DRI Data Legacy

A Data Managers Perspective

Phillip Harder

Purpose for DRI Data Legacy Management Plan

- Provide a framework to develop the DRI Data Legacy

- Increase the quality of the DRI Data Legacy

- Facilitate the sharing of data within and outside of DRI

- Provide a setting to showcase DRI's research

Plan Outline

- Phase 1: Metadata (6 months)

- Phase 2: Data (end of 2010)

Phase 1 - Metadata

- DRI needs to adopt a metadata standard:
  - convergence towards ISO 19115 in scientific community
- Standard would allow DRI data to be included in:
  - GEO, World Change Master Directory etc... and other data portals that use ISO 19115 as metadata standard

Phase 1 – Metadata

- Need a Metadata organization system
  - a simple list on the website, "one stop shop"
  - a searchable relational database
    - GeoNetwork Spatial Data Catalogue
      - http://geonetwork-opensource.org/
- Key to success:
  - System needs to be simple and accessible to DRI
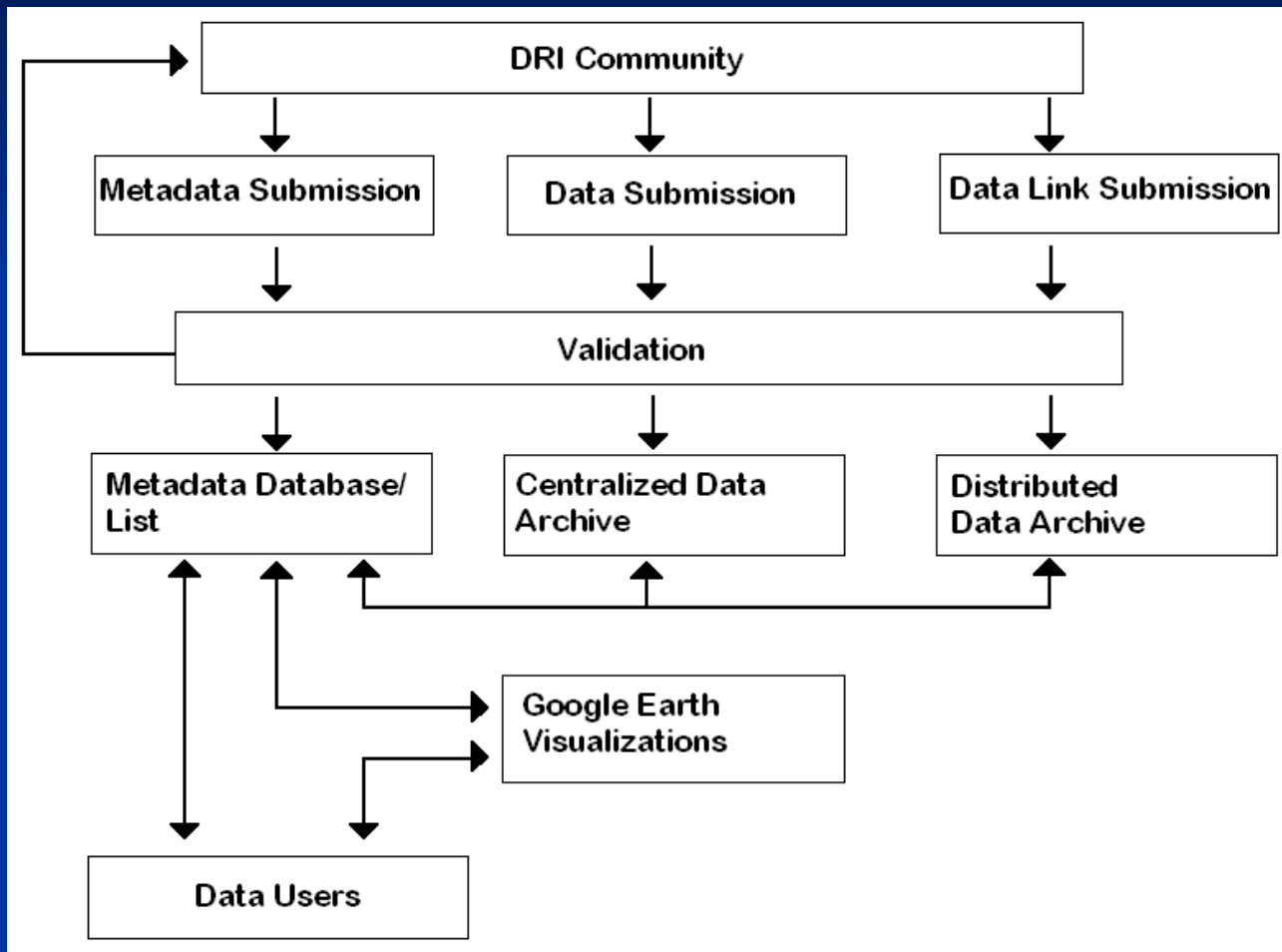 - **DRI COMMUNITY PARTICIPATION**

