USOE Data Warehouse Information Site

Important Data Concepts

- The focus of the data collection process for U-PASS/NCLB is on timeliness, accuracy and efficiency. New and changed data are only introduced to meet new legal requirements or improve data accuracy or efficiency.

- Timely reporting of complete and clean data to the USOE clearinghouse and data warehouse is the key element for successful data management and analysis for the districts and the USOE. The clearinghouse and warehouse encompass more than just U-PASS and NCLB. They help streamline data collection and increase the usefulness of all public education data in Utah.

- The success of both the clearinghouse and warehouse depends on understanding and cooperation between the USOE and the districts and schools.

- The USOE would like to designate a single data collection coordinator in each LEA as the USOE's primary data contact for all data collection matters and communications. This would be a "central" individual with whom the USOE can communicate concerning any data collection issues. The USOE has prepared a proposed list of these individuals.

The Seven Elements of Quality Data and Some Truths

1. **Accuracy:** Does the data accurately represent reality or a verifiable source?
2. **Integrity:** Is the structure of data and relationships among entities and attributes maintained regularly?
3. **Consistency:** Are data elements clearly defined and understood?
4. **Completeness:** Are all necessary data present?
5. **Validity:** Do data values fall within acceptable ranges defined by the business?
6. **Timeliness:** Are data available when needed? Don't let this ever override the ones above.
7. **Accessibility:** Are the data easily accessible, understandable and useable for generation of information?

- **To have quality data there must be a data coordinator on the business side of the organization to lead the data quality and sometimes cleansing effort.**

- **Don't think you can clean data just once. It is an ongoing effort that requires significant investments of time, money and resources.**

- **Data quality is a business problem and needs to be owned by the people who rely on the data; it's much more than good technology.**

- **The Gartner Group estimates that 25% or data large companies use for business analytics is flawed due to human error and poor processes.**

Tech types, including myself, often scoff, ignore or overlook these simple intiutive truths, but one does so at great risk