



Center for
Educator Compensation
Reform

TIF Challenges in Education Information Systems and Knowledge Management

An Introduction to Quality Data

Sara Kraemer

TIF Decision Support User Needs

- Transparency
- Validity of metrics
- Accuracy and replication
- Responsiveness and timeliness
- Granularity
- Interoperability

What are Quality Data?



- Accurate – Are the data right?
- Granular – Are the data detailed enough?
- Valid – Do the data represent reality?
- Integrated – Can data from multiple systems be connected?
- Relational – How does the organization of data affect data utility?
- Reducible – How can districts reduce data burden into meaningful analytics?
- Actionable – Do data consumers know what to do?

Challenges to Success in Decision Support



- Challenges can co-exist and compound each other
- Have social/organizational as well as technical roots
- Should be prioritized given grantees' constraints, priorities, and theory of action

For more on Data Quality



- View the videos on:
 - A Real Life Example of Data Quality Issues
 - Solutions for Data Quality Problems
 - Teacher Classification
 - Establishing Student-Teacher Linkages
 - A Real Life Example of Student-Teacher Linkages
- Read the Guide to Implementation module on Data Quality

TIF Challenges in Education Information Systems and Knowledge Management

A Real Life Example of Quality Data

Sara Kraemer

Example 1: Connecting Teacher Data From SIS and HR



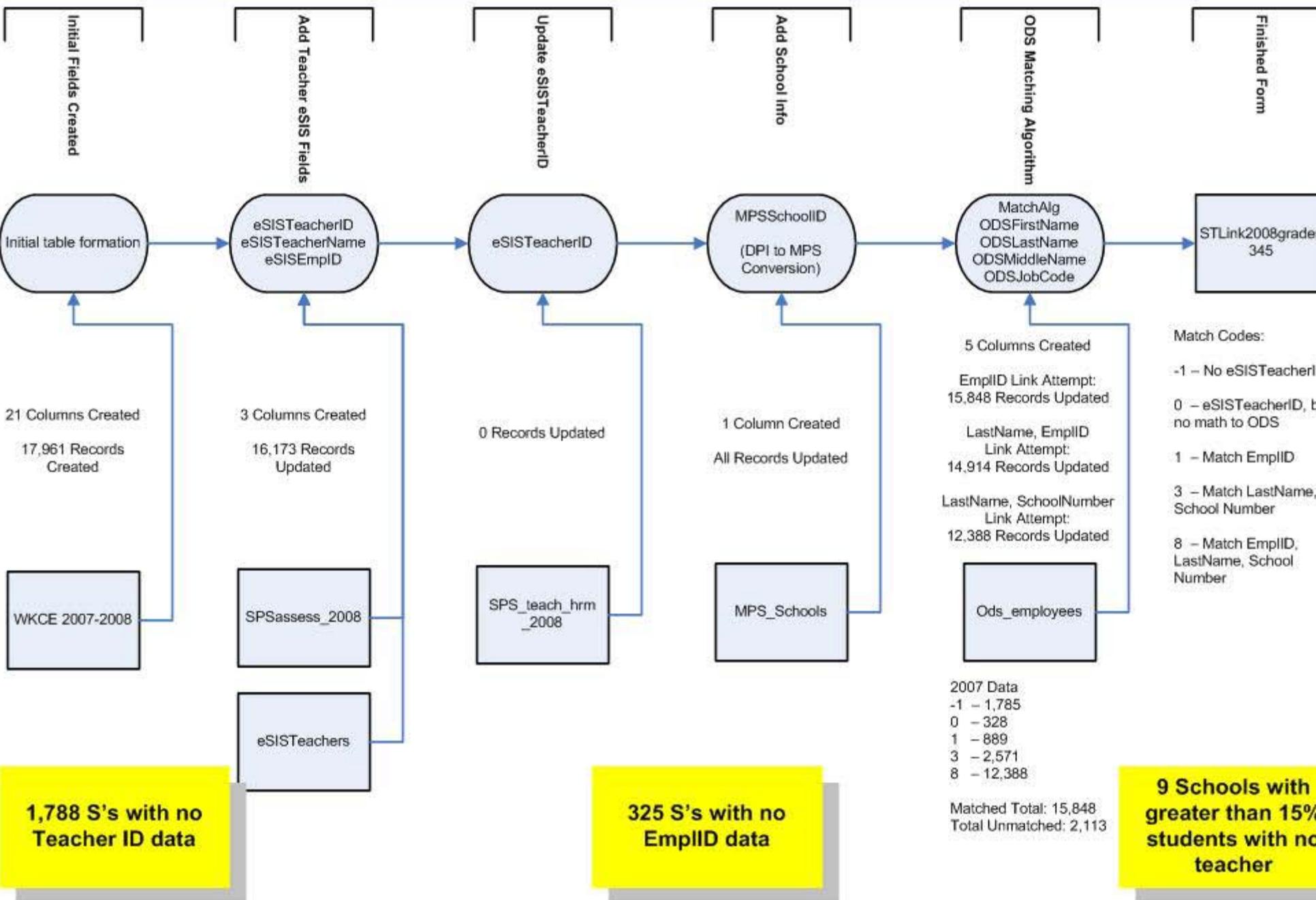
- Teachers in HR system did not match teachers in SIS ($\approx 70\%$ matched)
- Context:
 - Human Resources system (PeopleSoft) creates persistent and unique IDs (aka emplIDs)
 - SIS (eSIS) creates non-persistent but unique IDs (aka TeacherIDs)
 - School staff create and manage TeacherIDs throughout the school year, but especially during scheduling periods. Complex workflow not well represented by SIS interface
- Challenges represented: accuracy, validity, integration

Example 1 (continued)

