured, scheduling of data transmission and reporting, and similar matters. But the bottom line for the panel was that, whatever the details of eventual governance, participating states would have to feel that they collectively “owned” the operation.

A number of different types of organizations were discussed as possible “agents” to serve in this capacity. They included the following:

- The Federal Government through NCES. Not much support was given to this alternative because of the perceptions that creating such a “national database” would create within the higher education community. But the federal government was seen as a potential source of funding for to help establish such a capability.
- A SHEEO Agency Participating in the Exchange. This was a popular alternative, because the SUR housing this state’s data could itself serve as the secure environment in which matching could take place. Any agency taking on such a role, however, would have to have an SUR with Web-enabled relational database capabilities.
- A University Research Center on Higher Education. This alternative was seen as moderately attractive because it might involve members of the higher education research community in retention scholarship.
- An Independent Contractor. This might be a newly-created non-profit corporation, an interstate compact or regional higher education organization such as SREB or WICHE, or an established data provider like ACT (which currently runs data-gathering operations on behalf of clients like the NCAA).

Whatever entity eventually plays such a role, participants agreed that costs of the operation should be underwritten by a standard participation fee, supplemented by a schedule of fees for additional services. In

addition, panel members felt strongly that states would benefit the most from participation in such exchanges if they were supplemented by an infrastructure of conferences, seminars, and training opportunities that would see to what one participant called “the care and feeding of the Network.” The value of such activities has been clearly demonstrated by institution-based data-sharing consortia like the AAU and HEDS.

**Costs.** The panel estimated that annual operating costs associated with running the basic matching service would be about \$500,000. This could support a staff of four or five to conduct basic matching services for any number of states. Start-up costs would probably constitute another \$250,000.

**Conclusions and Next Steps.** In short, the panel concluded that accomplishing the required data exchanges was relatively straightforward from a technical standpoint and could be accomplished fairly cheaply. But they emphasized that participating states would have to be convinced that the value of participating would outweigh the costs—both monetary and political—that might be associated with doing so. As a result, a strong “case statement” would need to be developed that outlined these benefits and, particularly, provided assurances to participating states about issues of confidentiality and privacy that will surely arise as institutions are involved in discussions about how to move forward.

The panel also agreed that limited experiments in data exchange among a few states would be very helpful in both proving the concept and in developing prototype products to guide the implementation of the basic model described in this report. Such multi-state experiments should, at minimum:

- Assess the actual “added informational value” gained by supplementing currently-available data with out-of-state enrollment records for such measures as program completion.
- Uncover any issues associated with the proposed list of core data elements and establish consistent working definitions for all that are not already standardized.
- Demonstrate that student records can be exchanged in an efficient manner that complies fully with current regulations regarding confidentiality and privacy (including developing the necessary protocols, encryption routines, etc.).
- Develop prototype “best practice” approaches to reporting and analyzing longitudinal data.
- Experiment, if possible, with adding data from data sources beyond higher education such as K-12 and employment databases.

If these objectives can be accomplished through a few multi-state demonstration efforts, members of the panel believed that many states would see the value of expanding participation.

## Appendix

### Lumina Student Tracking Technical Issues Meeting Participant List

Emerson Elliott  
Director, Program Standards Development Project  
National Council for Accreditation of Teacher Education (NCATE)  
2010 Massachusetts Avenue NW,  
Suite 500  
Washington, DC 20036-1023  
Ph: 202-466-7496  
Fax: 202-296-6620  
Email: [emerson@ncate.org](mailto:emerson@ncate.org)

Charlie Lenth  
Senior Associate  
State Higher Education Executive Officers (SHEEO)  
700 Broadway, Suite 1200  
Denver, CO 80203-3460  
Ph: 303-299-3685  
Fax: 303-296-9016  
Email: [clenth@sheeo.org](mailto:clenth@sheeo.org)

Joseph L. Marks  
Director, Education Data Services  
Southern Regional Education Board (SREB)  
592 10<sup>th</sup> Street NW  
Atlanta, GA 30318  
Ph: 404-875-9211  
Fax: 404-872-1477  
Email: [joe.marks@sreb.org](mailto:joe.marks@sreb.org)

Tod Massa  
Institutional Research & Technology Services Director  
State Council of Higher Education for Virginia (SCHEV)  
101 N. 14<sup>th</sup> Street  
James Monroe Building  
Richmond, VA 23219  
Ph: 804-225-3147  
Fax: 804-225-2604  
Email: [TodMassa@schev.edu](mailto:TodMassa@schev.edu)

John Milam



Managing Director  
HigherEd.org, Inc.  
9 West Picadilly Street  
Winchester, VA 22601  
Ph: 540-722-6060  
Fax: 540-722-6163  
Email: [jmilam@highered.org](mailto:jmilam@highered.org)

Rhea Santos  
Assistant Director  
Information Systems  
Illinois Board of Higher Education  
431 East Adams, 2<sup>nd</sup> Floor  
Springfield, IL 62701-1418  
Ph: 217-557-7360  
Fax: 217-782-8548  
Email: [santos@ibhe.org](mailto:santos@ibhe.org)

Debra Stuart  
Executive Director State System Research  
Oklahoma State Regents for Higher Education (OSRHE)  
655 Research Parkway, Suite 200  
Oklahoma City, OK 73104  
Ph: 405-225-9176  
Fax: 405-225-9230  
Email: [dstuart@osrhe.edu](mailto:dstuart@osrhe.edu)

John Wittstruck  
Senior Research Associate  
Education Policy, Planning & Improvement Center (EPPIC)  
Coordinating Board for Higher Education  
Missouri Department of Higher Education  
3515 Amazonas Drive  
Jefferson City, MO 65109-5717  
Ph: 573-751-2361  
Fax: 573-751-6635  
Email: [john.wittstruck@mocbhe.gov](mailto:john.wittstruck@mocbhe.gov)

David Wright  
Senior Research Analyst  
State Higher Education Executive Officers (SHEEO)  
700 Broadway, Suite 1200  
Denver, CO 80203-3460  
Ph: 303-299-3677  
Email: [dwright@sheeo.org](mailto:dwright@sheeo.org)

## **Lumina Project Policy and Technical Agenda**

### **Site Visit to Tallahassee 4/15/04**

Peter Ewell and Marianne Boeke from NCHEMS visited Tallahassee to gather information on the operations of Florida's resource for linking state-level unit record information directed by Jay Pfeiffer. The office responsible for this resource has been through several iterations since experiments with linking UI wage records to postsecondary student records began in the mid-1980s in order to determine the effectiveness of employment training programs. It is now a comprehensive state-level resource called the Office of K-20 Education, Information, and Accountability. Its two main operations are a) the long-standing Florida Education and Training Placement Information Program (FETPIP) follow-up databases that create longitudinal records for all students leaving Florida public institutions to track further education and job placement and, b) the newly-created K-20 Data Warehouse that integrates all available information about students and former students in a relational database environment. Data contained in these databases are drawn from the state Education department (which collects district-level enrollment and performance data), the Community College System, the University System, records from the Florida's postsecondary financial aid program, the state UI-wage record system, military records, U.S. Post Office records, and (most recently) the National Student Clearing House.

#### **History**

Florida's unit record systems began like most others to drive funding formula allocations to institutions. They contain data drawn from system-level databases that were developed for this purpose and that continue to be maintained, so Jay's operation is entirely comprised of "second-order" data—they do not collect any data directly from institutions.

At the system level (e.g. the Community College System), institutions were at first somewhat resistant to the creation of an SUR database—"we began with an adversarial relationship and we had to sell it." The key to making this work was creation of an active task force of institutional representatives from all participating colleges, which operates largely on a consensus basis to provide guidance to the system.

FETPIP arose from the need to track Carl Perkins program outcomes. FL had been doing this the way everybody else did it at first, through survey follow up, then became an early-adopter of the UI-wage record link (along with TX and the University of MD). It was funded from the education side, but housed in the Department of Labor, with substantial "oversight" from the education side to ensure FERPA compliance.

The Independent College Association of Florida (ICAF) was at one time interested in being involved under past leadership, but is not so much lately. The original plan was for ICAF to coordinate data collection, then work with Jay's operation. Meetings with Presidents have been critical here to get buy-in and support. Independent colleges are included in FETPIP, reporting their graduates to the state annually [not sure if this is done independently to the state or through ICAF].

FL did create a legal framework within which this could happen, and this clearly has been an advantage. There is legislation requiring the production of particular kinds of performance measures for accountability and planning purposes—this is the background to creating Jay's operation. In addition, the K-12 system is governed by legislation that mandates the use of SSN in record-keeping and reporting (though this is currently under threat). But Jay downplays the need for this and feels that just as much could be done in states that lack a statutory framework. For example, he noted that the FL legislation includes dollar withholding for institutions that are not in compliance with data reporting provisions (at least for CCs), but this authority has never been used. Another important condition is the fact that the state auditor's office has the authority to audit the accuracy of institutionally-submitted data because it is used to make funding decisions. It is clear, though, that this formal legal framework sits in the background and is an important motivating factor even through "we work mostly through cooperation, not under mandate."

The K-20 Data Warehouse is a more recent development, and contains relational data from all three systems (K-12, CC, and University) together with a range of outcomes data. Independent institutions are not in the Warehouse. The recent K-20 governance change helped the creation of the Warehouse a lot because it provided a legal rationale for one K-20 database. Jay believes that it would have been able to integrate K-12 data without this development, but it certainly made the politics easier.

FL has had some limited experience in linking with other state SUR systems. Did this with the Georgia Board of Regents under an agreement.

