



DATA QUALITY 2

MSBO Certification course

Rob Dickinson, MPAAA Executive Director



Data Quality 2

- Session Agenda
 - *Where we Stand*
 - *Defining Data Quality & Data Governance*
 - *Quality Assurance*
 - Data Definitions & Types
 - QA methods
 - *Data Governance in practice*
 - *Data Privacy and FERPA*
- Questions

Where we stand

Data Quality Campaign (DQC)

- 10 Essential Elements for a Statewide Longitudinal Data System
 1. *Unique Student Identifier*
 2. *Student-level enrollment, demographic, program participation information*
 3. *Ability to match student level testing year-to-year*
 4. *Information on Untested students*
 5. *Teacher identification system*

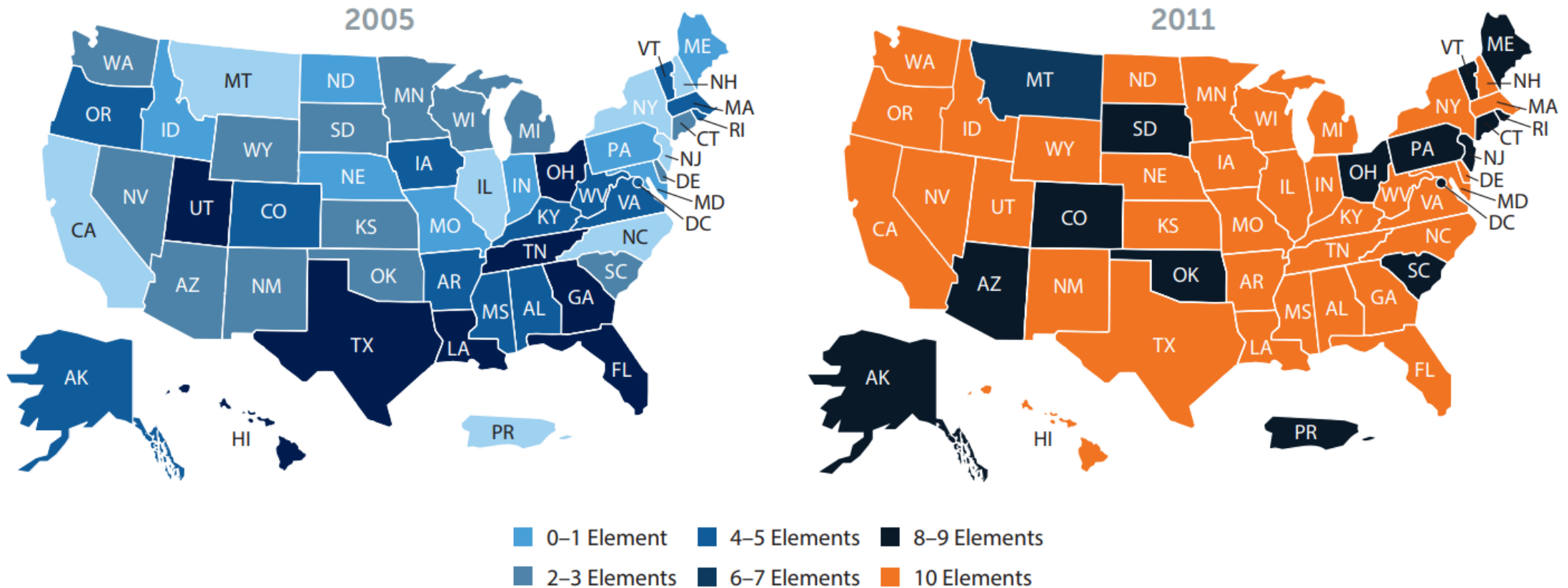
Where we stand

Data Quality Campaign (DQC)

- 10 Essential Elements for a Statewide Longitudinal Data System
 6. *Student-level transcript data*
 7. *Student-level college readiness test scores*
 8. *Student-level Graduation and Dropout data*
 9. *Match students from K-12 to postsecondary systems*
 10. *Statewide data audit system*