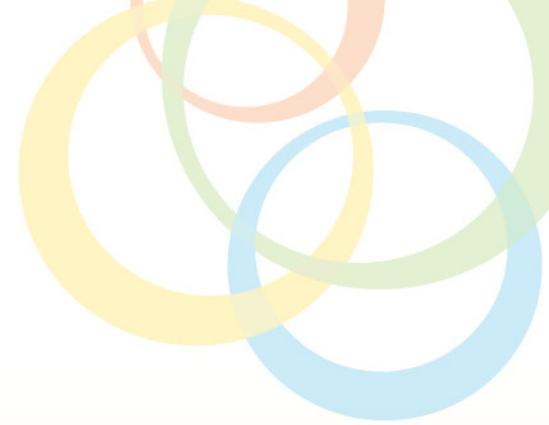


- School staff use SIS in a way that meets local scheduling needs – here are some actual teacher ‘names’:
 - Tch A - MRP2, Tch B - MRP1, Tch C - Sci6B, Tch D – Orchestra
- Some buildings use organizational structures that are not manageable with the data structure provided by the district
- Analyses:
 - Analyze matching patterns – Where is matching best? Worst?
 - For teachers assigned a grade level in SIS, roughly 15% (≈ 500) cannot be matched; disproportionate number in 8th and 9th
 - For teachers with no grade level in SIS, 55% ($\approx 1,500$) cannot be matched
 - Analyze workflows that affect data quality – Why is data quality compromised?
 - Create process flows for major tasks at schools such as scheduling, creating new rosters, keying teacher information

Creation of StudentTeacherLink2008grades345



For more on Data Quality



- View the videos on:
 - An Introduction to Quality Data
 - Solutions for Data Quality Problems
 - Teacher Classification
 - Establishing Student-Teacher Linkages
 - A Real Life Example of Student-Teacher Linkages
- Read the Guide to Implementation module on Data Quality

TIF Challenges in Education Information Systems and Knowledge Management

Solutions for Data Quality Problems

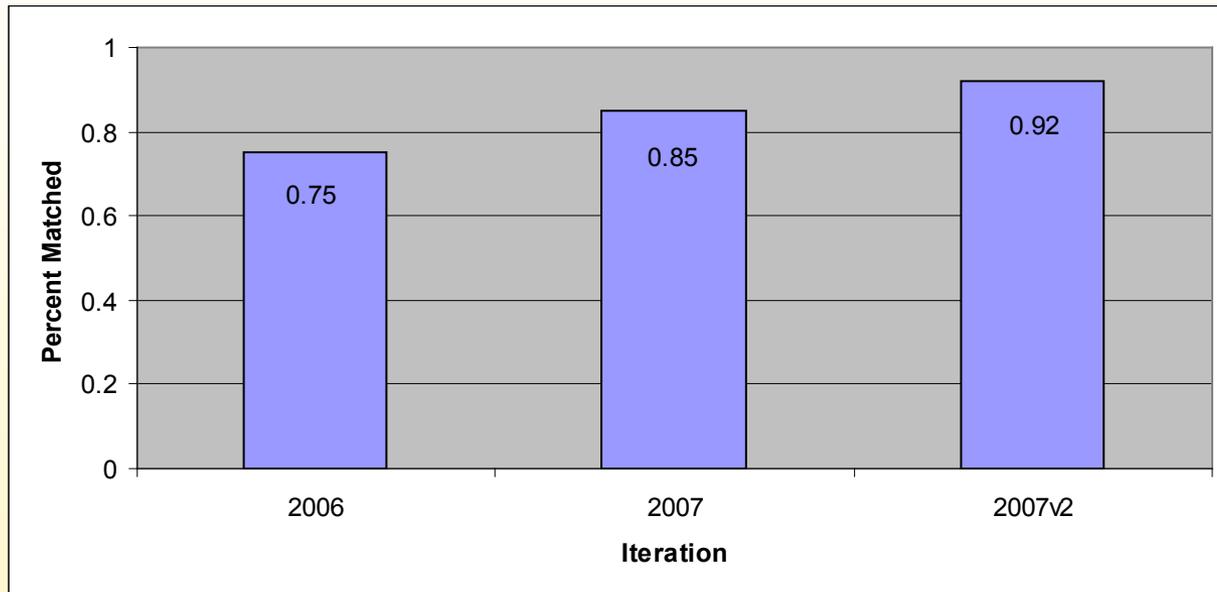
Sara Kraemer

Example 1 (continued)



- Solutions:
 - Build data quality checks for data-entry screens (e.g., leverage Oracle exception error) that use look up tables (improves integration)
 - Create data quality management tools (e.g., reports, training procedures)
 - Build support of stakeholders to emphasize quality – e.g., training, tech support,
 - Identify true needs of schools (e.g., scheduling logistics) and develop use-cases
 - Provide feedback to SIS vendor to improve underlying SIS data model

Percent of teachers with accurate HR data in SIS



For more on Data Quality



- View the videos on:
 - An Introduction to Quality Data
 - A Real Life Example of Data Quality Issues
 - Teacher Classification
 - Establishing Student-Teacher Linkages
 - A Real Life Example of Student-Teacher Linkages
- Read the Guide to Implementation module on Data Quality

TIF Challenges in Education Information Systems and Knowledge Management

Establishing Student-Teacher Linkages

Jeffery Watson

Example 2: Connecting Teachers to Students



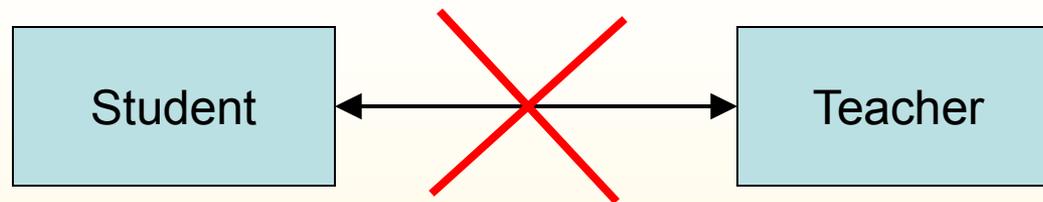
- Knowing what teachers taught what students is a critical linkage for TIF projects
- Context
 - Schools use a variety of organizational designs
 - SIS data structures for enrollment data may not capture non-traditional instructional models
 - Additional programs (e.g., after-school activities, pull-out specialists) exist
- Challenges: Validity, granularity, quality