### **Organizational Development**

Jay emphasized that an enterprise like this is "never finished," and that the continuing existence of even his own well-established operation can never be taken for granted. You always have to be making friends and keeping bridges mended, and you never know when something will come up that might change everything. An example cited was a bill in the legislature this year about banning the use of SSNs, advanced by the state student association. It looks like this is under control now, but it caught everybody by surprise and it has made everybody "gun-shy about SSNs in the current political climate." "We just never stopped," because every time there is a turnover (mentioned particular impact of term limits here) "you have to justify your existence—in an atmosphere that is pretty sensitive" about state offices, costs, and privacy issues.

The key to making this work is to continually emphasize "value-added" products that are of demonstrable use to somebody in helping them do their job. Simply mandating things doesn't work very well, even though they have the legal authority to do so. "Institutions are reluctant players in [integrated unit record database development] at first, and systems even more." Systems are particularly worried about turf issues when it comes to data—it's about who has control. And if institutions don't want to play, just leave them out until they begin to see the benefits of being included. You want to continually "touch people" with the value of the data you produce—especially longitudinal data—rather than emphasizing a legislative requirement. It is also sometimes a good idea to work closely with "lead institutions" in this process that will see the value in what you are doing and will be willing to put some effort into being a good prototype.

Another key to making this work is to have a constant set of "champions" in several places that will advocate for it. FL had less need of this because of its tradition of intrusive data reporting through legislation ("government by proviso"), but other states that do not have this tradition will need a network of champions. Jay cited the development of Ohio's system as being largely because of the SHEEO, Rod Chu. He saw this as both an advantage and a disadvantage because it is so tied to one person's initiative and very connected to his own SHEEO agenda, rather than a statewide agenda.

Jay emphasized the need to get formal agreements that spell out obligations clearly at all points—“we do agreements with everybody.” In FL there is no need to have a data exchange agreement with another state agency, but they do this with any other third party. Sometimes they even do it with another state agency—for example, FETPIP has an agreement with the Community College System “to clarify responsibility and to provide an audit trail.” It is always best to have the legal people involved with the party to an agreement at an early stage—both so they know and so they can spot potential problems. It is also important in these agreements to re-negotiate them annually on the basis of performance and delivery, in order to emphasize mutual accountability. Jay cited their agreement with the U.S. Postal Service as being the best general example to follow.

For multi-state data exchanges, the notion of having one SHEEO agency act as an “agent” for others certainly does have advantages. But one major potential drawback in pursuing this option is that the host agency would have access to data that others might not. There is a consequent need to craft the agreement in a manner that would preclude this—so that the host agency could not benefit in ways that others could not. [Pay particular attention to this for the KY-OBR agreement.] Does see some advantages to an interstate compact approach for this reason, but doubts that agencies like SREB or WICHE have the technical capacity to house and operate databases of this size and complexity.

A good model to follow in state-level exchange might be the federal pilot of the Wage Record Information System (WRIS). This is a national distributed exchange of UI wage record information among a number of participating states. It is not a national database per se, but rather a way of linking existing records [note: need to check this approach as a model because it sounds a lot like our decentralized “portal-based” approach discussed as the minority opinion in the Technical Issues meeting]. There is also a separate wage-record exchange initiative being coordinated by state Workforce Investment Act (WIA) offices that involves FL, OH, KY, PA, VA, WV, and MD.

### **FERPA and Privacy Issues**

The K-12 system had a FERPA lawsuit (Lake County) early in the process of integration of its unit-record data system. The outcome was a ruling in favor of collecting these data on the grounds that state law required it.

Jay sees most FERPA and privacy-related questions as not being substantive, but about “agency turf.” This is especially true in working with source systems, but if you can “work it out with them so that you are not getting anything from institutions” directly, this addresses a lot of everybody’s apprehensions.

Jay believes that there is no problem with the release of institution-level information contained within a system-level database to another third party like his office so long as it is clear that the second-order user (his office again) is clearly designated to be acting as an “agent” of the first through a written agreement. He believes that there is no requirement to go back to the institutions to get yet another release of the data to be used for this purpose. He does, however, believe that consulting with affected institutions—or maybe in some cases entering into a Memorandum of Understanding (MOU)—is always good politics.

Also sees no problem with FERPA’s “data destruction” requirement so long as the uses of data are for the “exempted” purposes of conducting educational research to improve programming. By definition, longitudinal work requires the indefinite preservation of some records and as long as the purposes require tracking, you are not “finished” with the study and do not have to destroy the records.

Jay emphatically believes that we should not be proactive about asking for FERPA clarification from the Family Policy Compliance Office or from individual state AGs offices unless absolutely necessary in a state. The responses will simply be ambiguous and will be written so that the person involved bears no responsibility. And putting the resulting ambiguity in writing will simply serve to make everyone more nervous.

With regard to Reauthorization, Jay also advocates keeping a low profile and not trying to get greater federal guidance. Worries about “homogenization” involved with any federal reporting framework that would get in the way of state efforts that are already far in advance of anything that the DOE would be able to do. He is particularly worried about the OMB Performance Measures initiative that may mandate worse measures for things like job placement for Perkins reporting than the ones FL now has in place.

But there may be a role for the federal government in establishing the proper “framework” for records confidentiality policy on a consistent basis. Jay cites the Computerized Records Matching Bill as a good example of helpful federal policy because it gives guidance and makes important distinctions without getting into the specifics of how things are done.

All data in the K-20 Warehouse are “anonymized” on entry into the system. Any data provided by a source system as an extract go two directions. The first is to a System Identification Database (SID) which determines if the record is a new person by comparing SSN and other directory information to records of people already in the database. This is done through 17 “business rules” that define various kinds of identification matches. If the person is not found, a new internal record identifier is assigned. If the person is found, the existing internal record identifier is passed to the Warehouse for linking with the new extract information. In the second path, the SSN is stripped from the extract record and replaced with a sequential code, which provisionally identifies the record in the Warehouse until the appropriate internal identifier (either new or existing) is determined. When this is done, the new information is then fully incorporated into the Warehouse.

### **National Student Clearinghouse**

Florida is working with the National Student Clearinghouse (NSC) to obtain data to load both FETPIP and the K-20 Warehouse, via the NSC’s reporting utility EdEvaluator. They can only do this for the state as a whole to drive their aggregate reporting, and cannot supply any of the resulting information to institutions. Institutions and systems have to approach the NSC on their own to get this information. This is an awkward arrangement, but anything else would destroy NSC’s ability to sell services to institutions.

They have an agreement with NSC that allows the use of the SSN because NSC is acting as “an agent of the state” for FERPA purposes. It is an “open use” subscription that allows unlimited use of the service for a particular period of time, with initial cost based on fall enrollment statistics. This does have the advantage of being able to construct and update cohorts based on repeated data matches over time. The state is essentially treated as one big institution under this agreement.

Jay’s initial reaction to the Lumina Project is that it would essentially duplicate the NSC, and thought that one of the main benefits of the NCHEMS effort would be to put market pressure on NSC so they would lower their prices. But noted that Clearinghouse does have other problems with the periodicity of data receipt which does not match term structures. Data are collected quarterly, and how Add/Drop is handled is unknown because it is essentially up to institutions to report. Also, consistency of data collection and definitions may vary a lot from institution to institution. Jay’s office has done some comparisons of the status of students matched in both NSC and FETPIP, and found that the two databases don’t agree in some cases—so this is worrisome as well.

In terms of data contents, noted that NSC does have a full-time/part-time data element, but this only refers to the student’s status when they first entered the institution.

### **Possible NCHEMS Center Proposal**

Jay and colleagues were quite supportive of the notion of an NCHEMS Center on State Policy and Student Progression. Jay especially noted that we might want to look at the structure and services of the Gartner Group as a model for ideas in developing the concept. The following were the kinds of activities that such a Center might engage in:

#### Facilitating Data Exchange:

- direct assistance in putting networks together as an extension of current Lumina work
- providing model agreements and organizational arrangements, including the FERPA/Privacy toolking
- providing model data definitions and best practices

#### State Unit Record Systems:

- Facilitate joint system development to address aging infrastructure and lack of personnel investment
- Model report development
- Research ideas and dissemination
- Emphasize nuances of data analysis and interpretation under different context (e.g. Hispanic performance and behavior very different in FL and TX/CA because very different populations involved)
- Regularly convene people responsible for these systems at the state level

#### Policy Development:

- Best practices in such areas as articulation and transfer
- What does and doesn't work in state policy to enhance progression
- State benchmarking on outcomes and policies (where to go for model states)
- Annual survey of states about what they are doing in terms of policy and about latest SUR status and initiatives
- 

#### Technical Assistance:

- In establishing data exchanges and SUR renewal
- Third party review of proposed new state initiatives to help them determine if what they intend makes sense and/or if anybody else has already done it
- 

### **Linking Student Enrollment Records Across State Lines: Policy Issues**

National Center for Higher Education Management Systems (NCHEMS)

Supported by a Grant from the Lumina Foundation for Education

**Background.** The need for more complete measurements of student success in the nation's colleges and universities is critical today. Unfortunately, current statistics on enrollment and college completion don't tell us enough about the details of student achievement to allow us to develop effective policies for student success. One problem is that the only national data on student retention and graduation are self-reported—either through ACT or the more-recently established NCES Graduation Rate Survey (GRS). State Unit Record (SUR) databases represent one way to circumvent the limits of self-reported data because they include individual records for every student enrolled in the colleges and universities they cover and can thus track student transfers from one institution to another. But state UR databases are currently limited by the fact that they only contain information about students enrolled in a particular state. And we know from national studies that over 54% of students who eventually receive a baccalaureate degree have attended two or more institutions, and 40% of these transfers crossed state lines. To obtain detailed estimates of actual persistence and completion under such circumstances, it would be necessary to link UR databases located in multiple states.

