What Do Geologists Want?
A Computer Scientist's Report on a Field Trip to Yosemite and Owens Valley

Yolanda Gil

Information Sciences Institute
and Department of Computer Science
University of Southern California
http://www.isi.edu/~gil
@yolandagil
gil@isi.edu
Abstract

What do thirteen geologists, twelve computer scientists, eleven compasses, ten bottles of hydrochloric acid, nine clean-up crews, eight ruggedized laptops, seven tablets, six-course dinners, five go-pro cameras, four vans, three high-school students, two linguists, and a psychologist have in common? A field trip organized by the NSF EarthCube EC3 project (Earth-Centered Communication for Cyberinfrastructure). The goal of the EC3 project is to understand the challenges of geologists to collect, manage, and integrate data collected in the field. By bringing computer scientists to see what geologists do in field trips, the project aims to uncover the cyberinfrastructure needs for field science. This talk will report on what I learned from the first field trip of the EC3 project, which took us to Yosemite National Park and Owens Valley on August 4-8, 2014. I will convey through pictures what we learned, and the opportunities for cyberinfrastructure in field science. I found that there are many opportunities for AI research, including speech recognition, sketching, data modeling and integration, geospatial reasoning and visualization, image processing, and robotics.
What Do These Have in Common?

• 13 geologists,
• 12 computer scientists,
• 11 compasses,
• 10 bottles of hydrochloric acid,
• 9 clean-up crews,
• 8 ruggedized laptops,
• 7 tablets,

• 6-course dinners,
• 5 go-pro cameras,
• 4 vans,
• 3 high-school students,
• 2 linguists, and
• 1 psychologist
Anna Zeng
Mission San Jose High School,
USC/ISI 2014 Summer Interns

Kevin Zeng

Paul Miyazaki
USC/ISI,
CS grad student
NSF’s EarthCube Initiative for Geosciences

Outcomes

Transform practices within the geosciences community spanning over the next decade

Provide unprecedented new capabilities to researchers and educators

Vastly improve the productivity of community

Accelerate research on the Earth system

Provide a knowledge management framework for the geosciences

http://www.earthcube.org
GeoSoft: Software Stewardship for Geosciences

Community
- Recommender system
- Interoperability

Publication
- Structured metadata
- Interactive advice

Learning
- Best practices
- Multimedia lessons

SoftCamp

TurboSoft

GeoSo@: Software Stewardship for Geosciences

Aug 22, 2014

gil@isi.edu
EC3

Earth-Centered Communication for Cyberinfrastructure — Challenges of Field Data Collection, Management, and Integration

This RCN will facilitate digitization of geological field data, enhance collaboration between the cyberinfrastructure community and those involved in geological field data collection, document what exists currently for field data collection issues, specifically targeting young investigators and the next generation. The RCN will conduct a series of workshops and townhalls at major conferences, catalog resources, and investigate data collection and digitization issues associated with digitization; and evaluate what is needed to improve the access and usability of geological field data. The RCN will focus on real-world, traceable, and long-term preservation of field data.

2014 Field Trip

EC3 is planning two field trips, one in Summer 2014 and one in Summer 2015, for researchers interested in contributing. All expenses will be reimbursed, and a small stipend will be provided. We are seeking to involve all different backgrounds in both the geosciences and the computer and information sciences community. Criteria for acceptance to the fieldtrip excursion will primarily be based on research expertise, and perspectives as possible. We strongly encourage graduate students applying to participate in the field trip forum should include a letter of recommendation.

The first field trip will take place on August 4-8, 2014 and will be to Yosemite National Park. Participants will stay in a field research center. A typical day schedule would include breakfast at the center, whole day out in the field including picnic lunch, dinner at the center, followed by group debrief and discussion.