Defining Student-Teacher Linkages

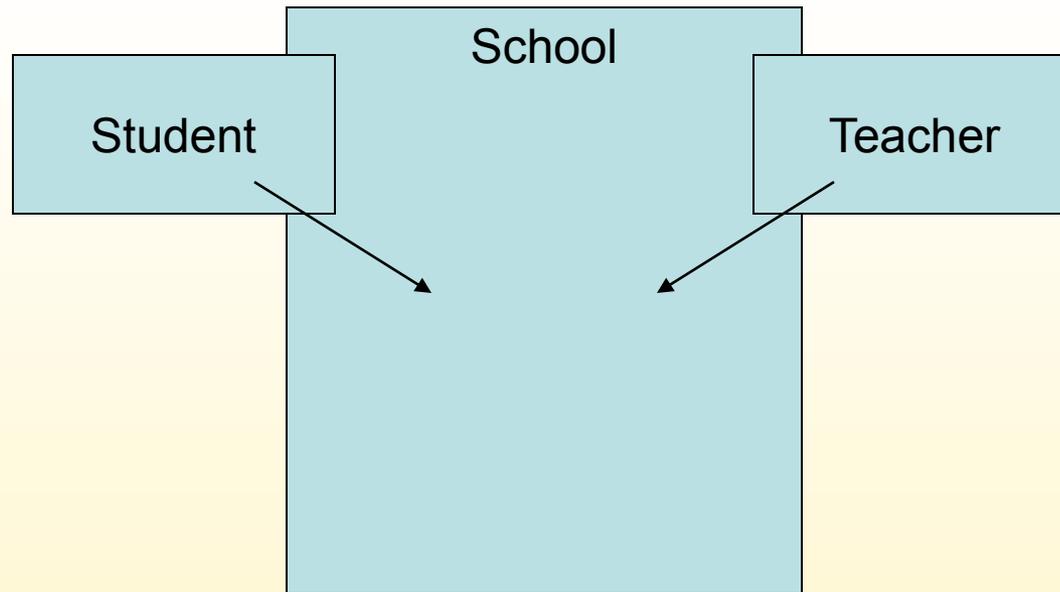


A record of which teachers and staff taught which students during a school year.

