CSP
Court Statistics Project

Caseload Highlights
Notes from the Field

**By practitioners —
for practitioners.**

www.courtstatistics.org

Creating and Sustaining Data Quality

In July 2009, Richard Schauffler, director of the Court Statistics Project (CSP), interviewed Craig Burlingame, the chief information officer of the state courts in Massachusetts. Among his other duties, Mr. Burlingame oversees the development of the new statewide case management system, MassCourts, and the IT infrastructure to support the statewide performance measurement system, CourtMetrics.

Richard Schauffler (CSP):

Massachusetts is in transition to a new statewide case management system called MassCourts. The issue of data quality—do you fix the current data now, never, or along the way?—must have been an integral part of that transition. How did you approach it?

Craig Burlingame (CB):

You have to recognize that the data that you have is the data that you have. For integrity purposes, we did some clean up as part of the conversion process to our new case management system, doing as much straight conversion of the historical data as we could, correcting the severe impurities along the way, but not spending a lot of time doing extensive clean up on all the legacy data.
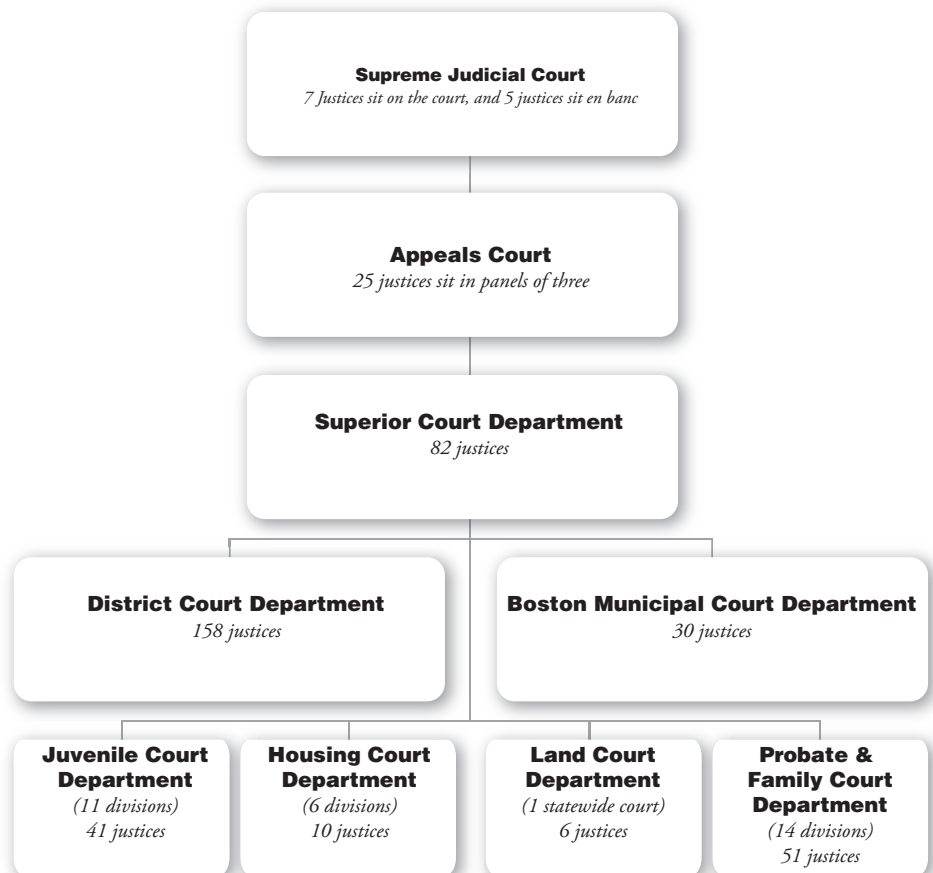
This is an iterative process; it is never going to be perfect the first time. In fact, I think in a lot of this "the perfect is the enemy of the good" maxim plays out. People overthink some of these decisions about how to get it perfect the first time; the reality is that there isn't a way to get it perfect the first time. You try to make it better than it is now, and then go through some iterations once you've had some experience using the data differently, and make further perfections along the way. To try over and over to get it perfect from the beginning is really just a waste of time.
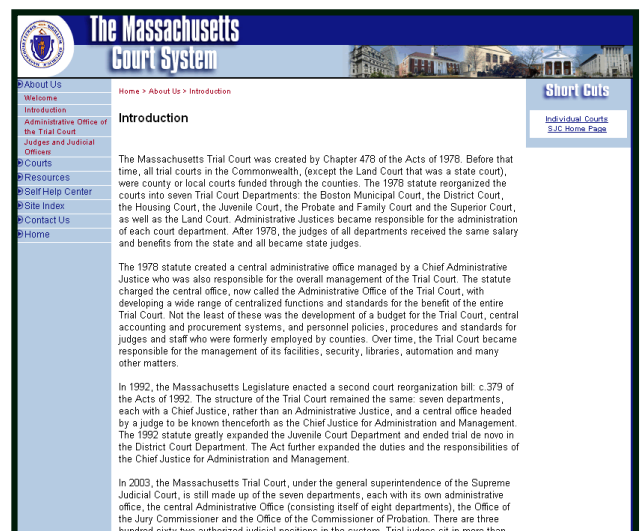
CSP: What kinds of data problems did you encounter in this transition?

