How do we compare simulator results with paleoseismic data?

- Paleoseismic Data Assimilation
  - Small, sparse, analog data sets with qualitative uncertainty
  - PRIMARY vs NON-PRIMARY data
  - Data management problems and data ontology
- Comparison of simulator results with paleoseismic data
  - Data scoring and validation
- Comparison of simulation results
  - Requires same input data
  - QuakeTables/QuakeSim

Primary vs Non-primary paleoseismic data

Table 1

Proposed categories for cataloguing paleoseismic data

Primary data	**Direct field measurements (site specific):** - surface or near-surface slip - slip rate - date of last earthquake - dates of multiple paleo-earthquakes - minimum, maximum, average or preferred recurrence - location of fault trace - sense of motion - fault dip
Non-primary	**Interpretation or extrapolation of primary data:** - characteristic or estimated recurrence interval - magnitude of past earthquakes - spatial distribution of slip - length of paleo-earthquake rupture - location of fault segments - spatially averaged properties of segments
Descriptive	- fault name - strand name - segment name

Grant, L. B. and M. M. Gould. PAGEOPH, 161, no. 11/12, 2295-2306, 2004.

*R*elationships between faults, data and data sources

Outdated model

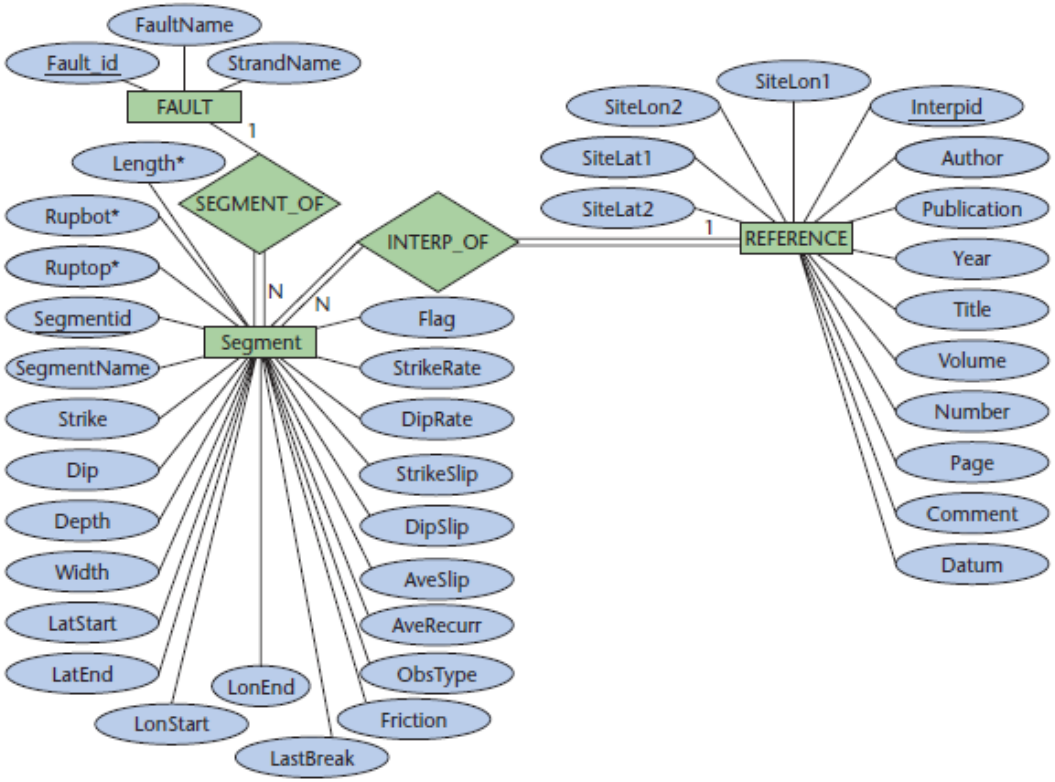


Figure 3. A simplified extended entity relationship model specification for the QuakeTables fault database shows relationships between attributes and faults.