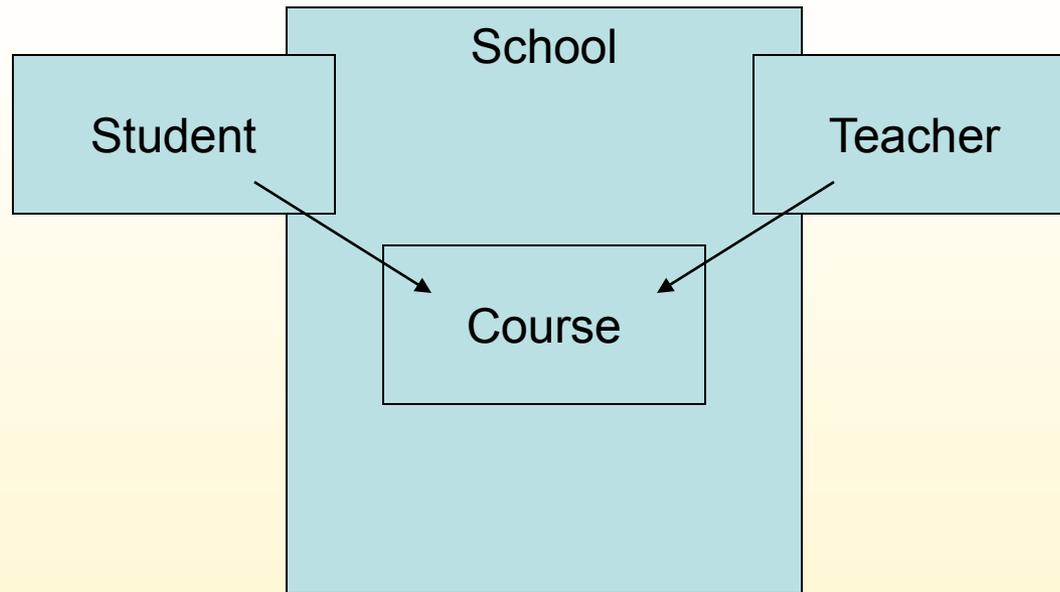
Defining Student-Teacher Linkages



Defining Student-Teacher Linkages



Defining Student-Teacher Linkages

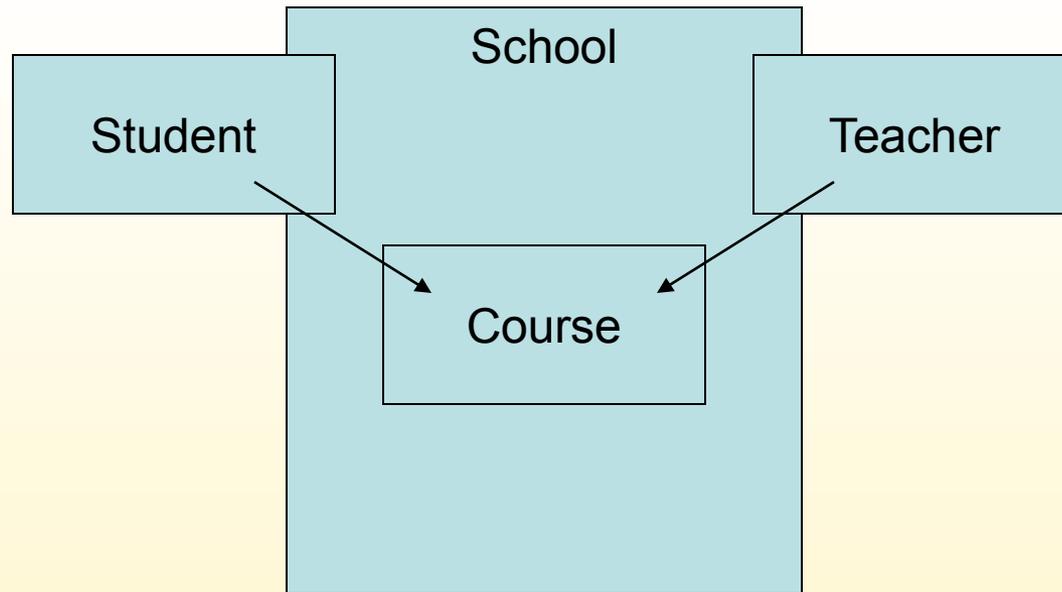


Four Easy Questions

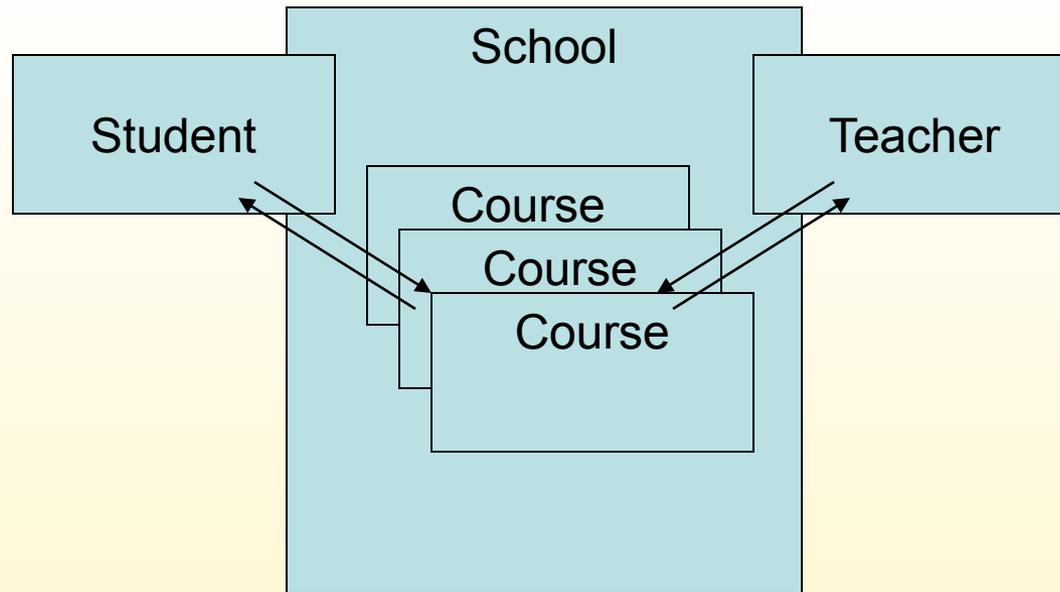
- What is a student?
- What is a teacher?
- What is a school?
- What is a course?