Applications can be submitted using this form. For further inquiries, contact Matty Mookerjee at matty.mookerjee@sonoma.edu.
“Requirements Elicitation”

• Understanding what technology might be useful for
  – Protocol analysis
  – Develop models of their current processes
  – Develop what-if scenarios for potential technologies

• Preparing for the Trip:
  1. What do their papers look like?
  2. What information do they collect in the field?
  3. What technology do they currently use?
Understanding kinematic data from the Moine thrust zone in terms of a kinematics-based mathematical model of deforming thrust wedges

Matty Mookerjee a,*,1, Gautam Mitra b

a Department of Geology, Sonoma State University, Rohnert Park, CA 94928, USA
b Department of Earth and Environmental Sciences, University of Rochester, Rochester, NY 14627, USA

ARTICLE INFO

Article history:
Received 25 January 2009
Received in revised form
20 August 2009
Accepted 21 August 2009
Available online 28 August 2009

Keywords:
Thrust sheet
Kinematics
Numerical model
Moine thrust
Shear zone
Strain

ABSTRACT

A salient–recess pair along the Moine thrust zone exhibits significant lateral, kinematic variations. Specifically, the more steeply dipping portion of the thrust surface, associated with the recess, shows evidence for greater strain magnitudes and more out-of-the-plane motion than the adjoining salient. Additionally, a subset of the grain-shape data has grain-shape approximations with their long axes oriented perpendicular to the regional transport direction. To more fully understand these kinematic features, we developed a kinematics-based mathematical model of thrust wedge deformation.

Our numerical model predicts an increase in strain magnitude as well as non-plane-strain flattening in the region of a thrust ramp, consistent with the strain patterns observed within the recess along the Moine thrust zone. Furthermore, in the absence of a lateral-confining boundary condition, much of the model thrust wedge experiences transport-perpendicular, maximum extension except for the region closest to the thrust where simple-shear deformation is dominant. Using our numerical model, we predict the amount of lateral-confining strain needed to induce plane-strain deformation ($\varepsilon_{\text{ps}}$) as well as the amount of lateral-confining strain necessary to make the long axis of the strain ellipsoid parallel with the transport direction ($\varepsilon_{\text{parallel}}$). We have found that the incremental $\varepsilon_{\text{parallel}}$ can be as much as 42% of the incremental strain magnitude. Thus, we consider transport-perpendicular lineations like those seen along the Moine thrust zone to be the result of lateral-confining boundary conditions that are weak enough to allow transport-perpendicular maximum extension.
Photomicrographs

Maps

Texture

X-rays
Strains

\[ t_{bc} = \left\{ \begin{array}{ll}
\frac{1}{\sqrt{6}} (\sqrt{\cos(\theta) - 2\gamma + \epsilon_y} - \sqrt{\cos(\theta) - 2\gamma + \epsilon_y}) & \\
\frac{2\sin(\theta)}{\sqrt{6}} (\sqrt{\cos(\theta) - 2\gamma + \epsilon_y}) & \\
\frac{1}{\sqrt{6}} (\sqrt{\cos(\theta) - 2\gamma + \epsilon_y}) & \\
\frac{1}{\sqrt{6}} (\sqrt{\cos(\theta) - 2\gamma + \epsilon_y}) & \\
\end{array} \right. \]

Model

Results

- Plane strain, i.e., \( k = 1.0 \)
- Non-plane strain, e.g., \( k = 0.5 \)
A Field Notebook
Mobile Apps for Field Research
Features of Mobile Apps

• Connection assumptions
• Data and metadata collected
  – Location
  – Timestamp
  – Strike/dip
  – Rock type
  – Pictures
• Data export to other formats
  – CSV
• Mapping
## Mobile Apps for Field Research