- Data management challenges
  - Annotation
  - Provenance and pedigree
  - Curation
  - Access control

Grant, L. B., Gould, M. M., Donnellan, A., McLeod, D., Chen, A. Y., Sung, S., Pierce, M., Fox, G. C., and Rundle, P., A Web-service based universal approach to heterogeneous fault databases, Computing in Science and Engineering*, July/Aug. 2005, p. 51- 57*

Comparison of simulator results with paleoseismic data

An example: Scoring the data

- *Van Aalsburg, J., Grant, L. B., Yakolev, G., Rundle, P. B., Rundle, J. B., Turcotte, D. L., and Donnellan, A. A feasibility study of data assimilation in numerical simulations of earthquake fault systems.* Physics of the Earth and Planetary Interiors, *163 (2007) 149-162, doi:10.1016/j.pepi.2007.04.020*

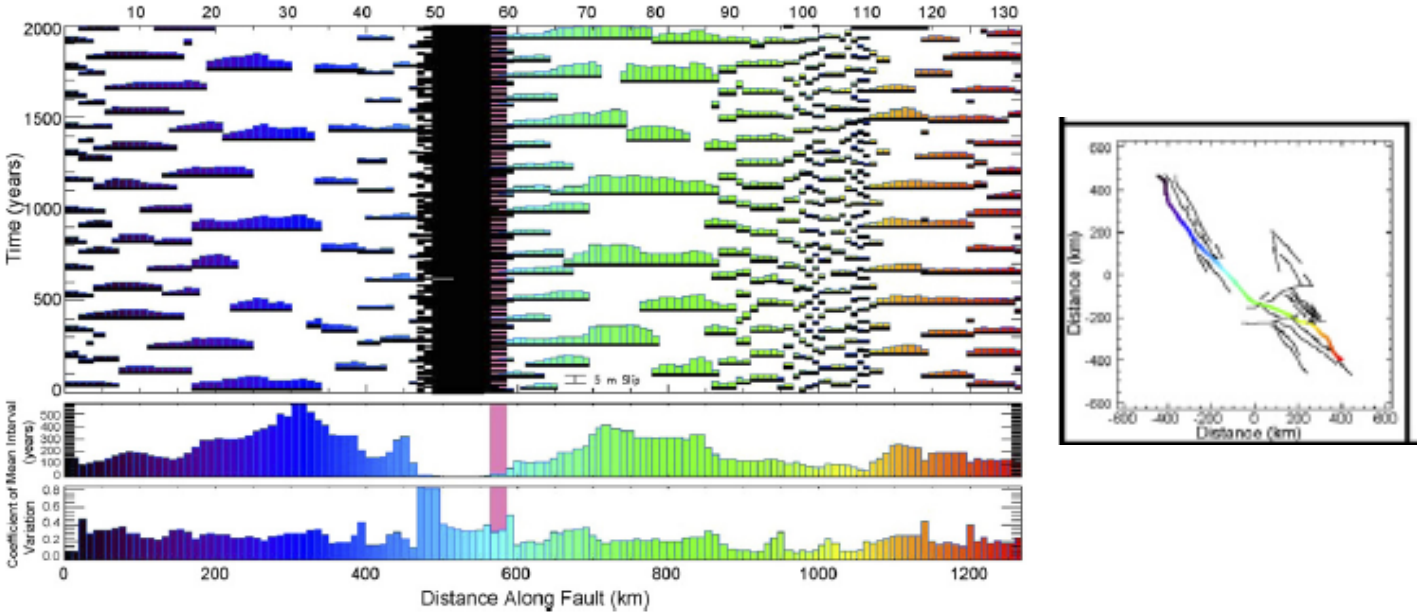


Fig. 2. Illustration of simulated earthquakes on the San Andreas fault. The slip for each earthquake over a 2000-year period is given as a function of the distance along the fault. North is to the left and south is to the right; shading corresponds to special shading in Fig. 1. Also included are the recurrence time mean and coefficient of variation for each segment.