CB: The biggest thing was cross-referencing between separate legacy systems. For example, on the criminal side our historical criminal database had been maintained by probation, but we also had a warrant management database that had been maintained by the clerk's office. Both databases had docket numbers, case numbers, and person identifiers that needed to be reconciled. We didn't want to create multiple instances of the same John Doe just because of slight variations in the spelling of the name. We wanted to ensure referential integrity and that's where most of the clean up happened.

Massachusetts Court Structure

Supreme Judicial Court
7 Justices sit on the court, and 5 justices sit en banc

Appeals Court
25 justices sit in panels of three

Superior Court Department
82 justices

District Court Department
158 justices

Boston Municipal Court Department
30 justices

Juvenile Court Department
(11 divisions)
41 justices

Housing Court Department
(6 divisions)
10 justices

Land Court Department
(1 statewide court)
6 justices

Probate & Family Court Department
(14 divisions)
51 justices

CSP: In civil cases, a common data quality problem is the absence of a final dispositional entry. Did you find problems like that?

CB: Yes, during the conversion process we worked with each of the trial court departments on that issue. In Massachusetts we have seven trial court departments that are responsible for different caseloads. For example, we have a department that handles juvenile cases, another that handles housing cases, and so on. We worked

with each department to establish some business rules for the conversion clean up. For example, if a case had certain characteristics and had been inactive for a certain period of time, that case would be recorded as a disposed case, using the proper kind of entry.

CSP: That kind of collaboration seems critical. How do you start to create the kind of organizational culture in which people see that the data is something with value, something that's practical, not just something they're doing for whoever they think cares about the numbers?

> **"If the critical data is right, the workflow is automatic; that is what provides the incentive to get it right and keep on getting it right."**

CB: I think you can use technology to an extent. For example, as part of the new MassCourts case management system, we've had a focus on identifying interfaces where correctly entered data can trigger further business activity—that is, it can automate subsequent uses of the data. For example, in January 2009 we started transmitting dispositional data to our state's department of motor vehicles for certain mandated dispositions in specific case types. We've reported about 70,000 of those dispositions, all triggered by the court staff simply entering the appropriate disposition code for the appropriate kind of charges. The system knows the business rules: which kinds of charges need to be reported; which kinds of charges need to be reported only when there's a conviction-type disposition; and which kinds of charges need to be reported no matter what the disposition. So just by virtue of the staff performing the business function for the court using the court's software application, the Department of Motor Vehicles (DMV) is also notified. Now, the users know that not only do they no longer have the additional task of manual reporting (formerly, a paper-based transaction), they also know that as long as they enter the correct data, that disposition is going to be automatically reported that night to the DMV.

There's a data quality component to this which is really important. For the DMV to do what they need to do they need certain pieces of accurate data as part of that case—a driver's license number is a significant data element for us. So we've developed some data quality tools within the framework of the application that help the court focus on specific data elements in certain cases without having to worry about some of those data elements in other cases where they're not as important. If the critical data is right, the workflow is automatic; that is what provides the incentive to get it right and keep on getting it right.

CSP: Do the court staff that are generating those records see error reports or somehow get a sense of how well they are doing?

CB: They do. We get a feedback report from the DMV, but more than that we use a tickler function that was already part of our software application. We built some of our own data quality queries that examine the cases that are candidates for reporting, even though the dispositions in those cases are not entered yet. These queries evaluate the relevant data that exists in those cases and identifies where the data is deficient. The system adds a tickler to those cases that articulates what the flaws are, for example, missing data or an invalid citation number. Court personnel know that whenever they notice one of these data quality ticklers, they should make that data correction prior to entering a new disposition or otherwise updating one of these cases. The nice part is that those ticklers are uniquely articulated, each with its own kind of code known to the staff. If the staff have some free time, they can run periodic reports and can do some clean-up work. Cleaning up the data is built into the daily workflow in this way; it is not an endless "special project."

