DATA QUALITY 1

MSBO Certification course
Rob Dickinson, MPAAA Executive Director

1

Data Quality 1

■ Session Agenda
  – *Putting DQ in context*
  – *Defining our terms*
  – *Why is quality data so hard to achieve?*
  – *Finding the problems*
  – *Fixing the problems*
  – *Data sharing/privacy – FOIA & FERPA*
■ Questions

2

History of Data

- Data – information recorded in an organized fashion for quick lookup and retrieval
- 150 AD – Ptolemy Star Catalog
- List of 1,100 stars, their constellations, brightness, position, etc.

3

History of Data

4

History of Data

- 1st large database – Untied States Social Security Number system
- Established in November of 1935 – 25 million numbers initially issued.
  - *1st number went to John Sweeney Jr, of New Rochelle, NY*
  - *Over 1,000 post offices (later SSA offices) had to coordinate issuance of numbers, recording of income.*

5

History of Data

- SSN – 293-47-2031
- Issued by Toledo OH office in 1971
- Issued to someone who's last name began with 'D'
- 31st person that year whose last name began with 'D' to get a SSN at the Toledo office in 1971

- Identity theft, anyone?

6

History of Data

- 1st data issue – Hilda Witcher
- 1936 – EH Ferree Company – New York
- New national contract for Woolworths Dept Store
- Clear box covers (new acetate plastic)
- Wanted something to put in the wallets to catch attention
- SS cards had been going out for around 3 months, lots of newspaper attention

7

History of Data

- 1st data issue – Hilda Witcher

8

Putting DQ in Context

- Data Quality is one part of larger model – Data Governance
- Data Governance:
  - *Policies, processes, and practices that control our data and ensure its quality*
  - *Hard to see directly, easier by example:*

9

Putting DQ in Context

- Where most organizations are:
  - *Data is defined inconsistently across systems*
  - *Student data is duplicated*
  - *Staff time wasted massaging data*
  - *Fragmented view of students exists*
  - *Accuracy issues in key data elements*
  - *Inefficient, leads to 11th hour scramble*

10

Putting DQ in Context

- The goal is:
  - *Key data elements sync across systems*
  - *Student information is not duplicated*
  - *Staff spends time analyzing, not verifying*
  - *Systems show a COMPLETE picture of student*
  - *Systems report efficiently for all compliance needs*
  - *Certification deadline is just another day*

11

Putting DQ in Context

- Not just data
  - *How well is staff trained on data definitions?*
  - *Are field 'owners' known to all?*
  - *How are staff informed of inevitable changes in these things?*
  - *Are staff encouraged to analyze data?*
  - *Does EVERY staff know data privacy rules, and live them?*
- All these things add up to Data Governance

12

Defining our terms

- Data Quality
  - *2 primary focuses*
- Quality Assurance
  - *Methods and processes to keep bad data from getting into systems*
- Quality control
  - *Ways to find and correct bad data once it's in our systems*

13

Defining our terms

- Pupil Accounting terms
  - *FTE – Full Time Equivalency*
  - *CEPI – Center for Educational Performance and Information*
  - *MSDS – Michigan Student Data System*
  - *General Collection – 3 time/year snapshot data collection*
  - *SRM – Student Record Maintenance*
  - *EEM – Educational Entity Master*

14

Defining our terms

■ Pupil Accounting terms
  – *UIC – Unique Identification Code*
  – *PIC – Personnel Identification Code*
  – *SIS/SMS/SRS – Student Information/Management/Record System*

15

Defining our terms

■ Pupil Accounting terms
  – *MDE – Michigan Department of Education*
  – *OAS – Office of Accountability Services*
  – *Secure Site – District user site maintained by OAS for districts to manage the testing of students*

16

Defining our terms

■ Data Quality
  – *Data that is fit for its intended use*
■ Not "Perfect data"

17

Why is it so hard?

■ Complex systems becoming interlinked
  – *Student records – Special Ed*
  – *Student records – Food service*
  – *Student records – Bus Routing*
  – *Student records – Personnel / HR*
  – *Student Records – Financial?*
  – *Student Records – Public portals/Websites*

18

Why is it so hard?