The Lumina-funded report *Following the Mobile Student* published in April of 2003 demonstrated that such linkages are feasible and that extant SUR databases contain information on a substantial proportion of the nation's students. This study identified and analyzed 46 SUR databases located in 39 states, and

established the fact that these databases contain information on at least seven core data elements covering a total of 74% of current undergraduate students. Although intended originally to house data about students enrolled in public colleges and universities, ten of the SUR databases identified contain information about independent college students as well, and these numbers are growing. And at least one state independent college association is establishing its own SUR database on a consortial basis. These findings constitute a solid foundation for pursuing the idea of linking SUR databases to develop a more detailed and comprehensive picture of student success and, more importantly, of the experiences of different kinds of students and the factors that lead to success. But translating this concept into reality is likely to raise numerous conceptual and policy issues. State record-keeping systems have been developed piecemeal over the past two decades and, because they were originally designed for quite different purposes, are organized and governed in different ways. And harnessing them for a larger national purpose requires convincing both state authorities and institutions that the resulting gains in information will be worth both the costs involved and the political risks entailed in sharing data.

To explore these issues more fully, NCHEMS conducted a series of interviews with selected policy leaders in the fall of 2003 including representatives of national organizations, state higher education agencies, and the policy research community (a list of those interviewed and the interview protocol used are presented in Appendix A). The majority of these interviews were conducted face-to-face in an hour-long format, and a summary was prepared for each. In addition, a two-day meeting of information-system experts drawn from state system offices and the research community was held in November of 2003 to investigate alternative design approaches, and this meeting also yielded important policy insights. This report provides a synthesis of conclusions drawn from both sources.

**The Need for Information.** Those interviewed were unanimous in their view that the goal of obtaining better statistics about student retention and completion by tracking former students across state lines was insufficient, in itself, to generate much public or political interest. Periodically-updated national longitudinal databases such as *High School and Beyond* were seen by many as already providing a good basis for most of the research that is needed, and data supplied by the National Student Clearinghouse can already be used by institutions wishing to enhance their graduation-rate statistics by identifying the subsequent enrollment behaviors of students who fail to complete a degree. Although those interviewed were favorably impressed with the conduct and conclusions of *Following the Mobile Student*, moreover, they also felt that the report focused too narrowly on developing a tool to link databases and lacked compelling policy reasons for states to want to participate—especially in the current climate of scarce resources that typifies most states. As a result, they were uniform in their opinion that further efforts needed a clearer purpose, linked to important policy questions that states really want and need to answer.

Some of these questions are centered on economic and workforce development, and how effectively the “educational pipeline” in a given state contributes to “educational capital” in various fields. For example, state policymakers want to know whether they are net importers or exporters of graduates in key fields and whether students are leaving the state to pursue particular kinds of degrees (and returning to work in the state, if they do study elsewhere). Similarly, state policymakers are interested in equity issues associated with participation in the complete educational pipeline from K-12 through graduate education such as whether particular populations (regional, ethnic, or economic) are excluded or underserved. Addressing such questions requires expanding the scope of data coverage to include information drawn from statewide K-12 information systems and from employment records like the Unemployment Insurance (UI) wage records maintained by every state. One SHEEO interviewed went so far as to say that obtaining linked UI wage-record data for former students from surrounding states would be more valuable than determining whether or not they completed their programs elsewhere. All of those interviewed also stressed that such information would be useful only if it were program-specific, allowing policymakers to examine student flows in high-demand fields of study or in areas leading to employment in fields deemed important for statewide economic development.

Other questions of interest to state policy leaders center on return on investment, especially in states that operate substantial scholarship programs. For example, state policymakers want to know whether the students supported by such programs ultimately complete degrees, where they complete them and in what fields, whether they remain in the state after they have graduated, and the extent to which they obtain

employment commensurate with the state's investment. They also want to know whether there are substantial differences in outcomes between state-supported students who attend public institutions and those who use scholarship funds to attend independent institutions, as well as the extent of these differences where they exist. Studies focused on such questions can help inform decisions about whether to continue to invest in high-tuition/high-aid policies, the proper mix between need-based and merit-based aid, and continuing sponsorship and support for independent institutions as part of the state's total higher education "asset." Parallel questions are associated with the return on investment for student exchange programs like the ones operated by the Western Interstate Commission on Higher Education (WICHE) under which students pay in-state or lower rates to enroll in out-of-state institutions in designated fields. Currently the success of such students must be tracked by creating separate databases because records for the affected students are contained in multiple state SURs.

Finally, information to demonstrate institutional accountability continues to be important to state policymakers including data about student retention and completion. Virtually all states have established publicly-reported performance indicator systems for public institutions, and most of these prominently include graduation rates broken down by gender and race/ethnicity. Many of these indicators systems report both within-institution and within-state rates that take into account students transferring to other institutions. Increasingly, moreover, policymakers are interested in time-to-degree, with the expectation that completing programs within advertised times-to-complete is efficient. Similarly, state policymakers are concerned about escalation in the numbers of credits completed before earning a degree—especially as students attend multiple institutions. Information about credits completed by program can thus not only provide insights into the overall efficiency of the higher education system but can also open up issues associated with the effectiveness of credit transfer policies within and across states.

Taken together, these insights suggest that further data-exchange development efforts should include data beyond that needed simply to obtain more accurate statistics on retention and program completion within postsecondary education. Instead, they should ultimately enable state policymakers to address compelling questions about return on investment, workforce development, and patterns of flow through the complete K-20 educational pipeline. These insights also suggest that such efforts be "problem-centered" rather than "tool-centered." That is, the case to be made to policymakers should concentrate on the specific questions and problems that enhanced information could allow states to address and not lead with the enhanced technical capabilities that a multi-state information base might provide.

**Organizational Issues.** When an early draft of *Following the Mobile Student* was reviewed by stakeholders from Washington-based higher education organizations in late 2002, major concerns were expressed about the possibility that this might lead to the development of a "national student tracking system." Following this feedback, the language of the report was modified to emphasize the virtues of linking state unit record databases in a more decentralized fashion. Interviews with policy leaders reinforced and extended this basic conclusion. First, almost no support was accorded to a federal solution, implemented through the Department of Education or the National Center for Education Statistics (NCES). Virtually all those interviewed felt that such an approach would be over-bureaucratized and would take too much time to develop. But many did view the federal government as a likely source of ultimate financial support for such a venture. Others thought that NCES might play an important role in helping to establish uniform data standards. Indeed, few of those interviewed expressed much support for *any* kind of centralized approach regardless of who governed or controlled it. Consistent with a decentralized problem-centered agenda, a majority felt that the most effective approach would be to create multiple regional data-exchange arrangements among interested contiguous states. Such multi-lateral arrangements might be quite flexible, with each involving a somewhat different combination of data contents and organizational arrangements. Furthermore, those interviewed did not feel that universal coverage was mandatory. One respondent, for instance, saw no need to "fill all the cells" in order for such a capacity to be useful for policy. Another, located in Florida, noted that he "probably didn't need to establish an agreement with [the state of] Washington." But interviewees did feel that all such arrangements should all be governed by a common set of data standards and operating principles.

While the primary reasons advanced in support of a decentralized approach were practical—principally the fact that only a few states would be of interest as data-exchange partners for any given state and the desire



to keep the effort problem-driven instead of tool-focused—some centered on politics and perceptions. One was the desirability of keeping a low profile for such efforts in order to defuse inevitable concerns about privacy. The main danger to be avoided here, as one interviewee put it, was “sounding like George Orwell” with respect to government knowledge of citizens’ whereabouts. Others emphasized the political drawbacks of advocating multi-state tracking in the policy context of the Patriot Act and other post-9/11 initiatives. All these concerns reinforced the notion of approaching the task as an as-needed data exchange initiative limited to addressing particular policy questions rather than advertising it as “building a tracking system.”

Policy leaders also agreed that the organizational challenges associated with establishing and running any data-exchange approach far surpassed any technical issues. Here a number of different alternatives were suggested. The majority proposed establishing relatively straightforward multi-lateral agreements among state agencies that would delineate the procedural guidelines to be followed when exchanging data and would establish the requisite data protection procedures. Representatives of multi-state organizations like WICHE and ECS further suggested that such agreements might be structured on the model of an interstate compact—a mechanism that is both simple and familiar to state leaders. Others felt that it might be better to structure these arrangements in terms of a performance contracts with a third-party entity that would perform the requisite data linking and reporting. Contractual arrangements have the advantage of legal recourse if confidentiality is violated because any such violations would constitute non-performance of a binding contractual obligation. At the same time, state authorities are accustomed to dealing with contracts when they retain consultants and other third parties who need access to student data. Still others representing institutional constituencies (such as AACRAO) felt that institutions should be involved in governing such an entity as a “mutual-benefit corporation. Such an organization might operate much like the National Student Clearinghouse. As an AACRAO representative put it, “we ought to have the people who supply the data and who are legally responsible for the data govern any such organization.”

While these responses yielded no consensus of opinion on the form that governance arrangements for a data-exchange initiative should take, they do suggest that both the substance and appearance of embarking upon a large, uniform, “national” effort should be avoided. They also suggest the need to proceed incrementally through limited, practical efforts with relatively low visibility and high payoff.

**Uneven State Capacity.** *Following the Mobile Student* established the fact that the data contents and architectures of extant SUR data systems were compatible, but because NCHEMS did not actually visit state SHEEO or system offices to directly examine these data systems, qualitative issues like system maintenance and staff capacity were not addressed. Interviews with policy leaders and those responsible for managing and operating SUR databases, however, reveal a significant pattern of uneven development with respect to such systems. First, many are getting quite old and are based on an IT architecture that reflects second-generation, batch-processing approaches typical of the “legacy” data systems developed in the mid-1980s. These systems are still quite capable, but changing them is difficult without a great deal of special programming. At the same time, largely because of substantial state budget cuts, many state agencies are laying off the programming staffs needed to update and maintain these systems. More recently-established SUR systems, in contrast, employ third and fourth-generation programming languages in a relational database environment and are often Web-enabled. Differences across states are equally great with respect to their abilities to use available data to generate meaningful policy information. Some state agencies have substantial experience in conducting complex, multi-faceted policy studies using longitudinal student enrollment and employment records. Others have almost no such experience or capacity, and employ only a few analytical staff whose responsibilities are centered on compliance reporting. These disparate capabilities also cut across one another. States like Missouri and Oklahoma, for instance, have a relatively sophisticated analytical capacity but employ hard-to-maintain legacy data systems, while states like Ohio have made substantial investments in state-of-the-art Web-enabled relational databases and are only now developing the analytical capacity needed to harness the contents of these databases as management information.