Where we stand

10 Essential Elements of Statewide Longitudinal Data Systems



Where we stand

Data Quality Campaign (DQC)

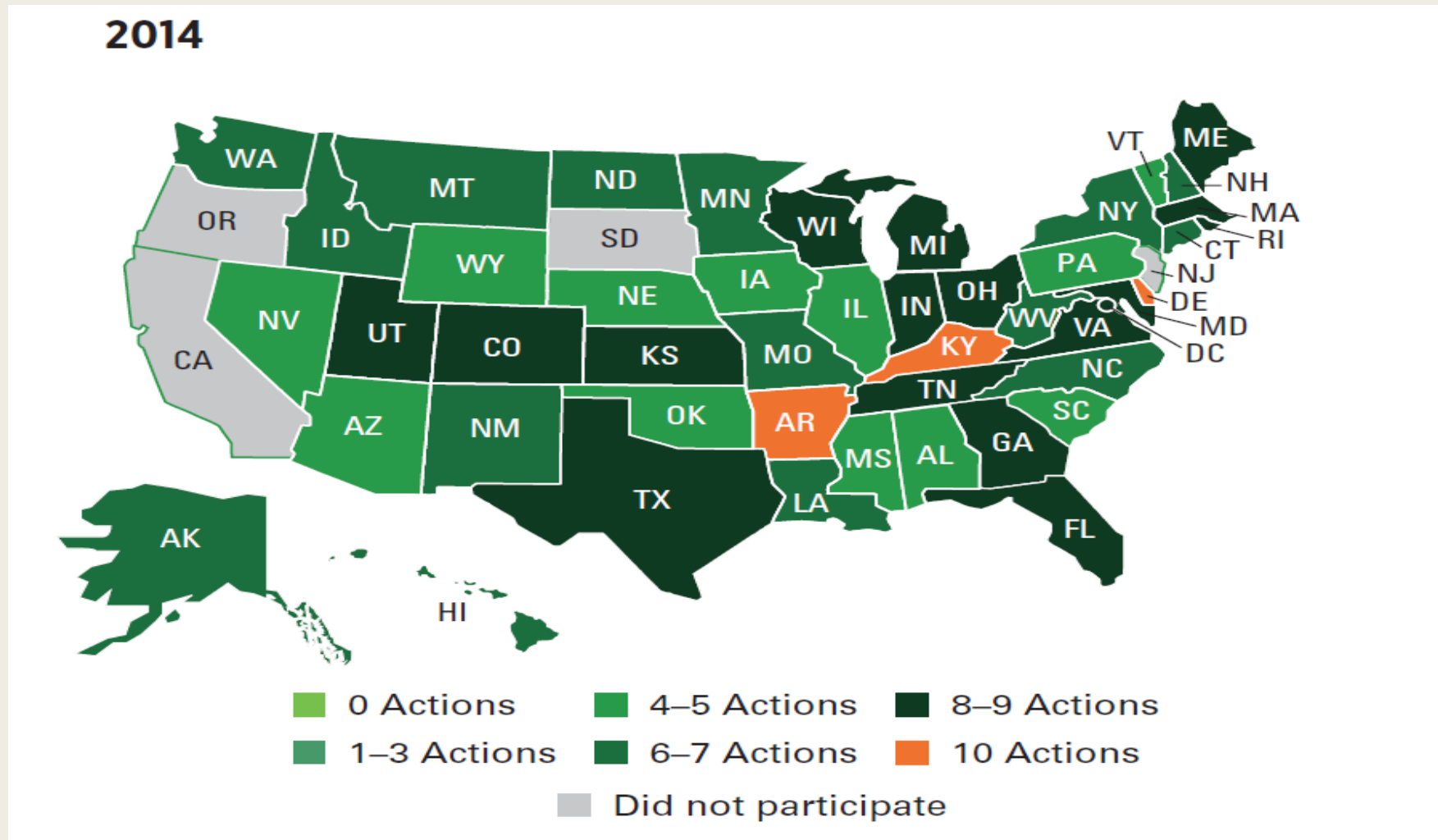
- 10 State Actions to Ensure Effective Data Use
 1. Link K-12 data to Early learning, postsecondary, and workforce data
 2. Create stable support for SLDS system
 3. Create governance structures
 4. Build state data repositories
 5. Data collection for more than compliance purposes