<table>
<thead>
<tr>
<th>App Name</th>
<th>Platform(s)</th>
<th>On-Start</th>
<th>To Use</th>
<th>Connection Assumptions</th>
<th>Metadata Collected</th>
<th>Data Export</th>
<th>User Input</th>
<th>User Experience</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lamport</td>
<td>iOS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>rock type, data view (alone/line), comment, location (all editable except timestamp &amp; geo code)</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps</td>
</tr>
<tr>
<td>FieldBox</td>
<td>iOS</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>rock type, data view, comment, location, format (all editable except timestamp &amp; geo code)</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps, Mapbox</td>
</tr>
<tr>
<td>Rocktagger</td>
<td>Android</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>rock type, plane type, comment, metadata, field, depth, elevation, photo/image/video, comment</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps, Mapbox</td>
</tr>
<tr>
<td>Geocampus</td>
<td>iOS</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>person recording the sample, photo/image/video, rock type, name, thickness, rock age, geologic age, tempeature, pressure, fossils, etc.</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps</td>
</tr>
<tr>
<td>Geology Sample Collector</td>
<td>Android</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>can create website for export</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps</td>
</tr>
<tr>
<td>Strike and Sp</td>
<td>Android</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>can create website for export</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps</td>
</tr>
<tr>
<td>GeoID</td>
<td>Android</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>rock unit, rock type, natural description, contact, fold, fault, joint, with picture if descriptions, more than sure how to get that working?</td>
<td>CSV</td>
<td>1</td>
<td>0</td>
<td>Google Maps</td>
</tr>
<tr>
<td>Collector</td>
<td>Android</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>can create website for export</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps</td>
</tr>
<tr>
<td>FreeNokia</td>
<td>iOS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>rock type, data view, photo/image/video, comment, location, format (all editable except timestamp &amp; geo code)</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps</td>
</tr>
<tr>
<td>Geolines Plus</td>
<td>Android</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>can create website for export</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps</td>
</tr>
<tr>
<td>eGIS Compass</td>
<td>iOS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>rock type, data view, photo/image/video, comment, location, format (all editable except timestamp &amp; geo code)</td>
<td>CSV</td>
<td>1</td>
<td>1</td>
<td>Google Maps</td>
</tr>
<tr>
<td>iGPS</td>
<td>Android</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>can create website for export</td>
<td>CSV</td>
<td>1</td>
<td>0</td>
<td>Google Maps</td>
</tr>
</tbody>
</table>
DAY 1

• Locations in Yosemite: Lembert Dome, Olmsted Point
  • Learning to use a Brunton compass
  • Measurement collection
Dan Vieira, Sonoma State U., UG Geo Major
“Some geothermal action”
No Service!!??!!
NEWS RELEASE

For Immediate Release
Date: August 2, 2014 Time: 8:00 AM

French Fire

Location: The fire is burning in the area between Rock Creek and Fish Camp in the San Joaquin River drainage, Bass Lake Recreation District, in the Sierra Nevada Mountains.

Acres Burned: 11,466
Containment: 15%
Fire Started: July 28, 2014
Cause: Under investigation
Total personnel assigned to the fire: 1560

Summary

Fire activity continued throughout the night in the north and northwest perimeters of the fire. The fire was started last night in the area near Wagners Mammoth Pool Resort on the northwest perimeter. The fire is expected to be active again in the northern perimeter later this evening, as temperatures increase and wind gusts are expected to cause spotting and to spread the fire into more dense fuels. The fire is currently 85% contained and is burning in heavy timber on the Sierra National Forest and is currently 11,466 acres and 15% contained. The fire is still several miles from the Inyo National Forest and poses no threat of spread to the Eastern Sierra communities.

The French Fire will continue to produce large amounts of smoke, which has impacted local communities. Smoke levels are typically greater in the morning and evening. Plan outdoor activities for times and places when smoke levels are expected to be lower. For up-to-date smoke conditions, visit the San Joaquin Valley Air Pollution District website at http://www.valleyair.org/wildfire.htm.

Evacuations

Residents living in the community of Arnold Meadow were notified to evacuate due to increased threat of the fire. Several summer homes in the Hogar Apple Ranch and Wagners Mammoth Pool Resort were evacuated earlier in the week.