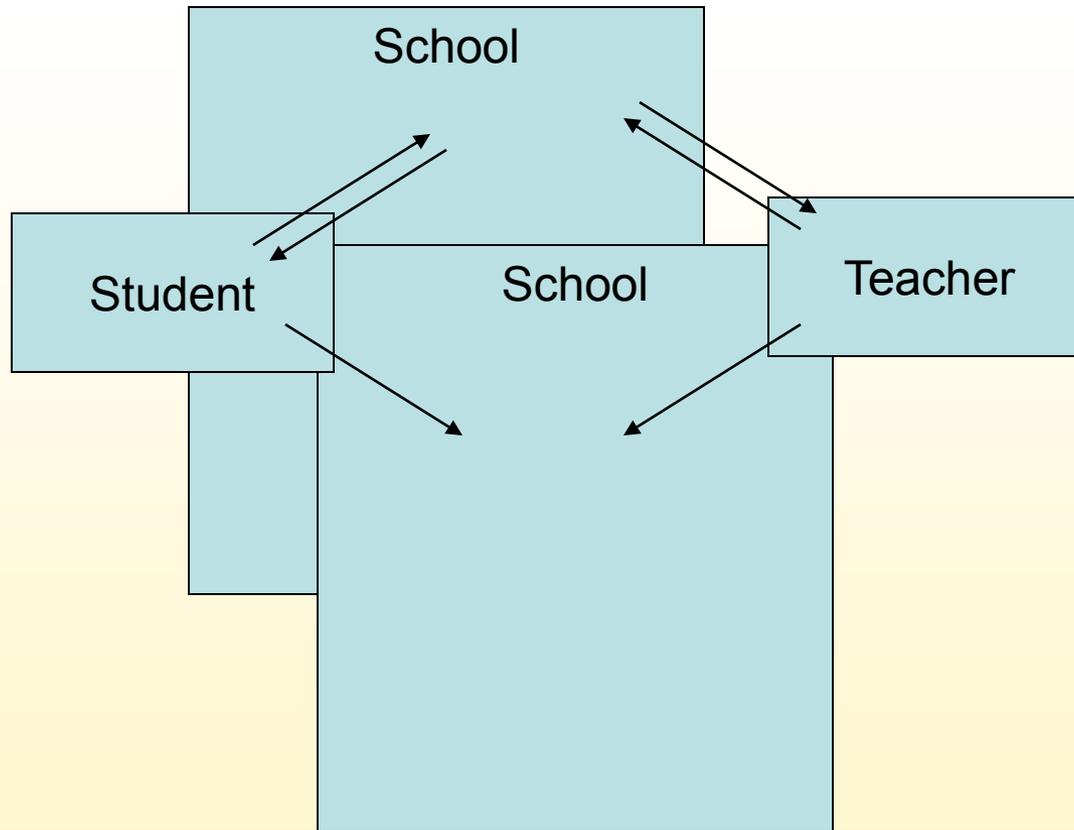
Warning: Reality Approaching



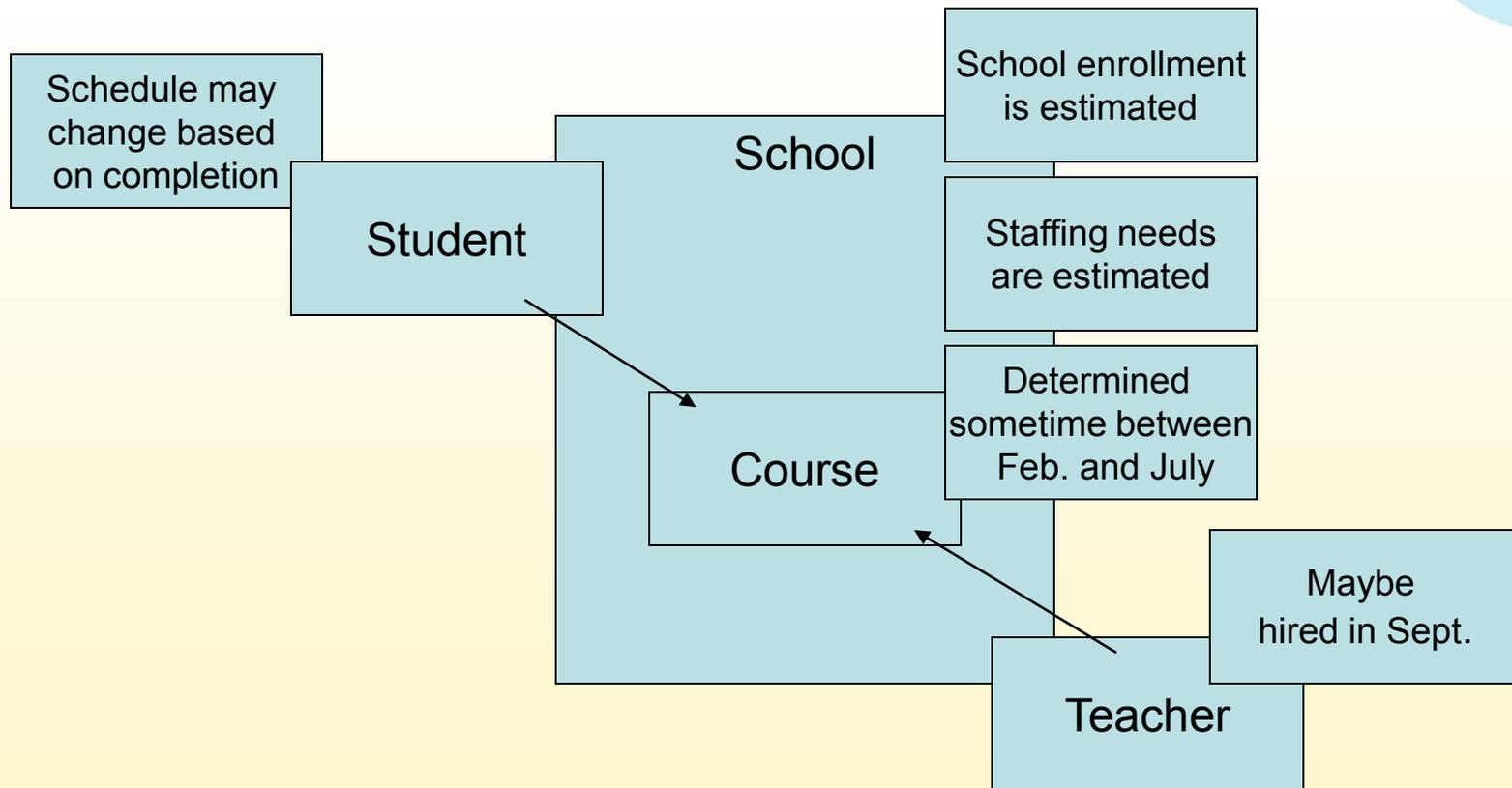
... movement within schools

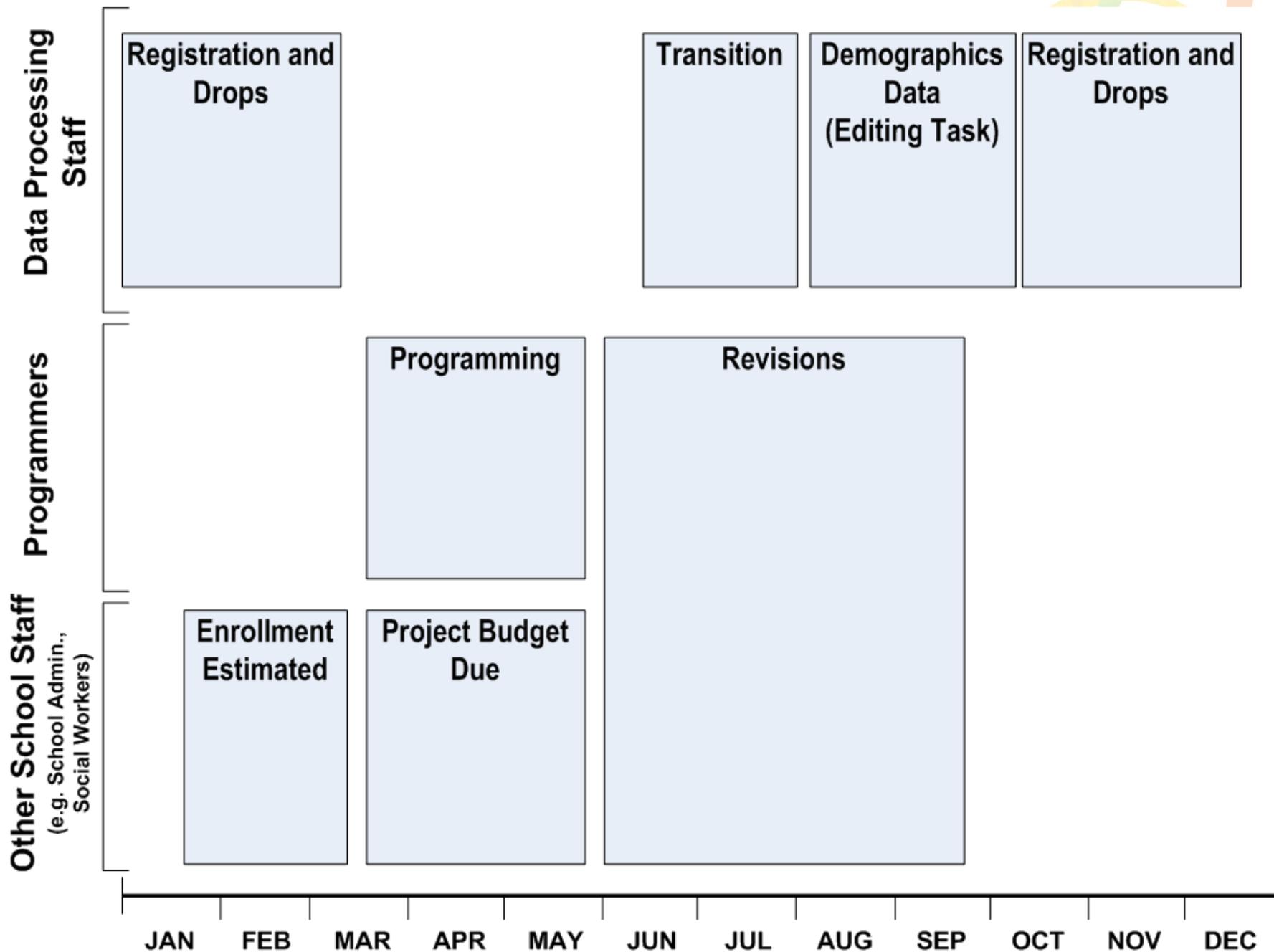