Comparison of simulator results with paleoseismic data

Scoring the data (Van Aalsburg et al. 2007)

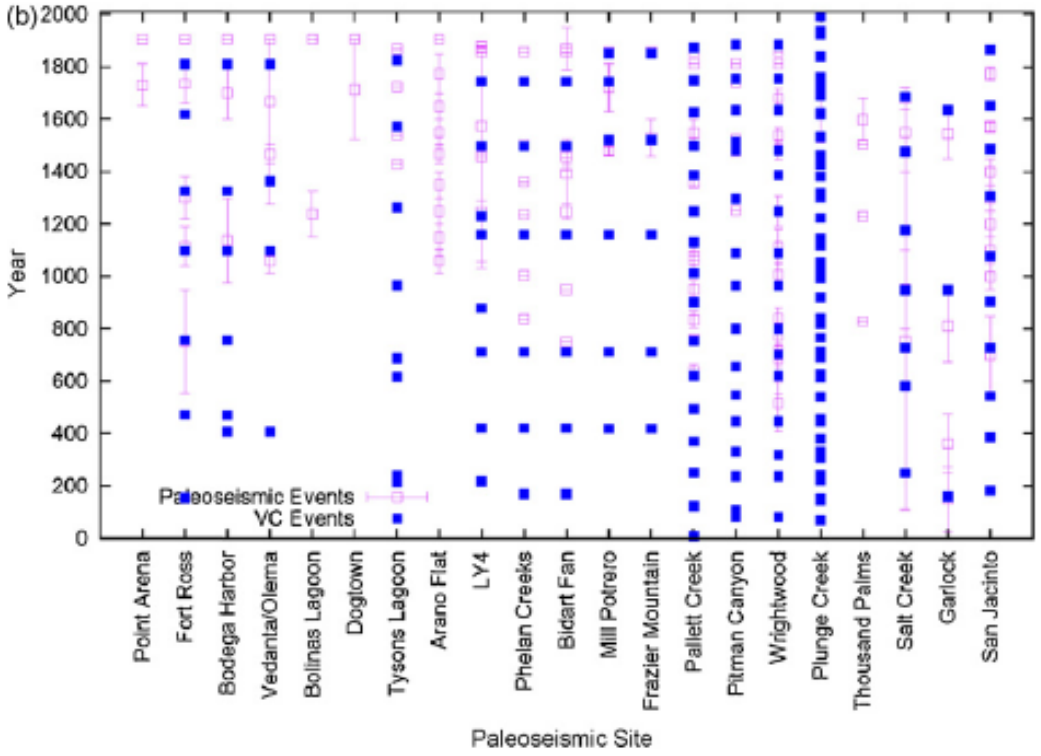


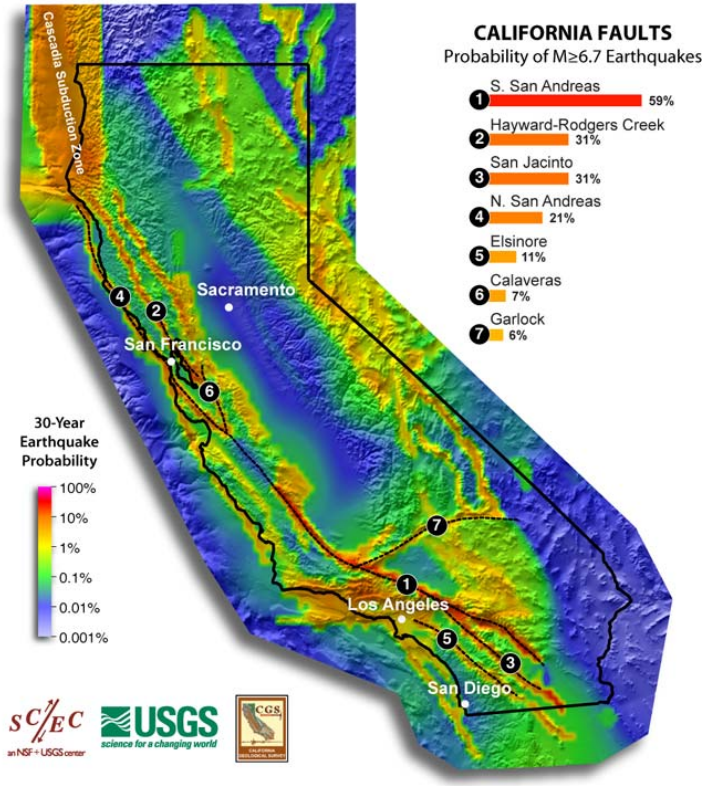
Fig. 5. Space–time diagrams of example *Virtual California* data for: (a) a low scoring year; (b) a high scoring year. Calendar year of the paleoseismic event is plotted against location of the paleoseismic site. The paleoseismic events are shown as red squares with error bars corresponding to the time window for that particular event. The *Virtual California* events during the time period are shown as solid dark squares. The times of the *Virtual California*. Note that the high scoring year has a much larger number of events which lie on or near the mean paleoseismic event time.

Feasibility study and comparison of results

– *Van Aalsburg, J., Rundle, J. B., Grant Ludwig, L. , Rundle, P. B., Yakolev, G., Turcotte, D. L., Donnellan, A., Tiampo, K. F.,and Fernandez-Torres, J.(in press). Space- and Time-Dependent Probabilitites for Earthquake Fault Systems from Numerical Simulations: Feasibility Study and First Results.* Pure and Applied Geophysics,

Relative spatial probabilities that at least 1 M > 7.0 earthquake will occur on a fault during the next 30 years (30-year probability per M > 7.0 event)