In the short run, this growing unevenness in state capacity suggests the need to develop data protocols that are extremely simple and robust—capable of extracting data under a wide range of technical circumstances and demanding relatively little of already stretched (or diminished) state IT staffs. One approach here

might be to employ more technically-advanced data systems in designated states as “lead” data warehouses, contracted to house and match data drawn from states with less technical and analytic capacity. In the longer term, this condition suggests that targeted investments in SUR capacity-building—including both a standardized relational database architecture that could easily be customized to the needs of individual states and appropriate staff development efforts—might be extremely effective in developing informed state policies about student progression and program completion. Part of this longer-term strategy might involve the development of “best practice” reporting and analytical tools that would both improve state capacities to use longitudinal information and that would avoid the need for states to continually “re-invent the wheel” when generating the information needed to support sound policy decisions.

**Institutional Issues.** In preparing *Following the Mobile Student*, NCHEMS essentially assumed that once institutions reported data to an SUR database, the further use of these data for the purposes of tracking students across state lines—if appropriate technical solutions and privacy safeguards could be found—would automatically follow. While this assumption remains true for the substantial majority of public institutions, it is far less true of the growing number of independent institutions that participate in such systems. Interviews with policymakers in the states that include independent institutions in their SURs reveal that the cooperation of these institutions is usually premised on a delicate set of mutual understandings about how supplied data will and will not be used. And even in the case of public institutions, Coordinating Boards often are quite sensitive to institutional complaints about reporting burdens and the potential misuse of data. As result, many interviewees noted that it would be wise for states to consult specifically with their institutional constituencies about cross-state use of SUR data because such uses of institutional data were not contemplated or addressed in the agreements that originally established such systems. Similarly, public and private institutions alike often see state-maintained SUR databases primarily in terms of reporting burden instead of an information resource. This is because so little attention is often paid to providing institutions with information products that they can use in their own planning and marketing efforts. The complaint from institutions under current circumstances is frequently expressed as follows: “data go up [to the state] but they never come back down [to the institution].” This suggests that institutional cooperation might be greatly improved by devoting more attention to providing institutions with SUR-generated data products that they can use. Prominent examples mentioned here included reporting student performance by field of study to feeder high schools and exchanging data on transfer student performance among institutions with well-established transfer paths across state lines.

Those interviewed were also asked specifically about what might be done to increase the participation of independent institutions in state-maintained SURs. Most acknowledged that such institutions (and especially their lobbyists in Washington) would be initially skittish about this matter. But the majority—including both state-level informants and representatives of the private college sector—believed that independent colleges could benefit greatly from the kind of information that a multi-state data capacity could generate. The most important of these benefits were seen for individual institutions in the form of enhanced information for planning and accountability reporting. Reported retention and completion rates for independent institutions, for example, can only increase if the universe of potential transfer-out destinations is expanded. With the inclusion of outcomes data, these information resources were seen by many of those interviewed to be even more powerful because they could be used in marketing. But policy leaders also saw potential *collective* benefits to private college participation in such an effort. For example, one product might be information to demonstrate the contribution of independent colleges to state policy objectives including increased numbers of graduates in key fields retained in the state and more efficient degree production. More narrowly, such information could be used to directly demonstrate return on investment for the often-substantial amounts of state scholarship funds that are spent at independent institutions. Finally, private colleges should want access to better data on student graduation and persistence because they will likely do better than public institutions on such measures. Although most policy leaders felt that independent colleges do not yet realize the potential value of these benefits, most felt that they could (and will) eventually be convinced.

Taken together, these findings make imperative the development of a succinct “case statement” to use with both state officials and institutional leaders that a) explains the rationale for multi-state data-exchange efforts and the payoffs that can be expected from engaging in them and b) clearly establishes the ground-

rules for participation and describes the safeguards in place to ensure confidentiality and guard against misuse.

**Alternatives to an SUR-based Approach.** Although the topic was not explicitly included in the interview protocol, the National Student Clearinghouse frequently came up in these conversations. These discussions revealed a distinct ambivalence about the role of the Clearinghouse as a source of information about student success. On the one hand, most of those interviewed acknowledged the importance of the Clearinghouse because of its impressive data coverage and its FERPA-compliant operating environment. In fact, some of those interviewed—including representatives from both state and national organizations—had recently explored working with the Clearinghouse to negotiate high-volume student-matching arrangements on behalf of their members. On the other hand, most of those interviewed were not satisfied with the Clearinghouse as an alternative to an SUR-based approach, though they believed that it could be useful in supplementing such an approach.

Reasons for this conclusion were many. First, state agency clients were discouraged by the cost of Clearinghouse services and reported being unable to obtain volume discounts for large numbers of transactions. Second, some reported that because the Clearinghouse was largely set up to serve institutional clients, it is neither configured nor inclined to respond appropriately to the data needs associated with state-level policy research. Third, most agreed that the limited data-element coverage of the records held by the Clearinghouse limited its value in enhancing graduation-rate statistics. Key data elements missing in the Clearinghouse deemed particularly valuable for state-level policy research included program of study, credit-completion or course-progression information, and some measure of enrollment intensity (e.g. full-time/part-time status). Finally, some (but not a majority) of those interviewed were uneasy about the Clearinghouse's relationship with the student loan industry. For example, AACRAO representatives worried about the Clearinghouse because it operates under FERPA as "agents" of the institutions, who legally retain ownership of the information; but the fact that the Clearinghouse is operated by lending institutions who may want to use the data for other purposes creates a clear conflict of interest. But the Clearinghouse is simultaneously seen as an example of a real "value-added" product that yields something that institutions want routinely in the form of transactional verification-of-credential services, not just "tracking" information which is only useful for research and accountability purposes.

Based on these responses, the National Student Clearinghouse should undoubtedly be a part of any effort to increase the coverage and comprehensiveness of national longitudinal data resources to investigate student progression. In the short term, for example, it represents the only way to obtain information about enrollments at the majority of independent institutions. But lack of data detail and concerns about ownership and cost make most policy leaders willing to support an alternative approach.

**Confidentiality and Privacy.** A major potential difficulty facing multi-state data exchanges raised by *Following the Mobile Student* centered on privacy and, in particular, on the requirements of the Family Educational Rights and Privacy Act (FERPA). Results of the interview process, however, suggest that the heart of the challenge is not FERPA and related regulations *per se*, but rather what might best be called "the perception of FERPA." The central difficulty is that state and institutional officials are simply not clear about what FERPA does and does not allow. This creates a situation under which both institutional attorneys and state Attorneys General are inclined to say "no" to almost data sharing arrangement simply to play it safe. States like Florida, Oklahoma, and Texas, in contrast, have developed excellent resource materials that allow them to do everything envisioned by *Following the Mobile Student* in a manner that complies fully with FERPA and other federal/state privacy policy guidelines. It is also important to recognize that there are many regulations governing the confidentiality of student records, and that some of them are contradictory. For example, FERPA, administered through the Department of Education, places clear limits on the kinds of research that can be performed and how individual student record data must be handled, while Carl Perkins and other occupational education programs administered through the Department of Labor not only allow, but *mandate*, the use of educational and wage-record data in required accountability reporting. Finally, several of those interviewed mentioned the growing interest of university Institutional Review Boards (IRBs) in getting involved in decisions about the use of student records for research purposes, though none of the cases mentioned ultimately prevented such use.

While these matters are serious, they chiefly arise because of ambiguity, not because of the actual content of legal restrictions. To address this situation, states need clear guidance on how to proceed in conducting SUR data exchanges based on the proven successful track record of states like Florida, Texas, and Oklahoma. This guidance should include specific examples of the kinds of protocols and agreements that need to be in place together with any requisite legal language. One state policy representative also suggested that it would be useful to determine if there is any body of actual case law regarding FERPA and related privacy regulations that might help states determine how to proceed. Several others noted the desirability of obtaining a clear ruling from the Department of Education on allowable uses of SUR data under FERPA, because the current situation of ambiguity allows each campus or state's legal counsel to make such determinations independently—frequently with different results. Progress in pursuing multi-state data exchanges under these conditions thus requires making an important strategic choice about how to proceed. On the one hand, it might be desirable to be proactive about seeking a clear legal opinion from the states involved in prototype data exchanges. On the other hand, doing so risks the possibility of raising issues where no objections have up to now been present. State policymakers for the most part advocated a middle ground here, which would involve proceeding incrementally and experimentally to prove the concept and addressing privacy issues as they arise, while at the same time being as thorough as possible in providing detailed guidance based on known state track records and legal opinions.

**Conclusions.** In sum, consultation with state and national policy leaders clarified a number of issues associated with pursuing the agenda suggested by *Following the Mobile Student*. Chief among these are:

- The need to provide states with a clear policy rationale for pursuing multi-lateral SUR data exchanges that centers on economic development, return on investment, and educational pipeline issues.
- The desirability of exploring multiple decentralized multi-lateral arrangements among states founded initially on specific regions and problems, rather than proceeding immediately with a more centralized and comprehensive approach to coordination.
- The eventual need to address considerable differences in state capacity to conduct longitudinal studies of student progress.
- The desirability of pursuing multiple sources of data about students beyond SURs, including the National Student Clearinghouse.
- The need to provide states with detailed guidance about how to address privacy concerns associated with exchanging student records based on the success of “best practice” states.