... and movement between schools

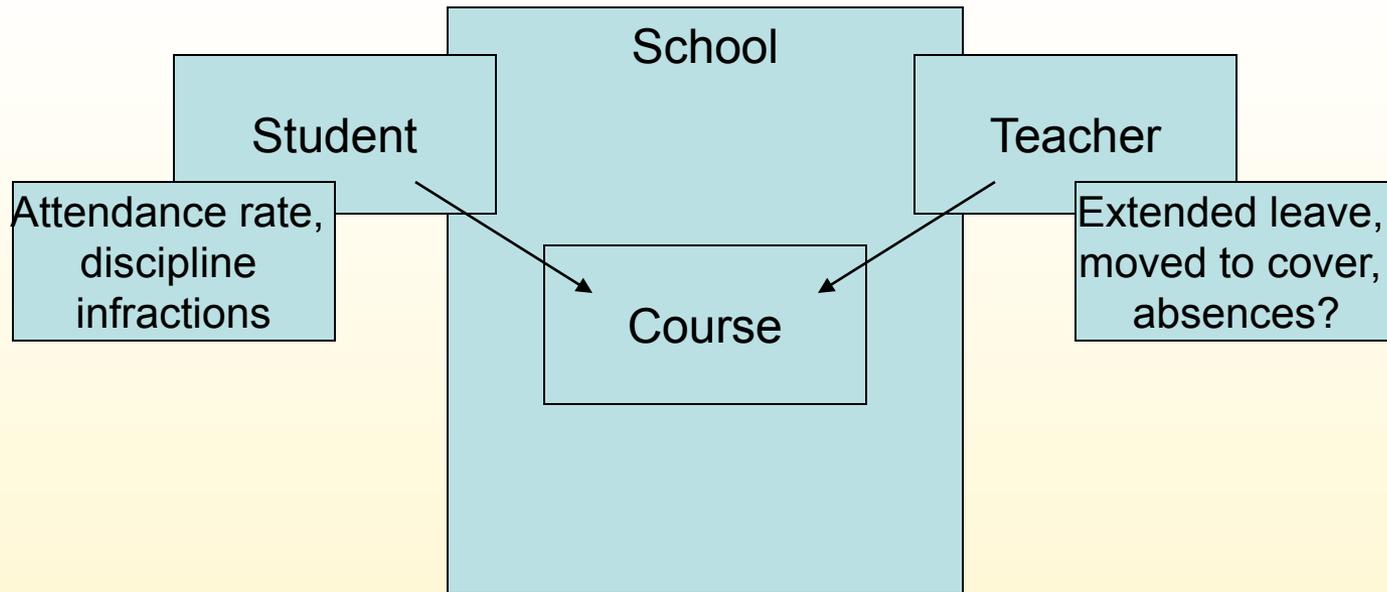


... and complicated workflows

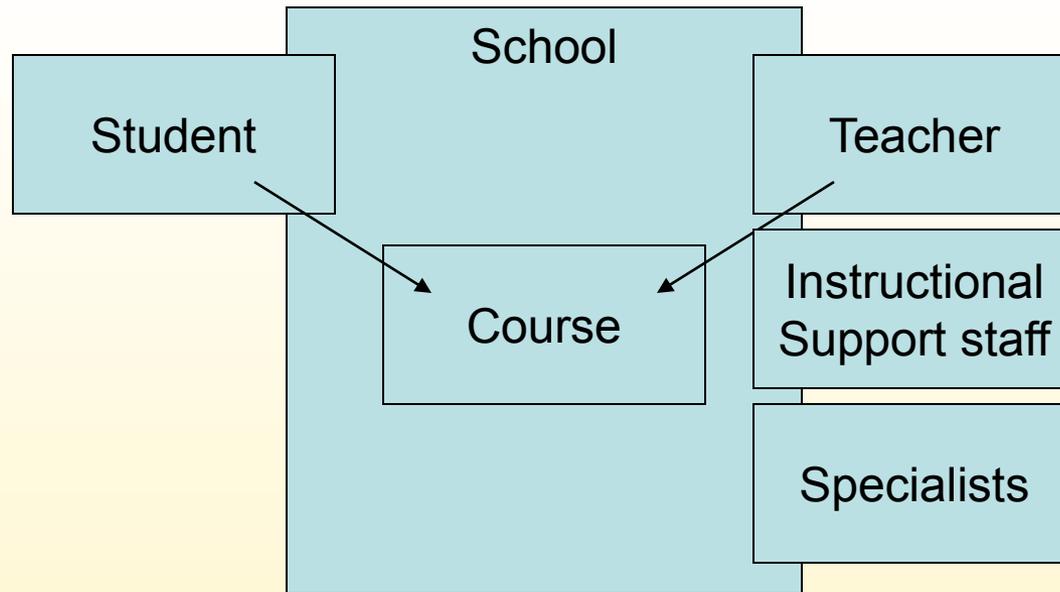