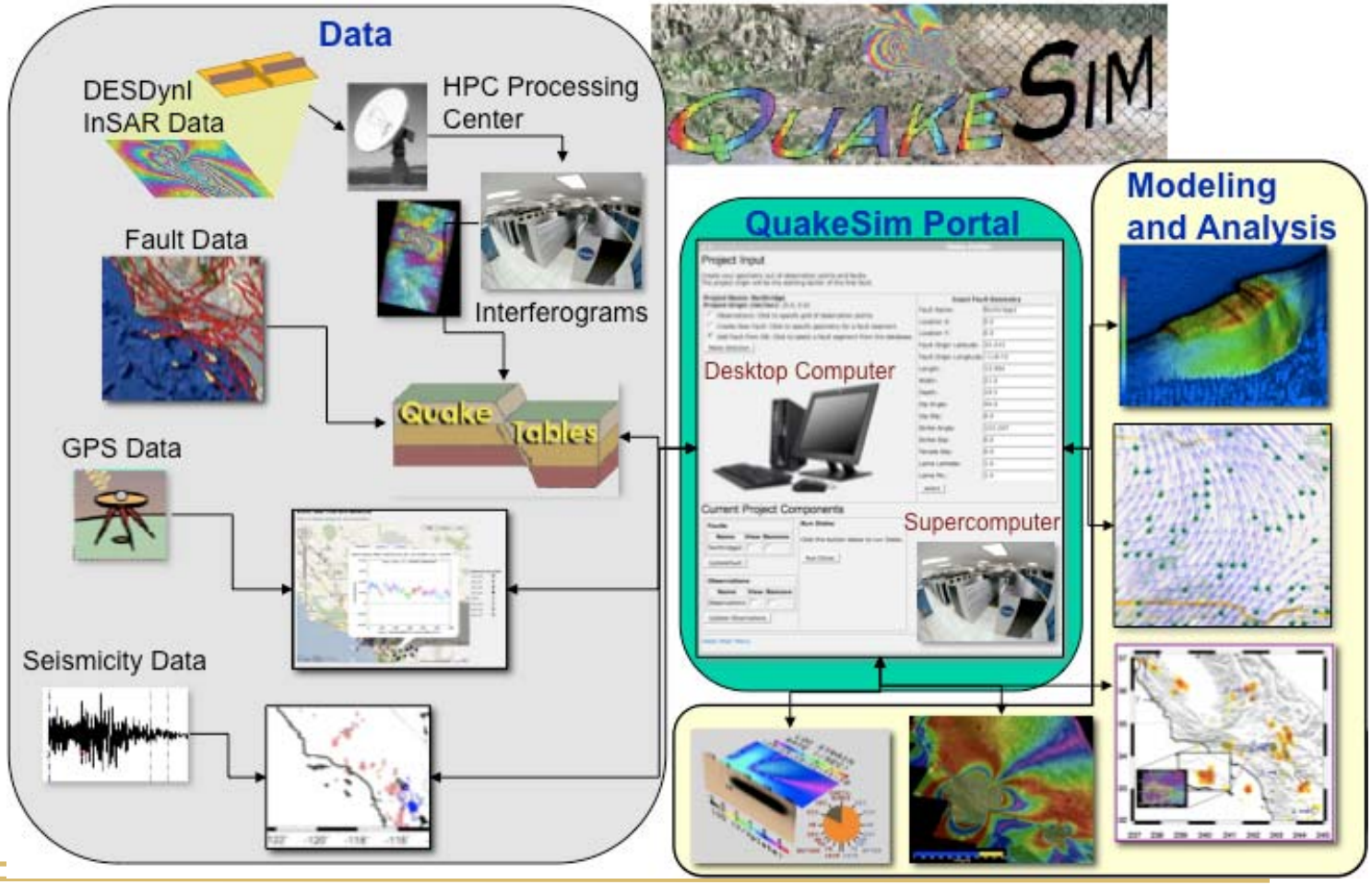
Fault	Eq. probability (%)	Fault length (km)
Bartlett Springs	3.1	85.0
Hayward	0.6	111.0
Hunting Creek—Berryessa	0.6	59.0
Maacama	2.5	179.0
Rodgers Creek	0.6	62.0
San Andreas North	32.6	467.0
San Andreas South	54.0	580.0
Garlock	5.3	234.0
White Wolf	0.8	47.0



For valid comparison, we need to use the <u>same</u> paleoseismic and fault data……