CSP: You seem to be putting the control and understanding of the process at the right level. Court staff are not waiting for some central authority to run a report and then send it to them and tell them what it means or what to do.

CB: That's true. Those kind of report-based data quality initiatives are in some ways burdensome. For the court staff, it means "I get this report, I've got to go pull those cases, I've got to find the records, I've got to find the corrected data, and I've got to fix it to get that case off the report." In our approach, because we're just flagging the case, the staff are not dependent on the report nor do they need to wait for another cycle of audit-report-retrieval-reenter-refile. Staff can fix the data quality issue when they are in the case in the course of regular business.

> ### "Cleaning up the data is built into the daily workflow in this way; it is not an endless 'special project.'"

CSP: Is the idea to adapt and extend that approach to the other kinds of cases?

CB: Exactly. This model is really a test to see if this method of approaching specialized data quality works. Clearly there are some generic data quality issues, data that needs to be in specific formats and certain codes that must be used according to specific business rules. Those issues are true across all case types. For now we are taking a targeted approach to narrow the focus of the data quality initiative.

The whole concept of designing interfaces to take burdens off court personnel, especially the burden of reporting, is a good way to get people to focus on data quality. Staff know they're not just entering the data for the purpose of having it in the database; they're also entering the data to drive the follow-on business processes.

So if you get the data in and you get it in right, those follow-on things are going to take care of themselves and you're not going to have that burden anymore.

CSP: How do you make sure the data entry codes are correct, and get people to use the right codes?

"The key to future data quality...is in the way you approach system design and code definition."

CB: On all of our application screens, all the codes are always translated into plain English. As soon as the staff person enters the code they see its descriptor; if the descriptor does not match the action that was intended, it is changed before it becomes embedded in the record. In a lot of legacy systems all that is recorded is the code; if you make a mistake, you don't really notice it.

You should never leave the user without the ability to do lookups. Whenever a user enters a code in our application, there's a supporting search function that can be used to locate a descriptive explanation of the various codes and what their values mean within the application via a pop-up window. If we make it easy in the interface, staff are much more likely to double-check and do a search to find the correct code. If you simply design these applications with a data field into which the user is supposed to enter code in A through Z, the likelihood that the right codes are going to get used is low.

CSP: How can courts do a better job of getting the interface to help enforce the data definitions and the business rules?

CB: The key to future data quality and correct code use is in the way you approach system design and code definition. We went through a fairly systematic process, even though we're using a vendor-based application. We built an eight-step implementation process for each of the departments as we transitioned them into the system. In the middle of those eight steps there is what we call the alignment process. Basically, this is when we look at their existing data and their existing business practices and make sure these are well aligned with the way the system works. The good part about the application we are using is that it is very soft, very code driven, very customizable in terms of the quantity and types of codes, and the derivative actions that happen after codes are used.

But therein lies part of the challenge. Because the software is so soft and there is so much flexibility, there is a lot of decision-making up front as to how it's going to be configured. So we go through this alignment process to make sure we've got the functionality and coding the way they need to be. And then we do a "model office"

in which we have people bring in real cases and sit down and use the system the way it's been configured. This is a validation cycle to ensure that we've accounted for all the types of dispositions, codes, reports, notices, etc. and that these can be handled properly to meet the court's business requirements before we ever go forward.

CSP: When you were developing these "model offices," did you discover that courts in the same trial court department were in fact doing things differently?

CB: Absolutely. In fact we had departments that were insistent they had uniform practice across all their divisions. But we could tell by looking at the legacy data there wasn't always consistency across divisions. More importantly, when we put these groups together during the alignment process (a mix of administrative office personnel and business process experts we selected from court divisions), it was not long before the differences emerged when trying to define "how we do it." When we got to what kinds of codes we were going to use for events on the court calendar, it was not too long before somebody said, "Well wait a minute, we don't use that kind of an event for this, we use this other kind of thing." We used those forums as an opportunity to talk through these issues and have the participants come to agreement as to how the system was going to be configured for all the divisions. We have insisted on a common configuration for each of the divisions within a department. We have not created local customizations just because Division X wants to do it differently than Division Y. That probably has been the hardest part of all this--getting all the divisions to come to agreement as to how the system should work for everybody.

CSP: How does this variation get started and what perpetuates it?

CB: I think over time people develop their own way of doing things that works for them. It doesn't mean they think it's the best way, it's just the way they've done it. So any alternative is going to be a change. And courts have change-averse cultures. A person may not have a good reason why they do things the way they do, but they also know they don't like change. Through collective discussion the vast majority of the time you come to some agreement as to what works for everybody.