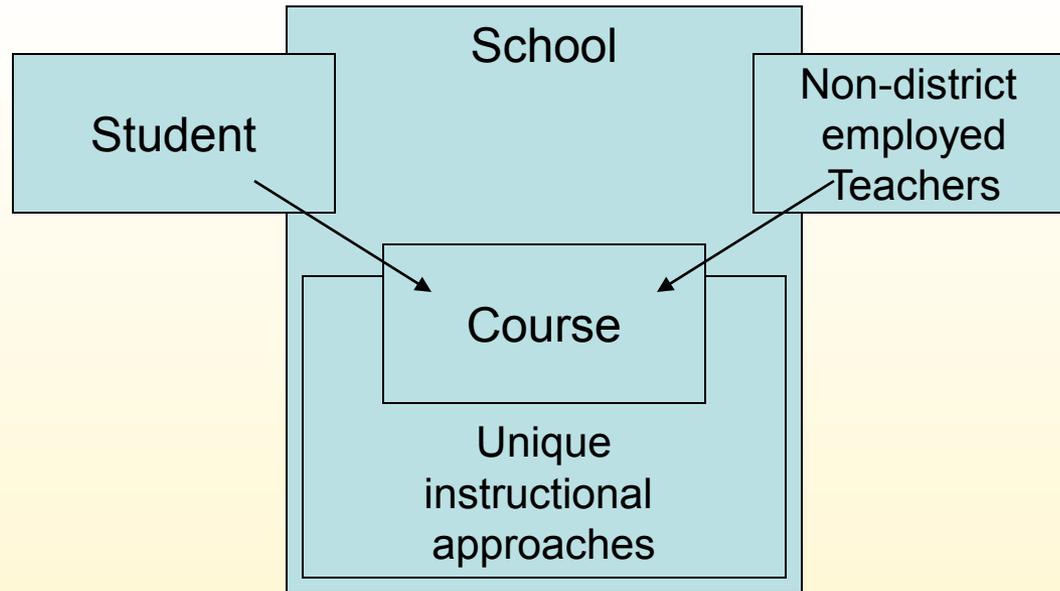
... and absence rates



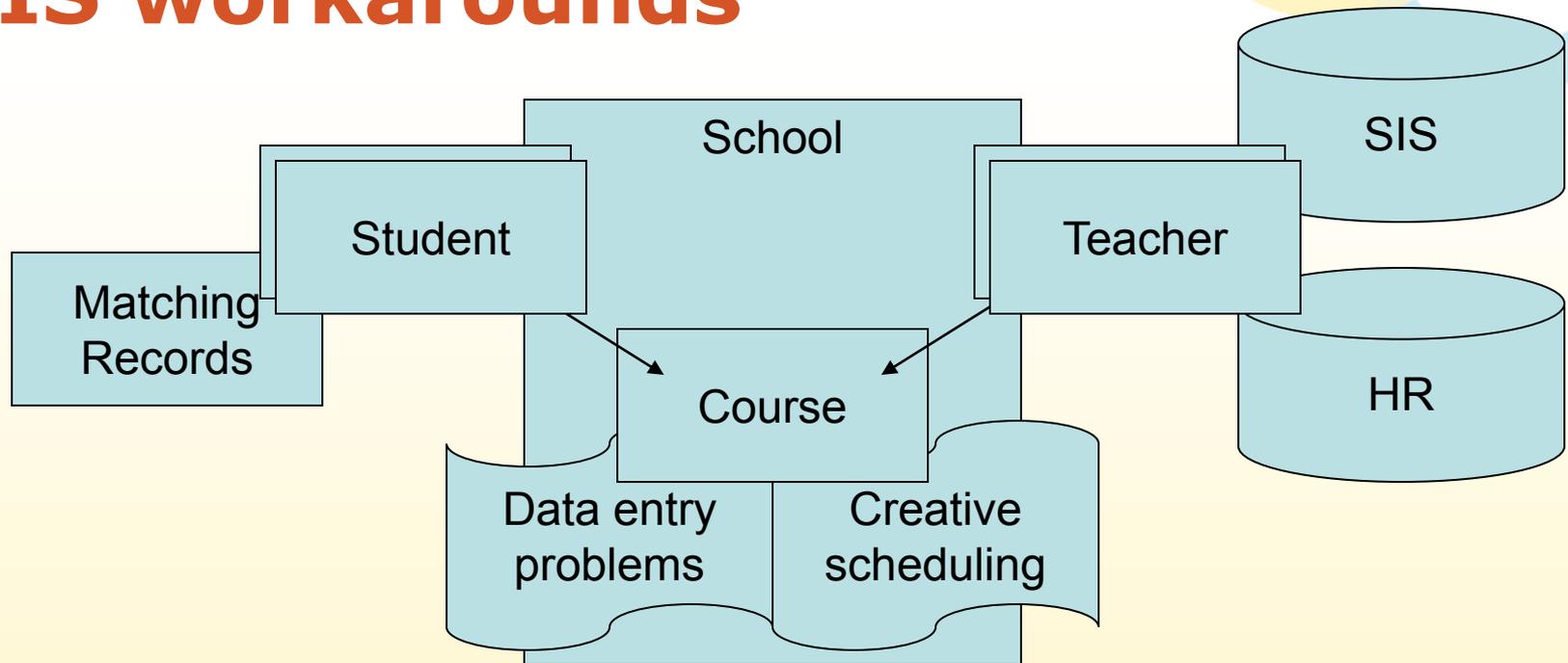
... Instructional strategies like grouping, pull-outs, room aides, team teaching (SAGE)