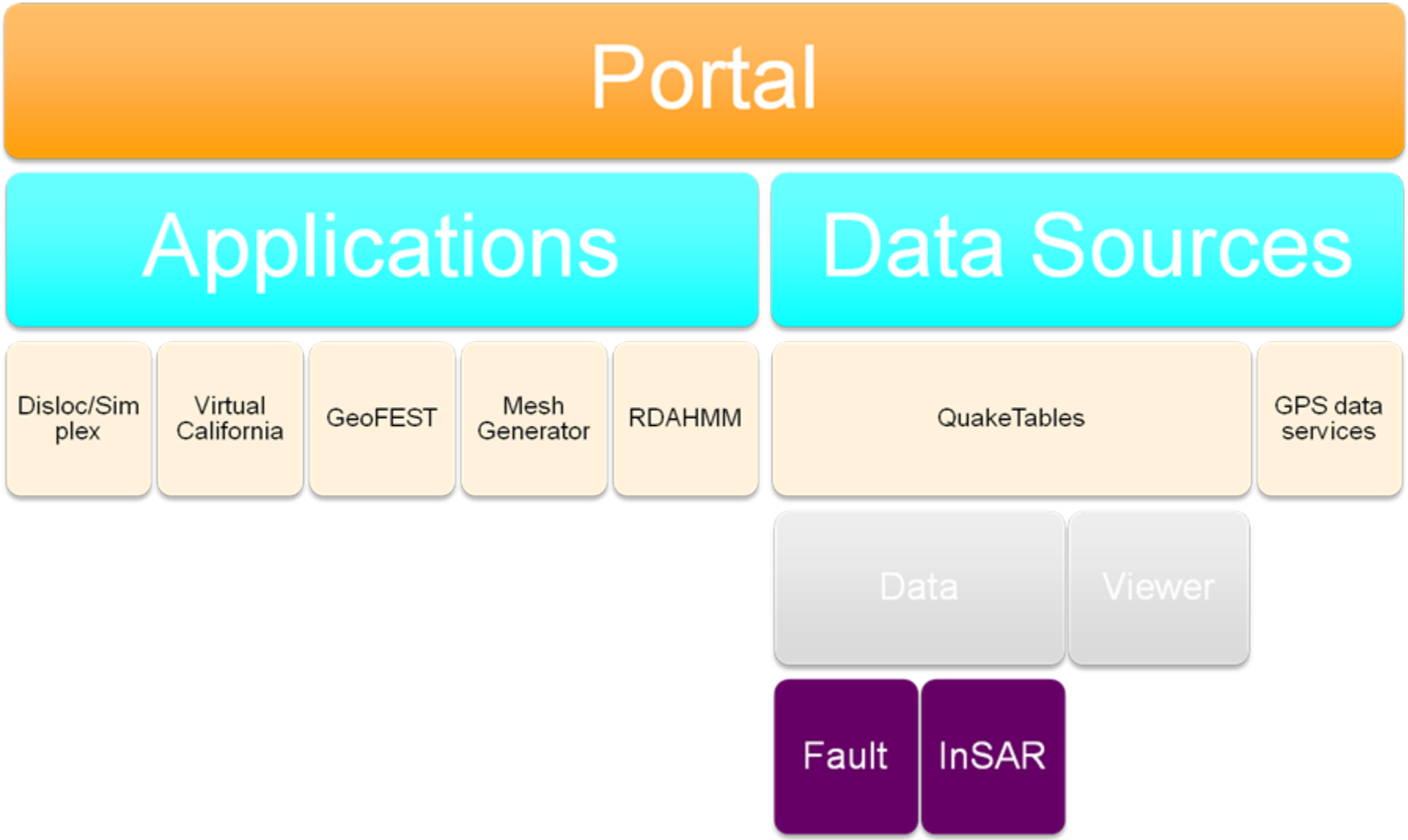
QuakeSim: Increasing Accessibility and Utility of Spaceborne and Ground-based Earthquake Fault Data

- Make crustal deformation, seismic and geologic data, and various earthquake simulation models available to the broader earthquake science community
- Focus on developing infrastructure to assimilate and model the large data volumes from UAVSAR and InSAR missions.
- Requires geologic faultdata to constrain and provide input to themodels



Andrea Donnellan, Jay Parker, Robert Granat, John Rundle, Lisa Grant, Dennis McLeod, Geoffrey Fox, Marlon Pierce, Rami Al-Ghanmi