From time to time you come to the situation where Judge X tells me I have to do it this way. And that's where you end up using the AOC participation in the process to reach out to Judge X and say wait a minute, we can't have local practices dictate the way we handle systems. We want personnel to be able to move from place to place and we want a court to be able to read data that comes from another court. And if you insist on doing it differently than everybody else, you're creating inconsistencies, misunderstanding, and so on. More often than not the AOC is successful in

explaining the statewide
perspective; the judge or staff
person who insists that
things be done their way just
doesn't understand that there are
global consequences
to that view.

CSP: How is training of court
personnel handled, from the
point of view of making sure staff
throughout the state know how
the system should be used and
making sure
that knowledge gets institutionalized?

CB: I think people learn better from their peers than they do from IT people like me.
We don't always talk in the terminology that's right for them anyway. So it's better
to have a clerk magistrate from a court that's a really good user, for example, or an
identity specialist from a probation department that's a really good identity specialist,
to come in and talk to their peers and their colleagues about how they use the system
as opposed to some techie that shows up and says here's the right widget to click on.

From the very beginning we designated what we called local user experts. There's
a local user expert in every court division that's using MassCourts. We did extra
training for them; we brought them into training ahead of their court so they could
see other users learn. We brought them in to work on implementations in other
courts so they could participate in go-lives and cut-overs in advance of getting to that
stage in their own courts. We then made them the focal point for support purposes
for both data quality questions and questions about how to use the system. So these
local user experts are a standing population of experts that we tap for all kinds of
things—to help us in communication, to help us in addressing data quality issues,
and to help us in data analysis questions. This has worked really well; it gives us
points of contact in each court, and it builds up local expertise that would have been
spread thin among a bunch of people if we didn't create a single resource who is
the go-to person. Now, if two employees in the same court have the same question
at different points in time, they're going to the same person and getting the same
answer. That question is getting resolved locally and they don't have to wait for
someone at the state level to get back to them.

Metric 1 : Clearance Rate

Q3 Case Activity June 30, 2009 to September 30, 2009

Court Department	New Cases	Disposed Cases	Clearance Rate	Clearance Rate Goal	Progress Towards Goal
Boston Municipal Court					
Civil	15,336	15,173	99%	105%	
Criminal	9,118	9,810	108%	105%	**Goal Achieved**
Sub-Total	24,454	24,983	102%	105%	
District Court					
Civil	31,283	32,362	103%	105%	
Criminal	48,295	46,963	97%	105%	
Sub-Total	79,578	79,325	100%	105%	
Housing	11,535	10,944	95%	105%	
Juvenile					
Civil	2,591	3,152	122%	105%	**Goal Achieved**
Criminal	5,495	5,657	103%	105%	
Sub-Total	8,086	8,809	109%	105%	**Goal Achieved**
Land	8,636	6,393	74%	105%	
Probate and Family	16,032	13,644	85%	105%	
Superior					
Civil	6,473	6,182	96%	105%	
Criminal	1,282	1,166	91%	105%	
Sub-Total	7,755	7,348	95%	105%	
Total	156,076	151,446	97%	105%	

National Center for State Courts

300 Newport Avenue

Williamsburg, Va 23185-4147

CSP
Court Statistics Project

RESEARCH DIVISION
800.616.6109

Court Statistics Project

Since 1975, the Court Statistics Project (CSP) has provided a comprehensive analysis of the work of state courts by gathering caseload data and creating meaningful comparisons for identifying trends, comparing caseloads, and highlighting policy issues. The CSP is supported by the Bureau of Justice Statistics and obtains policy direction from the Conference of State Court Administrators. A complete annual analysis of the work of the state trial and appellate courts will be found in *Examining the Work: An Analysis of 2007 State Court Caseloads*

Summing Up

CSP: You are launching new initiatives and everyone's excited. How do you sustain that over time, how do you institutionalize the new way of doing business?

CB: I don't know if there's magic bullet for this. I think the key is to have the relevant data used on a regular basis as a feedback loop to give the court information about their own activities. If I know these key data elements are going to be used to do statistical reports and to feed into an analysis of time standards, I'm going to be focused on maintaining that data accurately because it produces a byproduct that is of value to my court. For example, if I know my percentage of cases that are within time standards and I can get a sorted list of the oldest cases that are inactive, I'm going to be more focused on maintaining data quality permanently. The key to having people care about data quality is creating feedback loops not only for the AOC but for the local courts themselves.