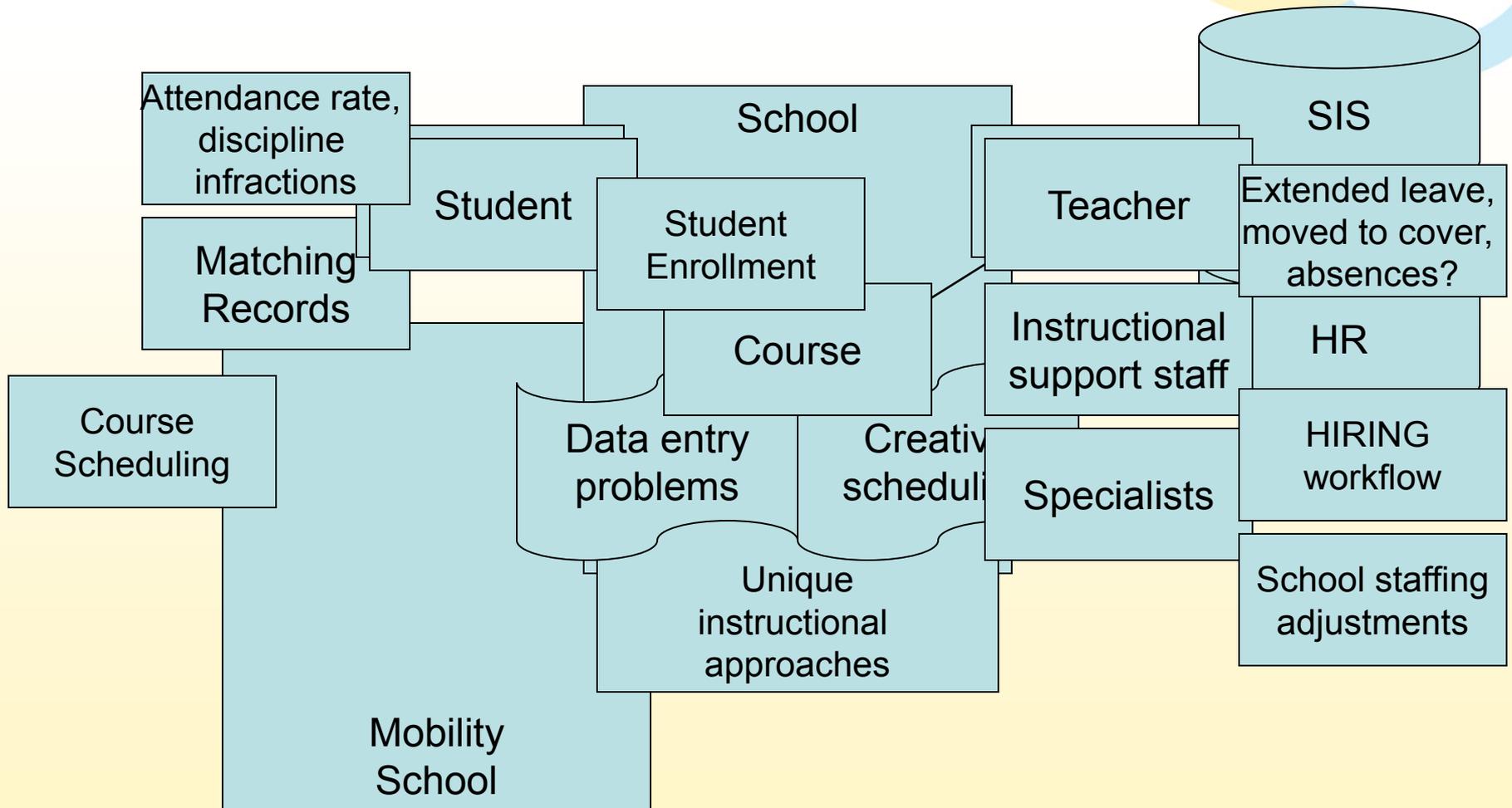
... non-traditional schools



... data errors, integration issues, SIS workarounds



Review of Student-Teacher Linkages



For more on Data Quality



- View the videos on:
 - An Introduction to Quality Data
 - Solutions for Data Quality Problems
 - Teacher Classification
 - A Real Life Example of Data Quality Issues
 - A Real Life Example of Student-Teacher Linkages
- Read the Guide to Implementation module on Data Quality

TIF Challenges in Education Information Systems and Knowledge Management

A Real Life Example of Student- Teacher Linkages

Jeffery Watson



Example 2 (continued)

- Mobility –
 - Introduces multiple teachers
- Do course titles in SIS reflect true curricular content?
- Team teaching –
 - Does SIS data indicate when team teaching is occurring? Who teaches what?
- Pull outs, tutoring, after-school programs (SEEs) –
 - Implications for VAA control variables

T_ID	E_ID	N	SG1	SG2	SG3	Size 1	Size 2	Size 3	Number of Schools	Number of Grades	Small	Large
####	####	34	217 .4			34			1	1	0	1
####	####	42	185 .5			42			1	1	0	1
####	####	2	144 .3	144. 5		1	1		1	2	1	0
####	####	2	174 .3	174. 4		1	1		1	2	1	0
####	####	31	303 .3	337. 3		30	1		2	1	1	1
####	####	4	357 .3	278. 5	357. 5	2	1	1	2	2	1	0

T_ID	E_ID	N	SG1	SG2	SG3	Size 1	Size 2	Size 3	Number of Schools	Number of Grades	Small	Large
####	####	34	217 .4			34			1	1	0	1
####	####	42	185 .5			42			1	1	0	1
####	####	2	144 .3	144. 5		1	1		1	2	1	0
####	####	2	174 .3	174. 4		1	1		1	2	1	0
####	####	31	303 .3	337. 3		30	1		2	1	1	1
####	####	4	357 .3	278. 5	357. 5	2	1	1	2	2	1	0

Distribution of teachers by # of schools and # of grades taught

	# of Schools	# of Grades			
		1	2	3	
June SPS Assess	1	792 158	119 71	29 28	940 257
	2	45 45	11 11	1 1	57 57
	3	1 1	0 0	0 0	1 1
Total by # of Grades		838 204	130 82	30 29	998 315

Solutions

- Audit data accuracy in SIS – use sampling, target initial analyses on grades that are easier to assess student – teacher linkages (assess quality)
- Examine capacity of SIS to track SES, team teaching, etc. (assess validity)
- Build incentives for schools to accurately record teacher of record; verify with teachers (improve validity)
 - Example: MPS requires teachers to build a course roster from a list of enrolled students. Redundant, but serves to validate the accuracy teacher / student links in SIS (improves quality through integration)
- Confirm accuracy of SIS data through phone calls and pen-and-paper questionnaires (quality and validity)

For more on Data Quality



- View the videos on:
 - An Introduction to Quality Data
 - Solutions for Data Quality Problems
 - Teacher Classification
 - A Real Life Example of Data Quality Issues
 - Establishing Student-Teacher Linkages
- Read the Guide to Implementation module on Data Quality

TIF Challenges in Education Information Systems and Knowledge Management

Teacher Classification

Jeffery Watson

Example 3: Classifying Teachers into Categories

- Teachers often teach across grades and content areas
- Context
 - What teachers teach both math and science?
 - What teachers teach more than one grade?
 - What is a course anyways?
- Challenges: Validity, reduction, accuracy

Example 3 (continued)

- Is there such a thing as a “math” teacher?
- Analyses
 - Create case logic for sorting course numbers into content areas
 - Count number of students in each course number, break out by grade level of student
- Example:
 - 20% taught students within a single grade and a single content area
 - 60% taught students across grades
 - 10% taught students within a single grade, but in math and science courses
 - 10% taught across grades and across math and science

Example 3 (continued)



- Solutions
 - Design an evaluation system that is aligned with the complex nature of schools, doesn't force teachers into categories, and captures the nature of teacher's jobs (improve validity)
 - Mine enrollment data rather than HR data (improves accuracy, validity)
 - If teachers must be categorized into a single grade or content area, then a couple of approaches might work
 - Use the number of students
 - Use the number of courses

Summary



- Each TIF project has unique IT needs and priorities
- Data quality is critical for most if not all TIF projects
- Data quality has several key components – these characteristics help us understand what to do first
- Improving data quality will involve both short- and long-term solutions
- Priorities should reflect constraints, priorities, and theory of action

For more on Data Quality



- View the videos on:
 - An Introduction to Quality Data
 - A Real Life Example of Data Quality Issues
 - Solutions for Data Quality Problems
 - Establishing Student-Teacher Linkages
 - A Real Life Example of Student-Teacher Linkages
- Read the Guide to Implementation module on Data Quality