- Projects are very goal driven, usually compliance driven & punitive (negative feedback only)
    - Fosters attitude "meet <u>THIS</u> requirement, then move on"
    - *Do you/anyone have time to analyze data, find & fix errors?*
    - *Does your work environment/culture value or invest in data quality?*

19

Why is it so hard?

- Software Development methods don't emphasize data quality
    - *Systems are evaluated on functions/features*
    - *Look of screen, how it functions*
    - *Information integrity is not valued as a decision-making criteria*
    - *If customers don't ask for data integrity, vendors won't build it.*

20

Why is it so hard?

- Data Quality exercise:
  - *Form being handed out*
  - *Create data input rules for date of birth*
  - *Ages where warning or errors SHOULD occur*
  - *Building or District level?*
  - *Work with your table*

21

Why is it so hard?

- Data problems are hard to find
  - *Data quality evaluated by different systems than capture it*
  - *Delays in time*
  - *Personnel, source data no longer available*
  - *Once problem propagates, much harder to root out.*

22

Why is it so hard?

- Input fitness
  - *Data is usually only made clean to the level needed by the person inputting it.*
  - *Example - building staff, working with parent*
  - *No incentive to maintain high data quality*
  - *Errors only show when data is summarized and integrated – which is usually at time of reporting*

23

Why is it so hard?

- Do you know all your data's "Intended use"?
  - *Data exists in our systems a LONG time*
  - *Impossible to know ALL intended uses at time of entry*
  - *Collection systems can't anticipate every future need*
  - *Reactive legislature adds to the problem*

24

Who Cares!?!

■ Why is quality data important?

25

Who Cares!?!

The price of quality data is far lower than the cost of the alternative

26

Who Cares!?!

- Costs of bad data
  - *Financial*
  - *Embarrassment*
  - *Reputation*

27

Finding the problem

- Who are the flag wavers?
  - *Who will know something is wrong?*
  - *How can they fix it, or raise flag?*
- How can YOU find errors?
  - *What data can you analyze?*
  - *What's be best time for analysis?*

28

Finding the problem

- Flag waivers
  - *Parents*
    - Parent portal
      - *Opportunity to send message on bad data*
        - Double edged sword

29

Finding the problem

- Flag waivers
  - *ETL processes & partners*
    - Data errors that occur when files transfer to other software
      - *Nightly with special ed system*
      - *Occasionally with transportation/other systems*

30

If you want it done right...

- Do it yourself!
  - *Compare summary data*
    - Do totals by various categories seem close?
      - *Need good feel for your data*

31

Finding the problem

- It can be challenging
  - *Student ethnicity – African American*
  - *CEPI data "010000" or "000100"?*
- You shouldn't rely on memory

32

Finding the problem

- Data analysis
  - *Can you run queries?*
  - *Download data from your SIS?*
  - *Can reports be dumped to Excel?*
- Beware of privacy issues!

33

Deciding to fix the problem

- Can the data be fixed?
- Should the data be fixed?

34

Can the data be fixed?

- Is the correction a new value for this field?
  - *No – Easy fix*
  - *Yes – Not so easy*

35

Can the data be fixed?

- New values for a field
- Mechanical issues
  - *Will new value fit nature of field?*
  - *Will it fit in size of field?*
- Policy/process Issue
  - *What else will new value affect?*

36

Can the data be fixed?

∎ New data values – Effects:
- – *All existing reports, queries*
- – *Any existing error checking*
- – *Other users of this field*
- – *Other systems that link at this data*

37

Should the data be fixed?

∎ Reasons NOT to fix data
- – *Fix has no net benefit*
- – *Cost greater than benefit*
- – *Lack of resources*
- – *Internal need greater than external*

38

Fixing the problem

- At what level should the fix be executed?
- How should the error be fixed?
- Who should fix the error?

39

What level?

- Sometimes, you don't fix the data
  - *Internal needs override external*
  - *Different uses at different levels*
- Options:
  - *Fix the PROCESS, not the data*
  - *Create a reporting field*

40

Create Reporting field

- Internal field used to break out, or consolidate data for reporting
- Place to build in checks
- Usually NOT available to users
- Change data submission routine to use this field

41

Fixing with Query

- ALWAYS use 2 step process
1. Run list of errors that shows bad data to be changed as it exists before the change
2. Use EXACT same logic to fix error

42

Fixing with Query

- Keep list of errors, showing how data WAS before the correction
- ALWAYS list before changing
- The more you can do, the more you can damage!