These points are already proving valuable as NCHEMS begins to pursue some actual demonstrations of data exchange with support from the Lumina Foundation. As these demonstration efforts unfold, NCHEMS will continue its consultations with policy leaders to obtain additional insights based on actual experience.

## **Appendix: List of Interviewees and Data-Collection Protocol**

### List of Interviewees (Policy Group):

Jerry Sullivan	American Association of Collegiate Registrars and Admissions Officers (AACRAO)
Barmak Nassirian	American Association of Collegiate Registrars and Admissions Officers (AACRAO)
Paul Lingenfelter	State Higher Education Executive Officers (SHEEO)
Hans L'Orange	State Higher Education Executive Officers (SHEEO)
Ted Sanders	Education Commission of the States (ECS)
Richard Eckman	Council of Independent Colleges (CIC)
Bill Proctor	Florida Council for Education Policy Research and Improvement
Pat Dallet	Florida Council for Education Policy Research and Improvement
Rod Chu	Ohio Board of Regents
David Longanecker	Western Interstate Commission on Higher Education (WICHE)
Brenda Albright	National Postsecondary Education Cooperative (NPEC)
David Swedlow	National Collegiate Access Network (NCAN)
Kent Weldon	Indiana Commission for Higher Education
Jay Pfeiffer	Florida Office of K-20 Education Information and Accountability

### List of SUR Technical Meeting Participants:

Emerson Elliott	National Council for Accreditation of Teacher Education (NCATE) [Former NCES Commissioner]
Charles Lenth	State Higher Education Executive Officers (SHEEO)
Joseph Marks	Southern Regional Education Board (SREB)
Tod Massa	State Council of Higher Education for Virginia
John Milam	HigherEd.org
Rhea Santos	Illinois Board of Higher Education
Debra Stuart	Oklahoma State Regents for Higher Education

John Wittstruck                      Missouri Coordinating Board for Higher Education

David Wright                         State Higher Education Executive Officers (SHEO)

Interview Protocol:

Begin with brief description of the NCHEMS Lumina Project

[All participants will have been sent a copy of *Following the Mobile Student*]

- Is the need for such a capacity really there, or are existing ways of estimating graduation rates based on national samples like *High School and Beyond* and limited datasets like the National Clearinghouse sufficient to address most state/national policy and accountability questions?
- How do you see this initiative in relation to the emerging Reauthorization debate? Are there particular issues here that we need to pay attention to or watch out for?
- Who would you expect to be the principal “customers” of such a capacity? Where is the demand coming from?
- What might be some strategies to induce greater numbers of independent institutions to participate in state-level unit record databases? What kinds of safeguards would they need and how extensive do you think participation might eventually be?
- With respect to the issues above, what about proprietary institutions?
- What particular issues do you see arising with respect to confidentiality and privacy, and how might such an initiative address these issues? And, more generally, how do you expect the matter of privacy and confidentiality to unfold politically in the next four or five years?
- How might such an initiative be structured from an organizational standpoint? Alternatives might include a) a series of bi-lateral or multi-lateral arrangements among individual states (or independent college associations) using a common data-exchange protocol, b) state-federal partnerships (such as seed funding from the federal government provided through a 1202 Commission-like arrangement or the use of existing channels like the NCES/SHEEO Network, and c) an independent non-governmental organization like the National Clearinghouse.

## Lumina Project Policy Agenda

Meeting with Jerry Sullivan and Barmak Nassirian (AACRAO)—10/24/03

Verifies that there is a very high level of interest in Congress right now for longitudinal data for accountability because of Reauthorization, but not clear what the proposal will be and the wider HE community does not want to advance a solution at the moment—the position is go slow and “if we have to do it, be pragmatic, but don’t build something prematurely.” Not clear that AACRAO entirely agrees with this position.

The question is what you really want to know from such a system...if it is only post-leaver outcomes for research or retention reporting purposes, linking databases in the way we propose may not add much value for money...but different stakeholder in the HE community will have very different points of view on this. The CCs need longitudinal tracking for Perkins and other programs no matter what else happens, and their enrollments are so volatile (and their data systems so underdeveloped) that they would clearly like anything that would enable them to better document successes. AASCU feels that state reporting is natural: they have to deal with public accountability anyway so they must supply outcomes information and have good tracking data, and are well served by SUR databases in this respect. Independent colleges split on this matter...NAICU remains skeptical of developing any kind of “national” outcomes statistics as an official line but AACRAO agrees that among certain independent college constituencies there may be support because of perceived *quid pro quos* on state student aid or the usefulness of such information to advocate for their contribution to state needs.

Systems like this will produce useful information, but it isn’t necessary to include the entire universe of institutions or students to get decent data for research or policy...can deal with the inevitable information gaps if policy information is the ultimate goal. They also agree with an approach that wouldn’t necessarily have to fill “all the cells” in a 50 by 50 state matrix—that is, one that would concentrate on areas of high interstate mobility where there is heavy enrollment volume already. And they gave no sense that existing research databases like HSB would fulfill the ultimate demand for accountability or research information that they saw coming.

AACRAO thinks the fifty-state solution will be much better than a single federal system...“50 variations on a theme” is what built a strong American system of higher education in Jerry’s view...but they may be a minority at One Dupont on this.

The best situation would be if a data-system approach could provide information that is genuinely useful for decision-making on a local or regional basis and build from there. And the best information will not just be useful for “research” purposes but will actually facilitate transactions and help institutions do their jobs better. Also, it would be important to start out with a posture of local utility and voluntary participation. [Barmak comment: “anybody can envision (and perhaps support) a voluntary system...the problem is when it becomes mandatory.”]

[Barmak] HE data systems are “stuck in the 70’s” with respect to their organization and conceptual structure. Current institutional and state system reliance is on large legacy systems with ASCII-format extract [typical] and data warehouse capability [at best] bolted on top of them for SUR reporting and tracking purposes. What we should be doing instead is “re-inventing” the whole approach to data systems based on experience in other industries like banking [an “XML approach”] based not on census- positioned extracts but on real “interoperability” capabilities among existing transactional systems. [The analogy used here was the migration from ASCII to Word-interoperability standards for word-processing systems.] The ultimate vision in this scheme would be electronic transcripts and portfolios anchored on competencies and certified by a third-party authority...Barmak admits this will take a while to develop, but it would be the

right way to go. But the main point [Jerry] is “we need to take control of these issues more generally with respect to data systems as a higher education community [successful examples cited are ELM and “CommonLine”?). Such a system [Barmak] ought to be based on individual transaction management, but could generate management information at the same time...like stock trading or banking now does [“you want the most up-to-date data possible and extracts won’t do this”]. This also implies agreement that the proper unit of analysis may not be “the higher education universe” as we currently conceive of it—we may need a more comprehensive “human capital database” instead that includes high school experiences and employment outcomes...[Barmak notes here that the largest National Clearinghouse customer at the moment is Walmart that wants to match to high school records for verification of attendance].

FERPA issues are real and growing. The National Clearinghouse is in AACRAO’s view in technical violation of FERPA because of the way they do record matching, but they have been granted an explicit exception. Jerry is very worried about the Clearinghouse because they operate under FERPA as “agents” of institutions, but they are essentially owned by the loan industry, creating a clear conflict of interest. On the other hand Barmak—though he agrees with these doubts—also cites the Clearinghouse as an example of a real “value-added” product that yields something that institutions want routinely (in the form of verification-of-credential services) not just tracking information, which is only useful for research or accountability purposes.

Governance issues for any data-exchange entity will be far more challenging to develop than any technical issues. The USDOE could own it, but “it wouldn’t work and it would cost too much.” One alternative is a “mutual benefit corporation” model that might be based on an NGO, but not one owned by the loan industry. Governance is the primary question for the HE community in the current Reauthorization environment...[Jerry] “we ought to have the people who supply the data and who are legally responsible for the data govern any such organization...” [and, by implication, Lumina may have some baggage in this respect too.]

Therefore, what could an NCHEMS effort do? Two answers: 1) demonstrate the feasibility of a decentralized solution based on “middleware” that might demonstrate the viability of an alternative to a “big federal solution” and, 2) check out why the costs of proposed alternatives are so different...e.g. why does it take billions of dollars to run E-ARMY-U and NDSL, while the Clearinghouse could achieve start up with equal [huge] numbers of records on only a three million dollar base?

**Bottom line:** Working at a state level to respond to clear policy and institutional needs on a voluntary basis makes sense to AACRAO, and any effort we make in this direction will be consistent with what they [appear to] want. This would be a state-level response based largely on the limited information demands of a particular region, and it might be better if several *different* solutions were advanced rather than a single “data exchange protocol” as we originally envisioned. We also need to keep our eye on Barmak’s “big solution,” which is very conceptually sound, but in my view far outside the vision of the community (either political or IR/research) at the moment.



## **Lumina Project Policy Agenda**

Meeting with Paul Lingenfelter and Hans L'Orange (SHEEO) – October 28, 2003

Believe that current data sources (e.g. HSS, Clearinghouse) may be capable of answering the degree attainment rate question as well as an SUR database approach. But these sources do not address the more complex questions of program enrollment flows or migration. “Our thinking is too much constrained by the traditional questions [like graduation rate].” Primary benefit of this project may be to move the focus of state questions if we frame them properly. They do see more rhetoric about the need for this kind of information at the federal level (but cautions about sources here—e.g. Jane Oates is high on NCES...).

Paul believes that the primary benefits for such a capacity will be to institutions in tracking their own graduates, for whatever reason. So policymakers may be less interested in this capacity than institutions will, because the enrollment flow question is not a pressing one for policymakers right now. [Note that SHEEO is thinking about pursuing a service for states to work with the Clearinghouse, and Paul sees this largely about enhancing state capacity to serve institutions better.]

