



# EDUCATION

- THE ARTS
- CHILD POLICY
- CIVIL JUSTICE
- EDUCATION
- ENERGY AND ENVIRONMENT
- HEALTH AND HEALTH CARE
- INTERNATIONAL AFFAIRS
- NATIONAL SECURITY
- POPULATION AND AGING
- PUBLIC SAFETY
- SCIENCE AND TECHNOLOGY
- SUBSTANCE ABUSE
- TERRORISM AND HOMELAND SECURITY
- TRANSPORTATION AND INFRASTRUCTURE
- WORKFORCE AND WORKPLACE

This PDF document was made available from [www.rand.org](http://www.rand.org) as a public service of the RAND Corporation.

[Jump down to document](#) ▼

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world.

## Support RAND

[Purchase this document](#)

[Browse Books & Publications](#)

[Make a charitable contribution](#)

## For More Information

Visit RAND at [www.rand.org](http://www.rand.org)

Explore [RAND Education](#)

View [document details](#)

## Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND PDFs to a non-RAND Web site is prohibited. RAND PDFs are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the RAND Corporation monograph series. RAND monographs present major research findings that address the challenges facing the public and private sectors. All RAND monographs undergo rigorous peer review to ensure high standards for research quality and objectivity.

# Toward a K–20 Student Unit Record Data System for California

---

Georges Vernez, Cathy Krop, Mirka Vuollo, Janet S. Hansen

Prepared for The William and Flora Hewlett Foundation

Distribution unlimited



RAND

EDUCATION

The research described in this report was conducted within RAND Education and supported by The William and Flora Hewlett Foundation.

**Library of Congress Cataloging-in-Publication Data** is available for this publication.

ISBN 978-0-8330-4206-4

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

**RAND**® is a registered trademark.

© Copyright 2008 RAND Corporation

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from RAND.

Published 2008 by the RAND Corporation

1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

1200 South Hayes Street, Arlington, VA 22202-5050

4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665

RAND URL: <http://www.rand.org>

To order RAND documents or to obtain additional information, contact

Distribution Services: Telephone: (310) 451-7002;

Fax: (310) 451-6915; Email: [order@rand.org](mailto:order@rand.org)

## Summary

---

To provide the information needed to help expand educational opportunities and improve the quality of education from kindergarten through college, U.S. states are increasingly building robust student data systems. These systems are commonly termed “student unit record” (SUR) systems because they contain individual electronic records of each student enrolled in an educational institution. SUR data systems permit the tracking of an individual student’s progress over time—from entry into elementary school to exit from college and eventually into the labor market. These data are vital to attempts to formulate policy and answer questions at the core of studies of educational effectiveness.

Currently, 18 states are able to match individual student records between their K–12 and postsecondary education systems. Several other states are developing this capability. California, however, lags behind other states. Despite the California legislature’s clear intent that a longitudinal student data management system able to track students from elementary school through college be developed and maintained, the state is, at this time, neither planning nor implementing such a system. A K–20 SUR system would enable the state, education administrators, faculty, and the public to address questions that currently cannot be addressed with current data systems—questions such as how California’s students are progressing over time; how course articulation between high school and college can be improved; how effective are programs to retain, prepare, and help students succeed in college; what classes of students may need special intervention and attention; and how well are students prepared to meet California’s future labor demands.

At the request of The William and Flora Hewlett Foundation, the RAND Corporation undertook this study to identify the issues and challenges involved in developing an integrated K–20 SUR data system for California. To this end, we

- reviewed and documented the characteristics of the SUR data systems currently in use or under development in all four segments of California education: K–12 public schools, community colleges, the California State University, and the University of California
- interviewed the administrators of SUR systems in California and other states
- reviewed the literature on longitudinal SUR systems and the experiences of various states with these systems
- interviewed key state-level stakeholders, leaders of education segments in California, and researchers.

## **California’s Current Student Data Systems**

California currently has seven student data systems either in use or under development:

- The California School Information System (CSIS), a mandated public K–12 system that contains limited enrollment and demographic data on all students and additional data on 60 percent of those students in districts participating voluntarily—currently in use
- The California Longitudinal Pupil Achievement Data System (CALPADS), a mandated public K–12 system—under development
- One system for each of California’s three public higher-education segments, the community colleges, California State University, and the University of California—all currently in use
- The California Post-Secondary Education Commission student database, a partially integrated system that brings together

student-level data from all three higher education segments—currently in use

- The California Partnership for Achieving Student Success (CalPASS), a voluntary and regional system that fully integrates student data from individual institutions in all four of California’s educational segments—currently in use and expanding.

Detailed information on each of the above student data systems is contained in Chapter Two.

## **Major Challenges and System Design Issues**

These seven systems give California the embryonic architecture that would enable it to develop and maintain a comprehensive K–20 longitudinal SUR data system. The state also has the necessary technical expertise. Furthermore, our respondents indicated that there are no technological barriers to developing such a data system.