Phase 2 - Data

- Nature of DRI datasets is unique:
  - Historic focus (1999-2004/2005) means data often collected by other agencies
- Need to define what a DRI dataset is:
  - What datasets have unique value and need to be preserved?
  - At what level does the data need to preserved?
    - Raw, QA/QC, Gridded etc….
- How should access to data be arranged that makes the most sense for DRI?
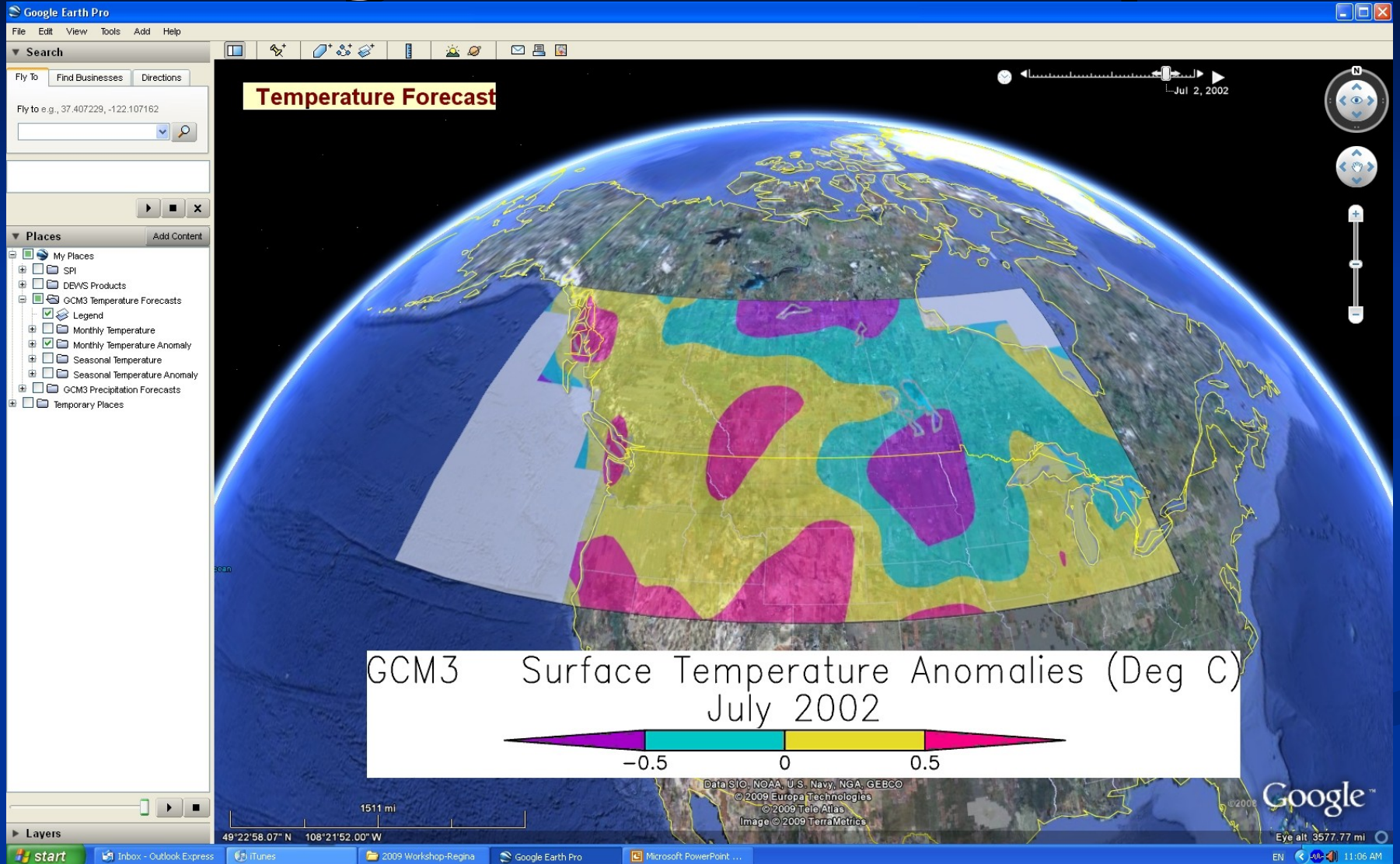 - Database, ftp, www, archive etc….