Have not heard as much about graduation rates in the recent conversations about Reauthorization as when this first came up a few months back. The focus seems to have shifted fairly decisively to costs. They see this as a “good distraction” because there will be less pressure for a “federal solution” to looking at graduation rates.

Hans was a bit surprised by our findings about the numbers of independent institutions in SUR databases, but acknowledges that everybody thinks it has been growing. Paul pointed out that independents (“except the top fifty”) have every reason ultimately to want to play [“if you are marketing, information is gold”]. And many of them are increasingly not in competition with one another but with publics.

Emphasize that the project is not about a “national” solution, but rather about regional or multi-lateral agendas that benefit real constituencies with real (and different) policy problems. This should not be about “getting a number” that is missing from the national policy dialogue. Graduation might not be the most compelling question in any case...rather whether graduates are retained in state or whether state scholarship funds are well invested in people who ultimately will work in the state. There has to be a simple, well-understood state reason to engage in this kind of activity.

Paul not “allergic” to the Clearinghouse as a source of information and would like to actively pursue a relationship. Does not understand why the Clearinghouse is seen as having suspect motives—part of the issue may be pricing, or simply the fact that they operate as a fee-for-service organization. “The higher education community should have the clout to extract from them what is needed.”

Above all, any approach needs to be simple...simplicity in concept and delivery is what makes the Clearinghouse appealing and cheap. States will vary enormously in their technical capacity to make this work and they are already badly stretched—technical folks are often the first to go during budget cuts. So need to assume variability across states and emphasize straightforward approaches.

Although Paul's first choice with respect to an organizational option would be to go with something like the Clearinghouse ("you get the most for the least"), he [and particularly Hans] are also attracted to the notion of a set of bilateral and multilateral arrangements among willing states ("letting them do their own thing"), but guided by a single data-definition and data-exchange protocol. In their view, the most effective contribution the Lumina project can make would be to build such a protocol.

Seeing a little bit of softening on the FERPA issue, but privacy still a major concern. Clearly the states that really want to do something are not seeing FERPA as standing in their way. The basic problem is lack of meaningful federal direction, which would ensure a bit more uniformity in the way state AGs are interpreting FERPA. Question raised: does FERPA apply to records generated (or disclosed) after a student has graduated? [We think it does.]

Watch out for one unintended consequence: yet another version of a "graduation rate" statistic that would "muddy the waters." We don't want "an NCES number, a state number, an NCAA number, and an NCHEMS number..."

**Bottom Line:** Approach this as a mutual-benefit set of tools to help solve regional policy issues—especially migration and human capital issues—instead of a "national" search for a better graduation rate number. Try to get several groups of states working on this at once focused on local issues of concern, but using (or building) a single workable model for data exchange that others can adopt

## Lumina Project Policy Agenda

Conversation with Ted Sanders (ECS) – October 28, 2003

Feels that the initiative as stated lacks a clear policy rationale—what kinds of questions would such a capacity be designed to answer? Agrees that the best answer to this question centers on “human capital” development in particular geographic regions. [Notes “regions” in two senses—multi-state aggregations (like WICHE or SREB) and conurbations involving parts of multiple states.] But given this, would like to see the focus expanded to involve the whole pipeline—from K-12 through employment. Focusing just on higher education is too narrow for the kinds of broad-based longitudinal analyses that are needed. And states are already familiar with trying to use longitudinal data through the stimulus provided by No Child Left Behind.

The “interstate compact” notion on which ECS is founded might prove an excellent model for creating data-exchange and access agreements in connection with this initiative. Ted would be interested in exploring this in partnership with NCHEMS later on in the project.

FERPA concerns certainly can be overcome. Cites example of the new National Center on Educational Accountability at UT-Austin, which has assembled data from multiple states using uniquely-identified K-12 student records that are linked through a locked key file. This might be useful to us as a model of how to create a third-party dataset. [Ted said he could put us in touch with the right people here—names mentioned were Tom Loos(?), Brad Duggan.] On FERPA and privacy, “the whole matter is one of perception”...the legal requirements are not that bad.

Sees questions about graduation rates being raised largely for political reasons as the presidential election season heats up...people will want to embarrass the administration by highlighting falsification of graduation rate statistics in Houston school district when Rod Paige was in charge...question will be “how do we get honest reporting,” so a federal solution may seem attractive to Congress.

**Bottom Line:** Need a clear and compelling rationale for states to want to do this in an environment that is otherwise very distracting for policymakers [“states are wound up now about as tight as you can wind them”]. Human capital approach may be the best way to sell this, and it will be important to demonstrate how data exchanges can actually be accomplished and be of benefit.

## **Lumina Project Policy Agenda**

Conversation with Richard Eckman (CIC) – November 10, 2004

There is only a short list of topics for accountability measures being considered as part of Re-Authorization, but the list is quite volatile (“the conversation always shifts ninety degrees”). Not hearing as much about graduation rates as last year right now...the big issue is cost (“not price”) and “accountability.” “Accountability” means “institutions taking the responsibility to report commensurable information about various aspects of performance,” but ACE thinks it could mean institutions doing the reporting on their own with their own definitions through a website...Rich does not think this will fly.

Accreditation not very united as a voice in Reauthorization...institutions are often unwilling to speak up in defense of voluntary accreditation because of their own bad experiences with accreditors (mostly SACS here). Accreditors are also not of one voice themselves and the feds are not treating them well in the recognition process even if they are seen as being on the “right side” (e.g. AALE). Don’t hear much these days about de-coupling accreditation from Title IV the way you did last year either.

Transfer of credit is a major issue for Congress right now...ACE is worried about this because it leads right back to the retention/completion issue, which they are happy to keep at a distance. [States are also interested in credit transferability] The worry is that the feds may mandate transfer of credit from any accredited institution to another (which would include the proprietaries and really heat things up). So One Dupont remains very ambivalent about this issue.

Independent colleges should want to get better graduation/persistence data and advertise it publicly because they will likely do better than public institutions on these measures...but they don’t yet realize that public reporting will be in their own best interest. The NAICU line is still to resist such reporting on the grounds of federal intrusiveness and unwanted burden. Independents should also be able to use such data to argue for a greater state role and support—the argument being the benefits to a state of supporting institutions already in place when there is lack of public capacity. Such data may demonstrate that independent colleges deliver more degrees for fewer dollars.

CIC would like to be in a position to advance a state policy agenda with respect to independent institutions. One issue is to help bear the burden of accommodating students displaced from closed programs at public institutions (Rich tells University of Nebraska/Nebraska Wesleyan story about this). Another fruitful area would be merit/need-based financial aid programs and what role independent colleges play in providing access and successful outcomes for such programs. Both of these agendas would require good statewide data.

Supports the notion of a decentralized, region-based approach that could be loosely organized in terms of data sharing instead of a “national” or “warehouse” solution. This would be generally seen as less threatening. Rich believes that NCHEMS retains a lot of trust and capability in the eyes of the community

and would likely be able to negotiate for something like this in preference to either a federal or a strictly state solution. “The worst case is that [some government agency] will step in and try to do something like this...everybody has a horror story.”

Good independent college associations to potentially work with include Kentucky (Gary Cox), Minnesota (David Laird), Pennsylvania (Don Ferris), Wisconsin (Rolf Wegenke ?), Ohio, and New York. Rich thinks it would be good if we included one of these explicitly as part of the plan.

Bottom Line. Verification of decentralized approach and need to have an emphasis on adding value to particular institutional sectors (especially demonstrating to a state the value of independent colleges). If we can do this visibly for CIC-type institutions, they may well want to play. Also, NCHEMS has a good deal of credibility in this arena and we should exploit this explicitly.

### **Lumina Project Policy Agenda**

Conversation with Bill Proctor and Pat Dallet – Florida (11/17/2003)

Believe that a decentralized, problem-centered approach is the right one in general. Should begin looking at places where there is likely to be high transfer volume (“we wouldn’t need much of a relationship with Oregon”). They pretty much have all the data they need in Florida from Jay Pfeiffer’s office and Jay is already working with two other states to get data on migrating students. But from the perspective of a policy shop, getting data from Jay can be a bit of a challenge because there is a lot of control exercised. Jay feels this is necessary to protect privacy and Bill and Pat admit that he probably would not gotten as far as he has if he hadn’t operated in this way. A warehouse is planned for public access to much of this data, but will report aggregate data only—outside users will not be able to link records directly.

Major examples of data use cited were about financial aid—the two big issues in Florida right now are pre-paid tuition and Bright Futures and they need data to show what happens to these students, many of whom may go out of state. Because students do not fill out a common state or federal form for these, they have no way to find them. Bill believes that having good data on family income would reveal that they are not serving those with incomes of \$40-60K effectively, but there is no way to investigate this right now.

SHEEO is trying to work a common discount deal with the National Clearinghouse, and that probably would add value. But Bill is not high on the Clearinghouse because of the money-making aspect of it...“you send them your data, then you have to buy it back.”

There is some nervousness about centralized records even in Florida. For example the Community Colleges have always been mistrustful, but this issue is largely about human resource systems not student data. Bill sees no reason why independent colleges should not want to be part of such a system. Many of them are in Florida through the leverage provided by state-based aid programs and independents have already used it to their advantage. The independents serve more minority students than public institutions in Florida, for example, and can document this. The independent colleges have student demographics in the system, but not course-level detail. The proprietaries are not players in this because there is no handle on them from the state like financial aid, although the large accredited for-profits might be a likely future target.

Florida is not “grandfathered out” of FERPA as rumors have it, but instead was able to develop an effective system through careful planning and knowing as much about the rules as possible (the source of this rumor is Dennis Carroll at NCES!). “Any state could do what we have done.” You have to have a sound

organizational structure and a capable person in charge who will do what needs to be done to protect the data and follow the right rules. But they admit that even Florida might have trouble setting up such a system today with “all the paranoia out there” about privacy. In fact, getting any kind of state initiative going now is difficult—“we couldn’t have set up common course numbering now either...”