43

Fixing with direct input

- One-off error, single fix
- Run through user interface whenever possible
- Allows any existing error traps to run

44

Who should fix the data?

- Users vs Finders
  - *Users are the inputters of the data in normal day to day usage*
  - *Finders are the data staff who collect/report/analyze the data*

45

Who should fix the data?

- User fix
  - *Closest to that particular piece of data*
  - *'Owner' for Data Governance*
  - *Typically small part of their job*
  - *Hard to get their attention*

46

Who should fix the data?

- Finder fix – District level staff
  - *Has greatest motivation to make data right*
  - *Able to concentrate on correction*
  - *Disconnecting users from their data*

47

Fixing the data:
A special case

Fixing errors in CEPI submissions

- A special case
  - *Two approaches*
    - Reload
    - Online correction

48

24

Errors in CEPI Submissions

Reload data

- Fix the error in your system, recreate the upload file, re-upload
  - *Eliminates synchronization errors*
  - *Longest process*

49

Errors in CEPI Submissions

Online correction

- Fix the error in CEPI online system
  - *Fast fix*
  - *Now your source data is different than your reported data*

50

Errors in CEPI Submissions

What is your transition point?

- First load w/many errors, reload
- At some point in process, switch to online
- Keep synchronization in mind

51

Make your life easier

What can you do to make data quality easier?

- Input forms match input screens
- Controls are LIMITED to avoid bad data
- Keep data originators there while you input

52

Make your life easier

District of Residence

- Do you have an ACCURATE way to lookup District for any given address in your area?

- County accessors office

- Call or online, avoiding commercial sites

53

Sharing Records

FOIA and FERPA

- Freedom of Information Act (FOIA)
  - *All District records not exempted may be requested*

- Federal Educational Rights and Privacy Act (FERPA)
  - *Allows student record to be seen by appropriate individuals*

54

Sharing Records

Release of records under FOIA

- Anyone can request records from a public body
- Ask to inspect, copy, or receive a copy of the record
- No age restrictions (cannot refuse the request of a minor BECAUSE they are a minor)
- 5 days to respond with the record, or request for more information, or a 10 day extension to prepare for complex requests
- Not required to CREATE documents, merely share existing ones

55

Sharing Records

Review of records under FERPA

- Parents have the right to view their children's educational records
- Students assume that right when they turn 18
- Districts have obligation not to share records with outside parties without parental consent
- Districts must provide parents notice of their rights under FERPA

56

Protecting Records
FOIA and FERPA

- **Freedom of Information Act (FOIA)**
  - *Exempt records are not required to be shared*
  - *Reasonable costs can be charged*
- **Federal Educational Rights and Privacy Act (FERPA)**
  - *Any record containing Personally Identifiable Information (PII) is protected*

57

Protecting Records
FOIA Exceptions

- Student records covered by FERPA
  - *FERPA trumps FOIA*
- Law enforcement records
- Records covered by attorney-client privilege
- Pending bids on public contracts
- Test questions and answers or scoring keys
- records relating to civil actions
- Requests by prisoners in County, state or federal correctional facilities

58

Getting Help

FOIA Assistance

– *Michigan Attorney General Website*
– *http://www.michigan.gov/ag/0,4534, 7-164-20988_18160---,00.html*

OR

– *Go to www.michigan.gov/ag*
– *Choose 'Opinions' from left hand menu*
– *Select 'FOIA / OMA'*

59

Getting Help

FOIA Assistance

– *Frequently Asked Questions*

– *Informational pamphlets to distribute*

– *All current legal citations*

– *Open Meetings handbook as well*

60

FERPA

Federal Educational Rights and Privacy Act

- Parents have the right to view their children's educational records

- Students assume that right when they turn 18

- Districts have obligation not to share records with outside parties without parental consent

- Districts must provide parents notice of their rights under FERPA

61

FERPA

What are educational records?

- Every record that the district holds that contains Personally Identifiable Information (PII)
    - *Exception – personal notes for personal use only*

- PII – Any information that discloses an individual's identity
    - *Even summary data that can reasonably point to an individual*

62

FERPA

Parents:

- FERPA defines a parent as a natural parent, a guardian, or an individual acting in the place of a parent or guardian. In the case of divorce or separation of a student's parents, both parents maintain full rights under FERPA unless some legally binding document or law revokes those rights.