Phase 2 - Data

- Possible Data Organization Route:
  - Online Metadata Database/List will serve as data portal
  - Available data - centralized archive
  - Unavailable data - distributed archive (links)
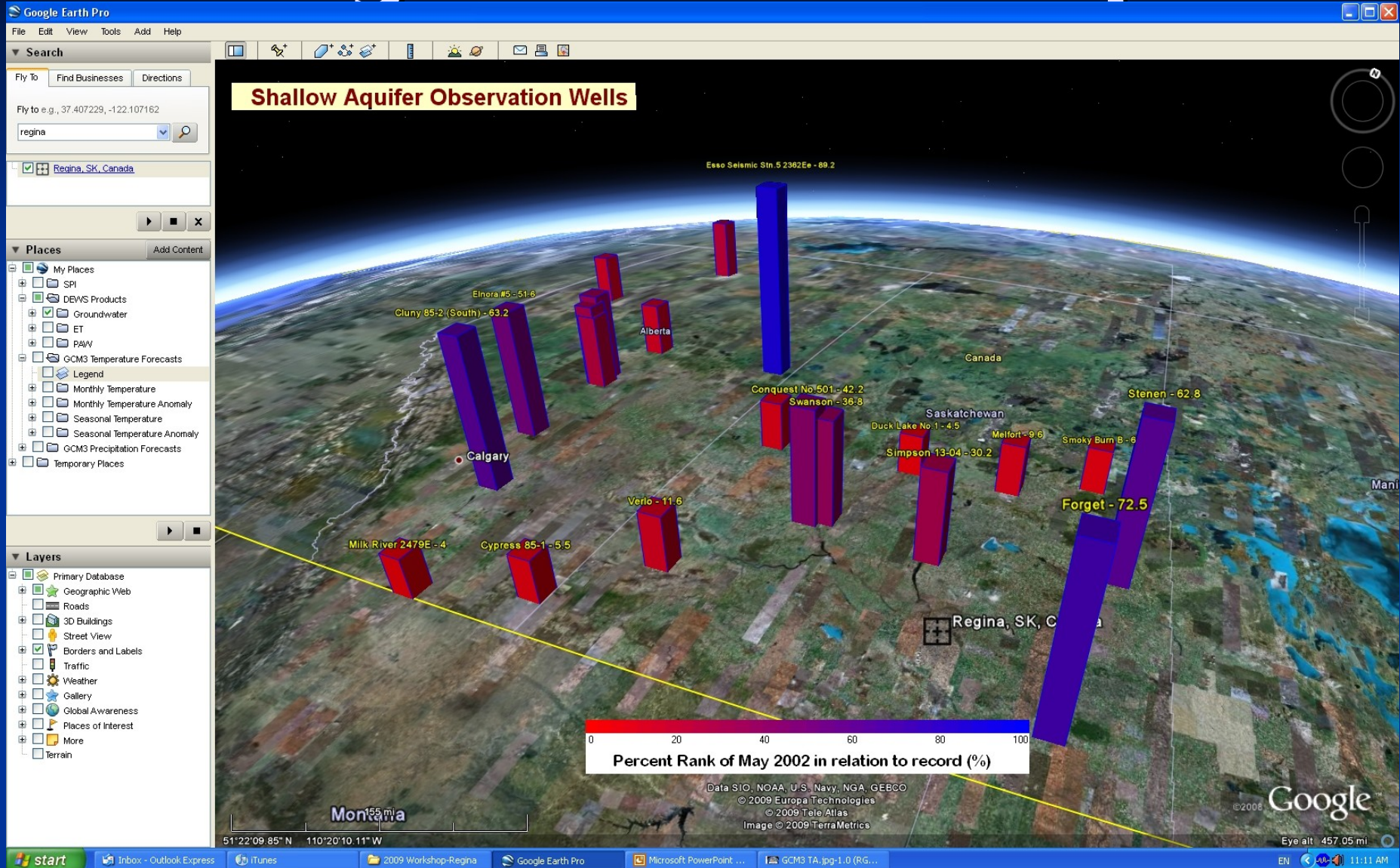- Google Earth can be used as an interactive data visualization tool