Where we stand

Data Quality Campaign (DQC)

- 10 State Actions to Ensure Effective Data Use
- 6. Create reports for educators, students, parents
- 7. Have reports provide longitudinal statistics
- 8. Develop a purposeful research agenda
- 9. Implement policies to build educators capacity to use data
- 10. Raise awareness of available data

Where we stand



Where we stand

Data Quality Campaign (DQC)

- *By 2014, Michigan met all elements and all Actions*
- *By 2017, all states with SLDS systems met all Elements*
 - *Several still working on Actions*
- *Michigan continues to lead in national data quality discussions*
- *CEPI staff involved in DQ policy and practice discussions at the national level*

Defining Data Quality

*Wikipedia - Data quality refers to the condition of a broad set of qualitative and quantitative variables. There are many **definitions of data quality** but **data** is generally considered high quality if it is "fit for [its] intended uses in operations, decision making and planning".*

- Not 'perfect', or 'error free'
- Involves both tangible (Quantitative) and intangible (Qualitative) measures

Quantitative measures

- Accuracy
- Integrity across systems
- Consistency
- Completeness
- Uniqueness
- Accessibility
- Precision
- Timeliness

Qualitative measures

- Relevance
- Usability
- Usefulness
- Believability
- Unambiguous
- Objectivity

Defining Data Quality

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

Data Governance

- 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*

Data Governance

- 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*

Data Governance

- 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

Putting DQ in Context

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

Quality Assurance

- Controlling data as it enters your systems
- Important part of system design/installation & maintenance
- 3 Major areas
 - *Data field design*
 - *Input control functions*
 - *System modification/customization*

Data Field Design

- Selecting the most appropriate type of field for the data it will hold, and assigning properties to that field to limit bad inputting.
- Field Types: Boolean, number, text, date
- Coded fields: Intrinsic, non-intrinsic
- Field Formats: Check boxes, buttons, selection lists, input fields

Field Types

Boolean

- ONLY 2 values - Yes/No, True/False
- Status (Participant status, Enrolled, Was Absent on Count day)
- Can NEVER hold a 3rd option
- Usually cannot be left blank
- Won't allow for any future re-definition

Field Types

Number

- Used for values, amounts
- Sometimes used for codes
- Significant digits are important
- Subtypes
 - *Integer* – 1, 2, 3 (no decimal)
 - *Currency* – Always 2 digits of decimal
 - *Floating Point* – No functional limits

Field Types

Text

- Used for list of values, string input
- WEAK choice for number only input
- Direct input – Almost impossible to analyze
- Using text for numbers
 - *Allows leading '0', fixed width*
 - *Only for list of codes*

Field Types

Dates

- Used for inputting dates, sometimes times
- Sometimes stored as number
- Usually built-in error checking for valid dates
- Allows date math
- Formatting for century (3/1/2016 vs 3/1/16)

Code Fields

- Stores limited list of values
- List determines field type (number, text, etc)
- Good error checking
- Adding & deleting values is a problem
- When creating – Intrinsic vs non-intrinsic
 - *Intrinsic* – the stored data conveys information
 - *Non-intrinsic* – stored value has no meaning on its own

Code Fields

Intrinsic or Non-intrinsic?

UIC

SSN

MSDS Exit codes '19'

MSDS Ethnicity codes '010000'

EEM District codes '41010'

EEM Building Codes '03921'

Code Fields

Intrinsic codes

- SSN, Gender, Special ed program codes
- Good
 - *Easy to understand*
 - *Built in error checking*
- Bad
 - *No privacy – allows guessing*
 - *Needs strong rules*
 - *Limits possible values*
 - *Needs to know all possible values*

Code Fields

Non-intrinsic codes

- UIC, EEM Building codes, MSDS Exit codes
- Good
 - *Not limited by rules*
 - *Can accommodate growth/change*
- Bad
 - *Has no value in itself, needs value chart/list*
 - *Can run into limits (field width)*
 - *Can only work if there is only 1 place generating values*