Making progress will depend on picking the right problems that require this kind of data, and for Florida these problems center on the effectiveness of state-based financial aid programs. Generating indicators that speak to creating greater efficiencies in the educational pipeline would be a high-demand area as well.

Bottom Line. Stick to a decentralized approach and learn all you can from Jay Pfeiffer about how FERPA and privacy need to be handled.

### **Lumina Project Policy Agenda**

Conversation with Rod Chu, Ohio Board of Regents (11/20/03)

Stimulus for new Ohio information system “was me...” When Rod interviewed he was struck from his business background by the lack of information-driven decision-making and wanted to start changing this because “all we [SHEEOs] have is the bully pulpit, and using information is the way to leverage this position.” They had a lot of data at the BOR but were not using it to inform policy or report to the public. The BOR was putting in a new information system anyway, so make it relational and quick-turn-around report oriented. Now “data are us...”

Particularly interested in linking to economic data in all of this—UI wage record and return on investment for state scholarship issues. Sees the greatest policy benefit for states beyond just things like graduation rates to broader social issues of mobility, individual earnings benefits, and economic return. If forced to make a choice, Rod actually would rather have a national linkage to UI wage records to track Ohio students economically than one that would indicate transfer (though certainly both are desirable).

To date, BOR has not tracked student enrollments into other states, but would be interested in doing so. They are exploring a relationship with the National Clearinghouse to follow up on state scholarship recipients particularly, but are concerned about the coverage of the Clearinghouse and the limited amount of data that can be generated. No comment at this point about the Clearinghouse as a partner [did not sense any negative affect toward the Clearinghouse because of their connection to the loan industry nor a concern about price].

Independent colleges in Ohio are included in the SUR largely as a quid pro quo for the receipt of state scholarship money—“you took the money, now you have to play.” Rod believes that cooperation would not have been possible without this and can’t see how any state without such an arrangement would induce independent college participation. Getting cooperation has been difficult, however, as the independent colleges have always had their own form of performance reporting where they could pick and choose what to report. Now the BOR puts out reports on their comparative performance which can contradict some of what they were saying, and this caused some friction. So far, there has not been much awareness of what the data can do for independents to make their case—for example that they might serve more diverse students or that they might be able to show higher performance and therefore justify higher prices. The independents also might benefit from the ability to do federal reporting more efficiently, though they have not recognized this yet...this was seen as a big benefit by the public institutions when better follow-up capacities were developed. Proprietaries are not an issue in Ohio because there are so few of them.

Confidentiality and FERPA issues have been treated very strictly by BOR at the individual student level. Rod made a distinction, though, between “individual privacy” and “institutional confidentiality” (raised by the independent colleges), which does not deserve protection...believes that “all this should be out in the open and let the data fall where they may.” But an exception should be made for performance reporting while the measures are being developed and are still experimental. Has not heard any groundswell of political or public opinion one way or another on the FERPA/Confidentiality issue, but expects that if there is a conflict, open access to information because of accountability demands will ultimately trump privacy.

Believes there is a major opportunity to drive Congress to think more comprehensively about higher education performance through the Re-authorization process. Federal approaches to performance need to get beyond institutional graduation rates (“which are insane”) to look at systemic performance and economic issues related to rate-of-return and human capital development. This would also be a major way to address the disconnect in federal policy between DOE and Labor Department in looking at performance.

Not in favor of NCES having anything to do with developing a national capacity to track students unless there are major changes in the way they operate. Federal data is “too slow in getting out” and too tightly restricted with respect to statistical standards and access to detail so the result is “we make policy with no data at all because NCES says the data aren’t perfect.” Also, if you go to the feds, you are locked into a bureaucratic approach that is basically about turf protection. [Rod had an encounter with Dennis Carroll through NPEC and was investigated for potential FERPA violations in the BOR’s database development work!]

Another potential backlash is in the K-12 accountability movement. K-12 information is comparatively underdeveloped in Ohio and the NCLB federal position is far too narrowly conceived on institutional performance. We need to make sure that we keep focused on the right state policy issues, which go way beyond individual institutional performance and accountability.

Bottom Line. Sees value in multi-state data and doesn’t see why it needs to be in the form of just a series of individual state partnerships for data-sharing. But the focus needs to be beyond just getting better graduation rates for individual institutions. And as much emphasis needs to be placed on the kinds of public reporting and policy uses of this information as simply developing a capacity to obtain and exchange multi-state data.

## **Lumina Policy Agenda**

Conversation with Brenda Albright, National Postsecondary Education Cooperative (NPEC) – 11/26/03

Purpose of this effort needs to be made clear for states to want to participate. Simply getting better data about graduation rates is not a good enough reason for a state to do this, especially in tight times when they may be cutting back on staff and budgetary discussions are overwhelming everything else. When she was in TN, moreover, Brenda found that reporting graduation rates in and of themselves rarely yielded much policy discussion because they weren't able to be linked to questions of resource allocation. She agreed that questions about time to degree or excess credit totals on graduation would be much more compelling. Having the capacity to link to workforce training issues and migration by program, or whether or not the state is getting return for its scholarship money would be much more compelling policy purposes. Also liked the idea of regional linkages around state borders. But simply research or accountability purposes will not be reason enough for states to want to invest in something like this.

Many of these systems are getting very old and are difficult to document and maintain. There might be major benefit to linking this project to an effort to upgrade and re-think these capabilities conceptually. Liked the idea of trying to interest Lumina in the notion of "capacity building" for state information systems centered on longitudinal data.

Reauthorization just as unclear now as it was six months ago. But Brenda has the impression that the early push for graduation rates as the central higher education accountability measure has dissipated somewhat and is being replaced by tuition cost as the central concern. But agrees that if the administration wants to push something it will likely be along the lines of NCLB and this may well fuel a demand more accountability-centered information. Also if this occurs, there will be have to be a concomitant effort to try to modify FERPA and other restrictions on information to allow greater access to data for generating the needed statistics. But the privacy issue is an old one and there are no new wrinkles in it. Actors like NAICU will always use FERPA as a way to block access to information they don't want released, and there



are plenty of established ways to keep records confidential without compromising their utility as the basis of policy information.

The rationale for independent college participation remains primarily through receipt of state scholarship funds. Brenda built the independents into the SUR in TN from the outset because of this. A lot depends on state culture and the historical relationship among the sectors within a state...at first, there will be mistrust because of what a state potentially could do with data like this, even if it doesn't intend to. Over time, if there is no misuse, trust grows and the state can do more things. But it is important to take the potential of misuse very seriously and talk about it visibly from the beginning to signal that you recognize it as a problem. There should be clear ground rules about what such data can and cannot be used for that all parties agree to from the outset.

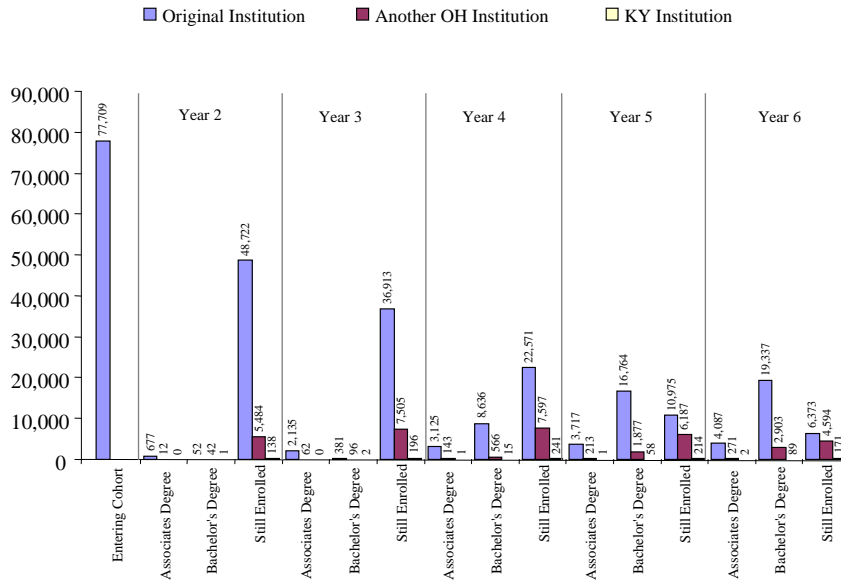
That said, Brenda favors an open policy of data use, once the proper safeguards on privacy are established and clear ground rules regarding data use are agreed to. She likes the idea of web access as FL is doing with their data warehouse and may even want to open some of this to the public to a limited extent. Certainly favors a situation in which the data can be accessed by researchers as well as other states. Not afraid of misuse under these conditions and feels that the resulting dialogues will ultimately be positive even though some constituencies might want to protect data.

Bottom Line. Be very clear about policy purposes for an effort like this, both in general and with respect to state participation. Simply getting better data about graduation and completion is not enough of a reason for states to want to participate in this, and purposes need to go beyond accountability.