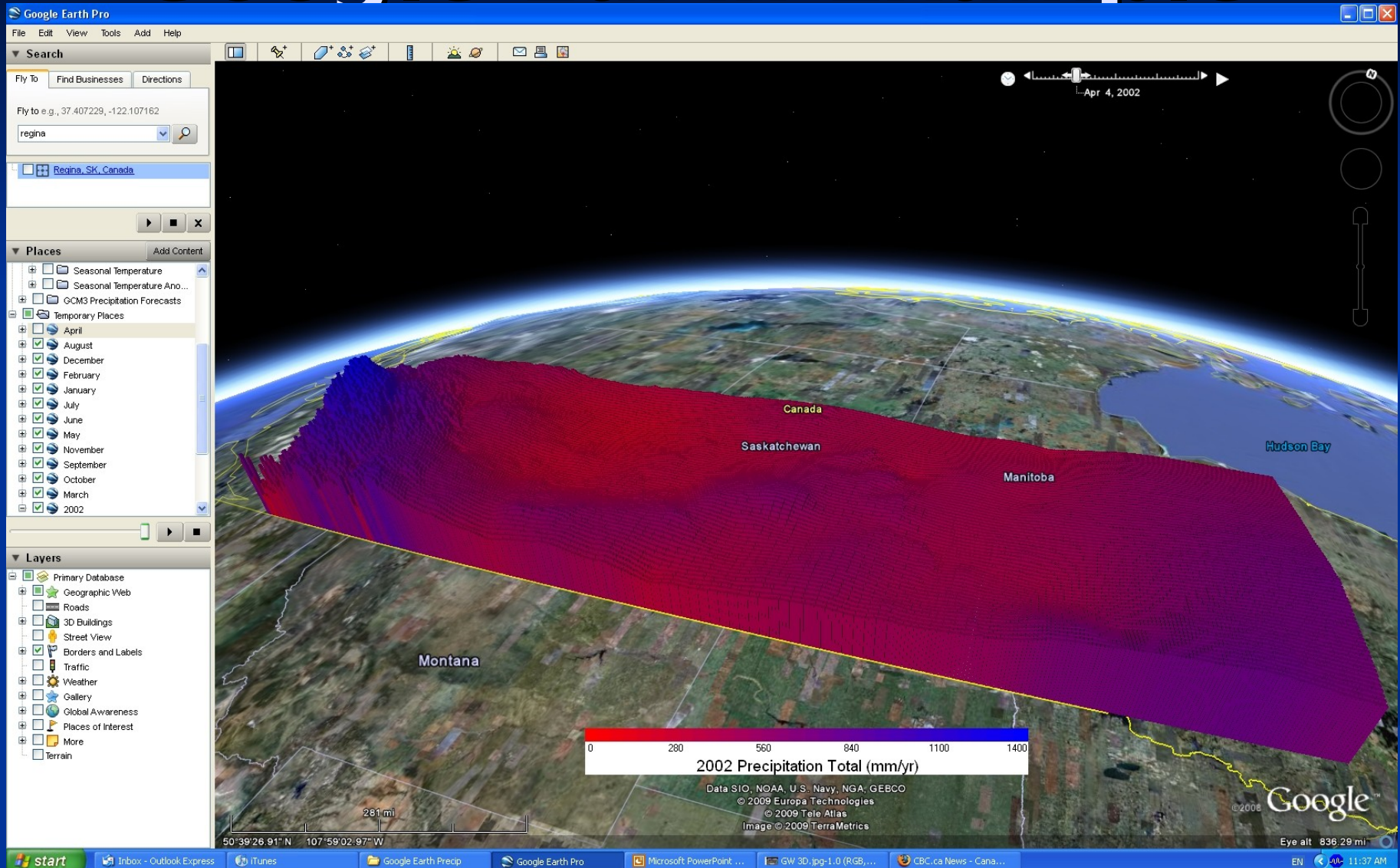
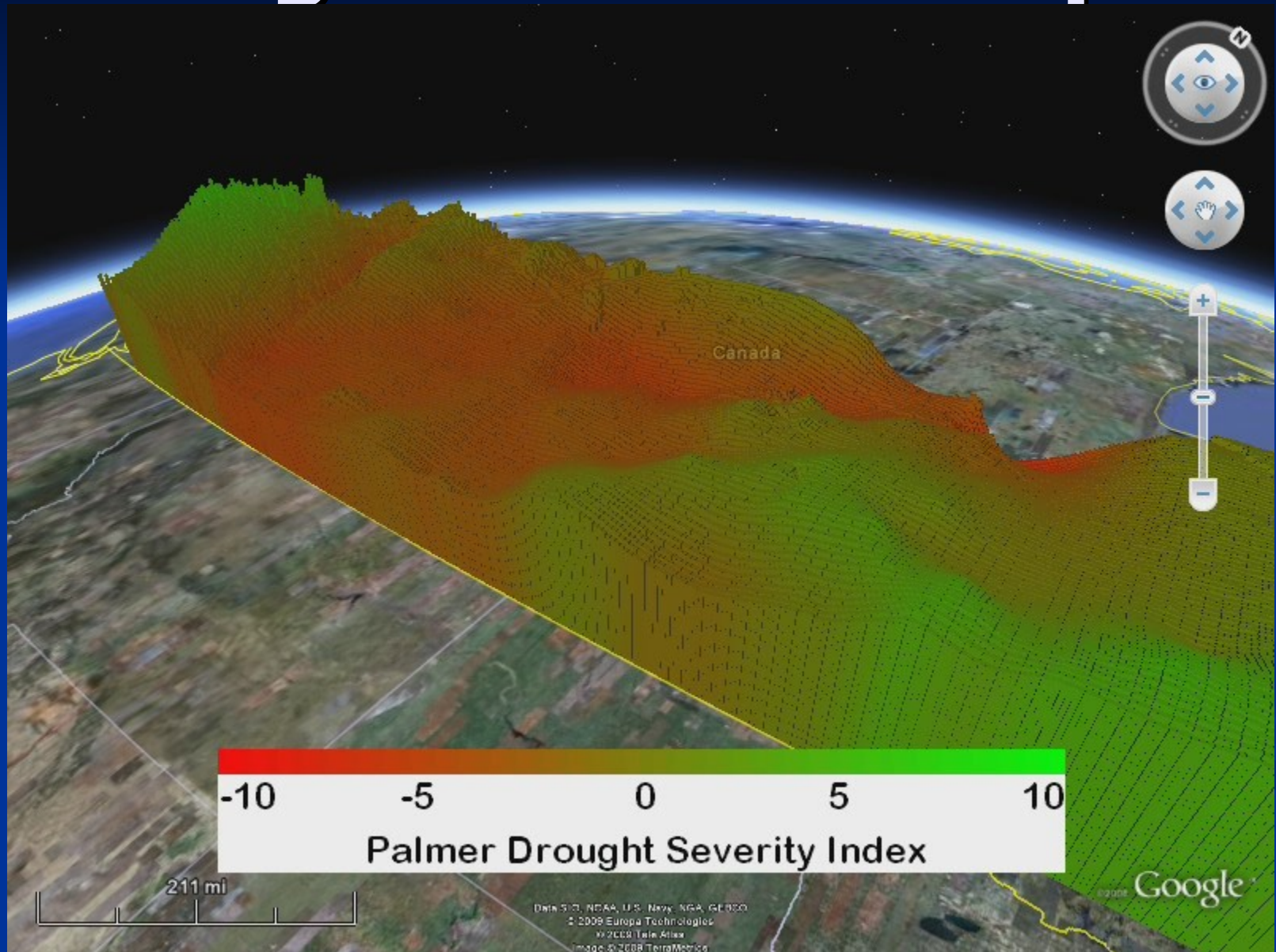# Data Management Schema

Google Earth Example

Google Earth Example

Google Earth Example

Google Earth Example

Questions that need to be answered:

What metadata standard will be used?

Where will the data legacy go after 2010?

How do we define what a DRI dataset is?

What DRI dataset needs to be included?

What level of data should be included?

Who should this data legacy be targeted towards?

Government, Researchers, Students, Public, all…

Data distribution policy?

A successful data legacy is completely dependent on participation!