- It is NOT the school's responsibility to hunt down or find those documents. Only to act on them when they are presented.

63

FERPA

Accessing and viewing records

- District has 45 days after a written request to make records available

- District must comply with reasonable requests for explanation of documents

- Copies do not have to be made, so long as parent can view documents

64

FERPA

Amending records

- Parent has the right to request that incorrect data be corrected in the record.
- School is obligated to CONSIDER the request.
- If the school decides to not change the data, they must inform the parent of their right to a hearing.
- If a hearing is held, and the parent the decision is made to not change the data, the parent has the right to puta statement into the student's file explaining their position on the contested information.
- School has to maintain that statement and produce it whenever it discloses the contested data on the student.

65

FERPA

Amending records

- Under FERPA, school are not required to consider requests to change:
  - *A grade for a student because they feel the student should have been given a better grade*
  - *Opinions*
  - *Disability placement decisions regarding the student*
  - *Discipline decisions regarding the student*

66

FERPA

Student rights

- When a student turns 18, their FERPA rights transfer from the parents to the student
  - *Also applies when a student enters college at any age*
- Parents may no longer view students records, unless:
  - *Student is listed as a dependent on parent's current tax records*
  - *Pursuant to a health emergency*
  - *Student consents*

67

FERPA

Sharing protected records

- District may not disclose information without parental consent, unless:
  - *Directory Information*
  - *Exempt recipient*

68

FERPA

Directory information

- Information that is not harmful, or intrusive to privacy
  - *Name*
  - *Address & other contact information*
  - *Date and place of birth*
  - *May NEVER include Social security number*
- District may determine what it constitutes directory information

69

FERPA

Directory information

- Student numbers
  - *Can be considered directory information*
  - *Can be displayed on student ID*
- UNLESS
  - *Number is used as a password, without other PIN or counter code, to gain access to student data*
  - *In that case, it cannot be on the badge, and is not eligible to be considered directory information*

70

FERPA

Directory information

- May be disclosed without consent if:
  - *Annually, notice is given to all parents of what records district deems to be directory*
  - *Gives parents the option to opt out of disclosure*
- If parent opts out, Directory info for that student MUST NOT BE PROVIDED.
  - *Most SIS have FERPA opt-out check boxes*
  - *Important that custom reports reference that status*

71

FERPA

Exempt recipients (within the district)

- Records may be shared with:
  - *Employees of the district with a 'legitimate educational interest'*
  - *Teachers, administrators*
  - *Attorneys, counselors, nurses*
  - *IT staff*
  - *Contractors, consultants, 3rd parties providing services*

72

FERPA

Exempt recipients (outside the district)

■ Records may be shared with:
- *Public health and law enforcement*
- *Emergency responders*
  - ■ Only for imminent or already existing emergencies. Not as preparation for future hypothetical events

73

FERPA

Exempt recipients (outside the district)

■ Records may be shared with:
- *School a student transfers to*
- *Accrediting organizations*
- *Local, state or federal officials with audit or program evaluation needs*
- *Organizations the student has applied to for admission for school, or for financial aid*
- *To comply with a subpoena or judicial order*

74

FERPA

Annual Notice

- Must include:
  - *Parent's rights to inspect records*
  - *Right to request corrections*
  - *Right of consent to disclosure of PII*
  - *Procedure for inspecting records*
  - *Procedure to requesting corrections*
  - *Information of reporting complaints to USED Family Policy Compliance Office*
- Frequently included in student handbook

75

FERPA

Training assistance

- Privacy Technical Assistance Center (PTAC)
  - *Part of US Department of Education*
- https://studentprivacy.ed.gov/
- Rules, current case law
- Training videos
- Will track learning, provide documentation of compliance
- Excellent source for training staff

76

Getting Help

- Pupil Accountant
  - *Closest, knows your data best*
- ISD Auditor
  - *Knows the rules and regulations*
  - *Does your Pupil Accountant want to be a liaison?*

77

Getting Help

- CEPI Helpdesk
  - *(517) 335-0505, Option 3*
  - *cepi@michigan.gov*
- MPAAA
  - *Rob@mpaaa.org*
  - *(517) 853-1413*

78