Still, a number of major challenges will have to be addressed before an integrated K–20 SUR system can become a reality in California. Each of the state’s four education segments has developed its own policies and administrative practices and all have developed strong separate cultures and identities as well as a protective mindset. A consensus among all stakeholders that building such a system is desirable in the first place will have to be developed. Beyond that point, it is also likely to be difficult to develop a consensus on the purpose of the system—who should hold decisionmaking authority over its development and maintenance, who should have access to it, and who should operate it—because this may require that the current overseers of the four education segments give up some decisionmaking authority. Given California’s budget woes and its experience with cost overruns when developing some previous state data systems, providing the necessary funding may well prove to be a roadblock, even though the funds would be minimal compared with the billions spent on education in the state.

In addition, the state will need to gain a consensus among the four education segments and other stakeholders on various system design issues:

- determining how comprehensive the system should be
- standardizing data elements across segments
- improving the quality of the data collected
- selecting a common student identifier across segments
- protecting student privacy.

But as experience in other states bears witness, all of these design issues are solvable.

## **What Next for California?**

Overcoming the interrelated challenges involved in developing an integrated K–20 SUR data system in California, and eventually one even more comprehensive, may be a daunting task that could take several years. In our view, the state will need to take several steps to resolve these challenges:

### **1. Complete the Design and Implementation of CALPADS**

The first step is to complete the design and implementation of CALPADS. Much remains to be accomplished before the system is operational in 2009–10 as anticipated—including selecting a vendor, which is not expected until fall 2007, and providing the funding to build the capability of Local Education Agencies (LEAs) and schools to collect and transmit student data electronically. Funding for this purpose was deleted from the state’s 2007–08 budget.

### **2. Identify a Champion to Be an Advocate for a K–20 Data SUR**

According to our California respondents, getting the process started will require an influential individual or group of individuals to perform several key functions:

- champion the development of an integrated K–20 SUR data system
- gather the necessary support from stakeholders and the public
- develop a consensus on the system’s governance and design
- sponsor the legislative action that would mandate and define its purpose and capabilities and identify its limitations.

Potential candidates include the governor, a well-connected business leader, or a legislator. Alternatively, a K–20 SUR could be championed by a commission that might include a representation of committed stakeholders and potential users.

### **3. Obtain Legislative Authority**

Even though new legislative action is not needed to develop and maintain a K–20 SUR data system, our California respondents agreed unanimously that such legislation would be required to enable the state to develop an *integrated* K–20 SUR system. Many of the issues involved in developing such a data system are so sensitive and important that they need to be aired, debated, and eventually resolved in an open public debate. Nor is it likely that the disparate views on governance of the data system (i.e., who holds decisionmaking authority and who owns the system and access to the data) could be resolved without legislative authority.

The California legislature would need to vet and address the following questions:

- What purposes would a K–20 SUR data system primarily serve?
- Who should have decisionmaking authority over the design of the system?
- Who should operate the system and where should it be located?
- Who should have access to the data it contains?
- What common student identifier should be used?
- What is the minimal content to include in the data file?
- What level of funding should be allocated to develop, maintain, and use the system?

In this report, we have presented a number of possible answers to these questions—various options that the legislature may consider, some that would lead to a more comprehensive integrated K–20 system and some to a narrower set of linkable individual student databases.

#### **4. Build the K–20 Student Data System Incrementally**

It will take several years to fully develop an integrated K–20 data system that has maximal utility for policy and administrative decisionmakers at the state as well as at the segmental and individual institution levels. Consequently, we recommend taking an incremental approach. A possible development sequence over a period of four to five years would be:

- Integrate the four existing segmental data systems “as is.” Even this simple merger of the three current postsecondary student data files with CALPADS would provide useful and more accurate information than is currently available on dropout and transfer rates, student mobility, and educational progress.
- Add data elements already being collected at the individual school and campus levels but not currently forwarded to the central student database of each segment—such as individual courses and grades.
- Link the K–20 student data file to other state and federal data systems—such as preschool, employment and wage, welfare, foster care, corrections, military, and private university.
- Add data elements not currently being collected by one or more the segments—but only if deemed desirable and cost-effective after several years of experience using the data system and after the three steps just outlined have been implemented.

Overall, California should take a long-term perspective toward improving the quality and standardization of the data that schools and campuses collect. In the short term, those responsible for collecting and entering data should receive training to ensure that they use appropriate quality control and privacy protection practices. But experience suggests that such training will go only so far and that the quality

of the data will improve primarily through use and user feedback to schools and campuses.

### **5. Develop an “Objective” Analytical Capability and Expertise**

To make maximum use of the integrated K–20 student data and effectively respond to the needs of individual schools, colleges, universities, and state officials, California will need to take one last step: An analytical capability, adequately funded, should be developed either within or independent of the organization operating the K–20 data system. Not only will this capability enable those assigned this responsibility to develop the information required for policymaking, it would also address one of the principal concerns of our respondents in the state’s higher education segments—that their data might be misinterpreted either because users are unfamiliar with how a segment operates or because the segment has made administrative or policy changes. With an experienced analytical team, the frequency of such problems would be significantly reduced.

Experience with the use of SUR data systems in California and other states suggests that ten to 20 analysts might be required to provide the level of capability and expertise required.