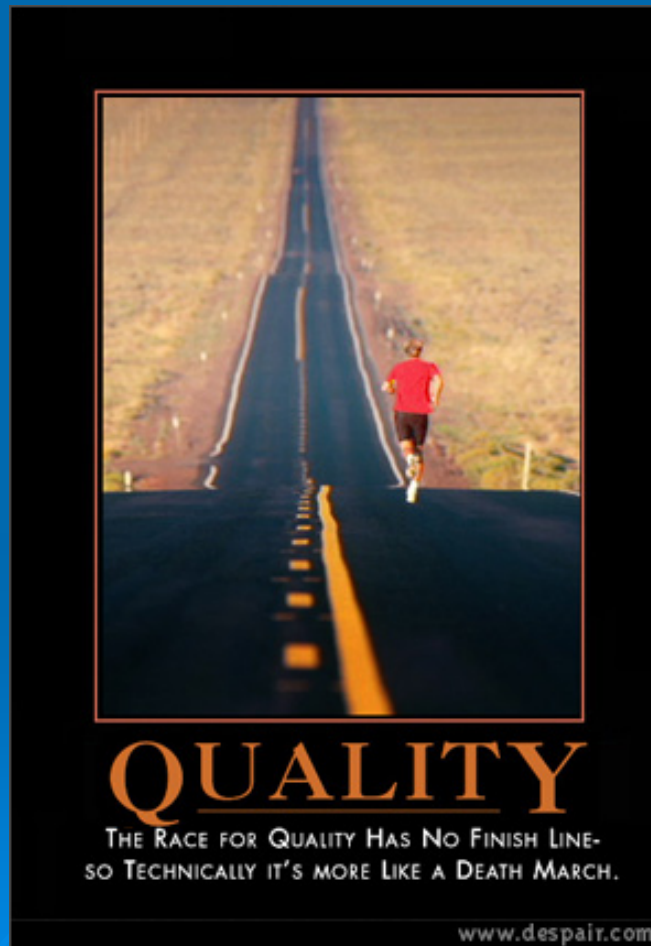# Workshop to define quality management standards for data completeness in derived data products

Las Cruces, NM 31-Jan-2007 to 01-Feb-2007

Motivations

- ➤ LTER IM mail list thread about Trends submissions
  - How does Climdb deal with missing daily data in monthly/yearly aggregated products?

- ➤ ASM discussions
  - How are incomplete & flagged data being handled for Trends?
  - How should info about data quality be managed and communicated in PASTA architecture (or other federated databases)
  - Similar issues being discussed in CUAHSI, SEEK, Grasslands NPP DB

- ➤ Many maturing and emerging synthesis efforts – good time to standardize on approach to derived data quality
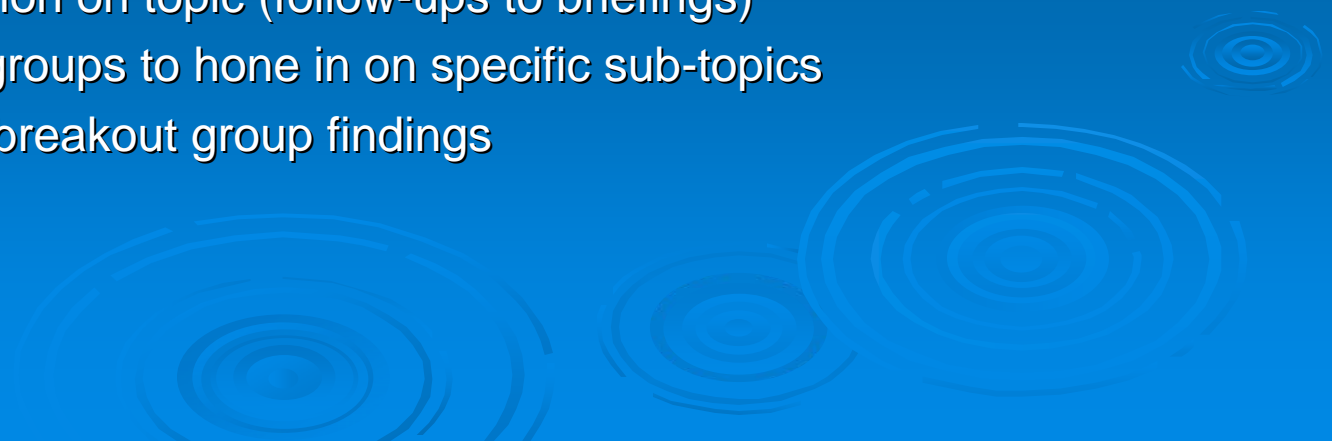
Workshop Goals

➢ Bring together people with broad experience to discuss quality management for derived data

➢ Try to identify core set of best practices for:
- Data & metadata providers
- Synthetic databases and data management systems
- Synthesis tools and other CI

➢ Develop a plan to draft standards (or adapt existing ones), gather community input, and publish findings

➢ Also tie in with EML development, EML BP revisions

Agenda

- ➤ **Four Sessions**
  - Current practices regarding missing, qualified values in primary data
  - Performing synthesis with incomplete, qualified primary data
  - Providing quality assurance for derived data
  - Drafting standards and best practices (writing/outlining findings)

- ➤ **Format for Sessions 1-3**
  - Short briefings and Q/A
  - Group discussion on topic (follow-ups to briefings)
  - 2-3 Breakout groups to hone in on specific sub-topics
  - Discussion of breakout group findings

Things to Avoid