Field Formats

The interface that controls how the data is entered

- Checkboxes, radio buttons
 - *Boolean data, 1 choice among very few*
- Lists, Dropdown lists
 - *List choices available, one or more than 1*
- Input box
 - *Most freeform, hardest to control input*

Field Formats

The screenshot shows a Windows form titled "Form1" with the following controls:

- Status
- Race
 - Hispanic
 - African American
 - Caucasian
 - Native American
- Last name:
- Grade:
 - 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 11
 - 12
- Breeds:
 - Afgan
 - Airdale
 - Boxer
 - Bulldog
 - Bassetthour
 - Collie
 - Greyhound
 - Bloodhound
 - Mastiff
 - Bullmastiff
 - Shepard
 - St Bernard
 - Pointer

QA Methods

Ways to ensure data is entered into your systems correctly

- Error checking at input
- Training for input staff
- Error checking routines run at regular intervals
- New screens, reports, queries follow same rules

Error checking at Input

Prevent bad data from getting into the system

- Data Types, field formats
- Error checking rules behind the field
- Make it difficult to allow non-standard data to be input
 - *Can't make it so hard that it is ignored*
 - *'Are You Sure?'*

Training for Input Staff

Make sure staff entering data is aware of it's importance

- Initial training

- *Bring new staff up to speed*
- *Familiar with systems*

- Recurring training

- *Letting everyone know what's new, changed*
- *Reminders on problem areas*

Error checking routines

Frequently run reports/queries designed to find errors soon after input

- Find and fix before it is used, propagated to other systems
- Nightly, over weekend, end of attendance period
- Can be system report, email, faxed, etc.
- Do you fix, or do they?
- Balance of finding errors vs overwhelming users

Error checking routines

New screens or reports or queries **MUST** follow same QA rules to prevent 'trapdoor' errors

- New screens – 'All in one' or audit screens
- Reports – Do report generators follow security rules?
- Queries – double edged sword
 - *Easy to run, change data quickly*
 - *Usually avoids all the user input rules*
 - *Can fix - or break - large amounts of data very fast*

Data Governance

- Data Horror stories
 - *Bear Sterns, 2002*
 - *Japan Stock Market, 2005*
 - *SID data – West Michigan, 2011*

- Impact of poor data governance

Data Governance

Data Governance Strategy

- Overall vision for improvement
- Program Implementation plan
- Linking data Quality back to District policies and objectives
 - *How does good data make education easier?*

Data Governance

Technology & Architecture

- Flexibility to change
- Open and Common Standards
- Data accessibility among systems
- End-to-end data security

Data Governance

Governance Organization

- D.G. recognized at a organizational level
- Data quality as an embedded competency for ALL staff
- Data Stewards recognized and known
- Senior Stakeholders recognized and known

Data Governance

D.G. Processes

- Correction processes
- Root cause analysis
- Best practices and methods
- Focus on Improvement
 - *Starting on Key elements*
- Supply chain approach

Data Governance

D.G. Policies

- Common definitions
- Data Standards
- Review of Policies and Standards
- Defined Controls

Data Governance

Data Monitoring/Investigation

- Qualitative understanding of issues
- Key data pieces identified
- Ongoing monitoring
- Tracking of issues for Improvement

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

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*

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
- Parents DO have the right to request that inaccuracies be corrected
 - *Is not a mechanism to dispute grades, alter opinions, eliminate discipline records*

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 not view students records, unless:
 - *Student is listed as a dependent on parent's current tax records*
 - *Pursuant to a health emergency*
 - *Student consents*

FERPA

Sharing protected records

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

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

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 new/custom reports reference that status*

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*

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

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*

FERPA

Correcting 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 put a 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.

FERPA

Correcting records

- Under FERPA, schools 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*

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*

FERPA

Training assistance

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

Getting Help

- CEPI Helpdesk

- (517) 335-0505, Option 3

- cepi@michigan.gov

- MPAAA

- Rob@mpaaa.org

- (517) 853-1413