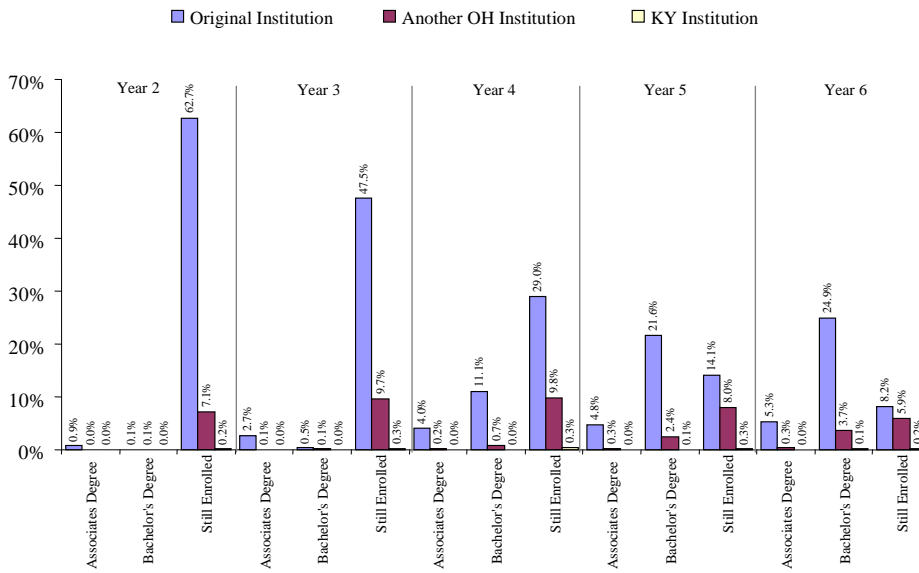
## **Appendix A**

### **Selected Results of Ohio/Kentucky Data Match**

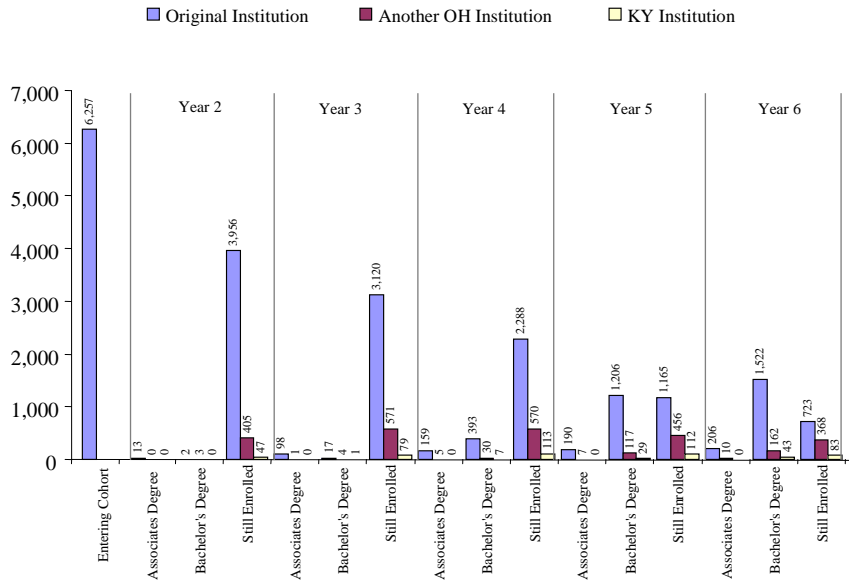
## Ohio Students Transferring and Completing in Other Ohio and Kentucky Institutions



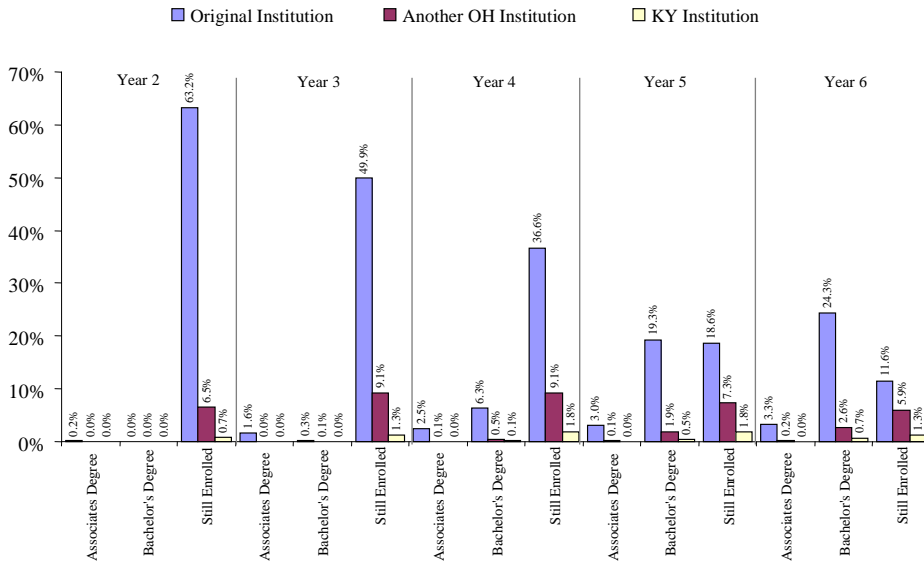
## Ohio Students Transferring and Completing Other Ohio and Kentucky Institutions (Percentages)



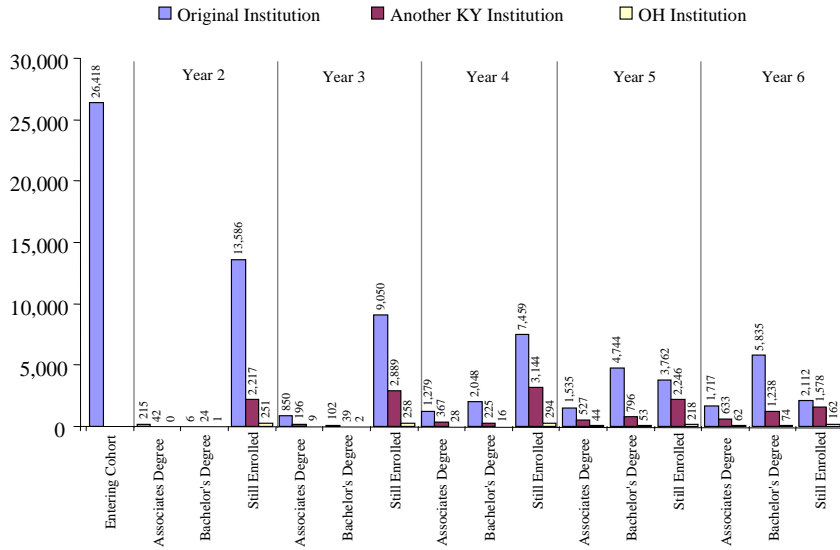
### Cincinnati Area Students Transferring and Completing in Other Ohio and Kentucky Institutions



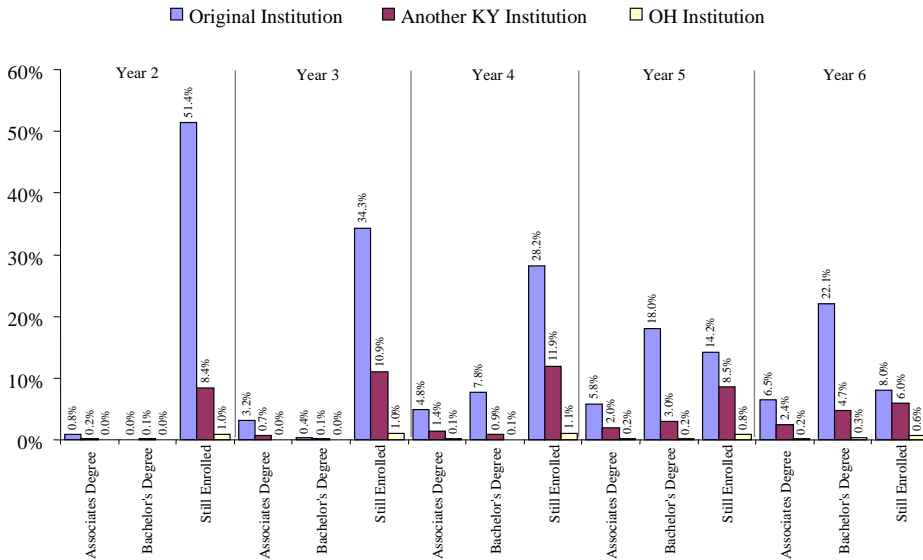
### Cincinnati Area Students Transferring and Completing in Other Ohio and Kentucky Institutions (Percentages)



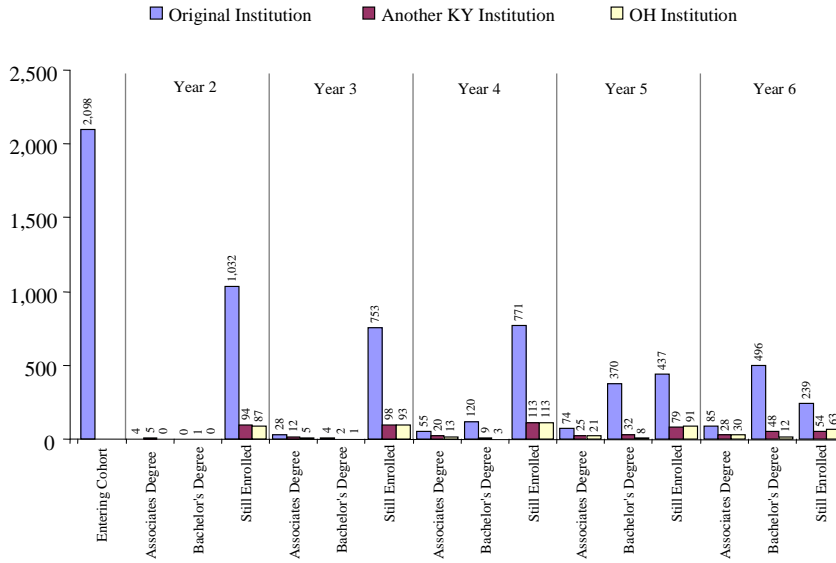
### Kentucky Students Transferring and Completing in Other Kentucky and Ohio Institutions



### Kentucky Students Transferring and Completing in Other Kentucky and Ohio Institutions (Percentages)



### Northern Kentucky University Students Transferring and Completing in Other Kentucky and Ohio Institutions



### Northern Kentucky University Students Transferring and Completing in Other Kentucky and Ohio Institutions (Percentages)