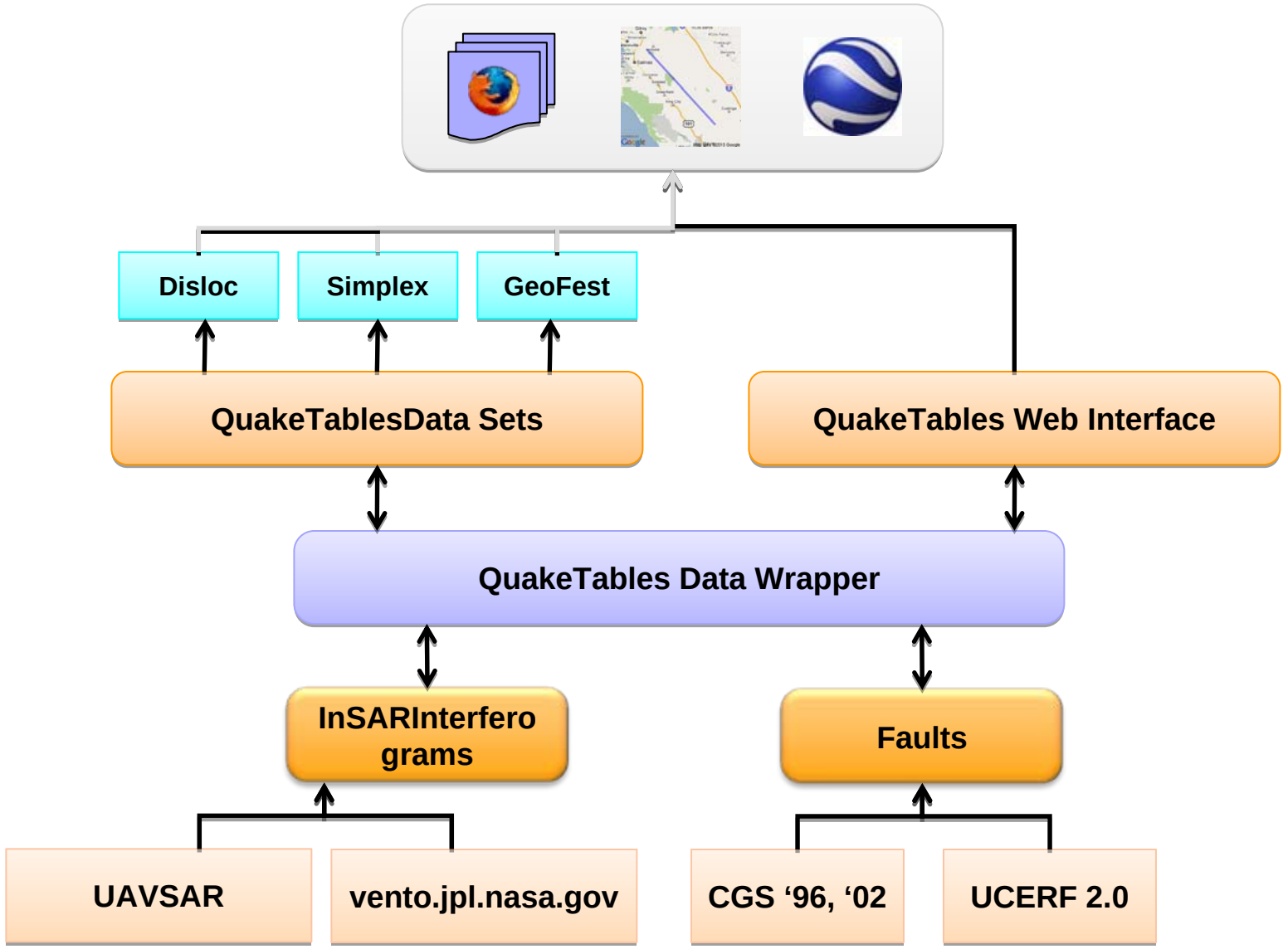
Architecture



QuakeTables

- QuakeTables provides QuakeSim with a uniform methodology to access data
- Access to QuakeTables is provided through a web interface and an API for data delivery
- The Ontology of QuakeTables provides access to:
  - Paleoseismic Fault Data: CGS, UCERF and Virtual California
  - InSARInterferograms
- Data is categorized into self consistent datasets that could be queried in its original form or a derivation of the original dataset
- Access to mapping features is provided through QuakeTables
- QuakeTables Web Interface
  - The QuakeTables web interface provides users with direct access to the QuakeTables federated data
  - Users can browse, map and navigate the available datasets

QuakeTables Architecture



QuakeTables Web Interface