Closures

A Forest Service Area Road and Trail Closure, Forest Order # 15-14-07, has been issued for the fire area. More information about the Forest Closure and the closure can be found at: http://sinyrf.com/wildfire/forest_order_15_14_07.html.

South Central Sierra Interagency Incident Management Team

Fire Information: (559) 877-2605
Follow this Incident on Twitter @Sierra_NF
For additional information visit: http://inciweb.nwcg.gov/incident/4013/

Sierra National Forest

Bishop, Calif., August 2, 2014 – Smoke from the French Fire near North Fork, CA continues to affect air quality in some Eastern Sierra communities including Bishop and Mammoth Lakes.

The French Fire started on July 23rd and the cause is still under investigation. The fire is burning in heavy timber on the Sierra National Forest and is currently 11,466 acres and 15% contained. The fire is still several miles from the Inyo National Forest and poses no threat of spread to Eastern Sierra communities.

Because of recent weather patterns and wind direction, smoke from the French Fire is pushing east, over the Sierra Crest into the northern portion of the Inyo National Forest. It is likely that the smoke will continue to impact these areas over the next several days, reducing visibility and affecting air quality.

For more information on this fire contact the French Fire Information Line @ 559-877-2505 or check the internet via InciWeb at http://inciweb.nwcg.gov/incident/4013/

Local air quality information can be found at http://www.qbaaped.org/data/index.htm

-30-
Entrance Fee Waiver for Resource Education Study
National Park Service
Yosemite National Park
(Please print clearly or type)

School Name: Sonoma State University
Department: Geology

Mailing Address: 1801 E. Cotati Ave.
Rohnert Park, CA 94928

Phone: 707-664-2002
Fax: 707-664-3975

Instructor: Matty Mookerjee
Contact Person: Matty Mookerjee

Date of Entry: 08/04/2014
Date of Departure: 08/04/2014

Number of Buses: 0
Vans: 5
Autos: 0

Number of Students: 10
Grade(s): College
Total Number in Group: 35

(s) of Study: Geology and Cyberinfrastructure

Park Area(s) to be Visited: Olmsted Point, Lembert Dome, Tioga Pass
Matty Mookerjee, Sonoma State U., Structural Geologist
Ruggedized Laptops
Lembert Dome
Terry Pavlis,
UT El Paso,
Structural Geologist
Brunton compass
But how do you open this thing?
Granite Composition
Classifying Rocks
Not an “Outcrop”
Measuring Strike and Dip
8/4/14
EC EC3 field trip

Rain (sucky)
Can’t see rocks
Don’t think best
Yosemite Lembert Dome
Granite slab

Granodiorite
5-7%

Fracture

Bearing/strike N
Dip 56° below horizon, to the west

Geo goniometer 17° (from south)
63° inclination

Should name Chimes to account for mix-up
Clinometer
Inconsistent readings!!??!
Dog Lake
Stables
Parking
Soda Springs Trailhead to Glen Aulin
Mike Williams, UMass Amherst, Petrologist
White Mountain Research Center
Cyberinfrastructure Opportunities (I)

• Current technology is not terribly useful to take measurements in the field
  – “Strike and dip” measurements are only reliable (mostly) from a compass
  – Digital devices (eg clinometer) are often incorrect
  – Mobile/Tablet devices give incorrect readings
  – Mobile/Tablet apps are not conveniently designed
    • Requirements are far from met in those devices

• When technology works, scientist can be an order of magnitude more efficient (10x)
  – Important to collect as much data as possible
Cyberinfrastructure Opportunities (II)

• Many observations are still qualitative
  – E.g., “Vein is finer grained than the rock it is on, it is mostly quartz and feldespar, fracture is very irregular”, “black grains are needle-like showing cleavage”

• Pictures are helpful
  – But pictures need to be geolocated, linked to notes, annotated with highlights
  – Do not convey 3D features
Cyberinfrastructure Opportunities (III)

• Samples only advertised through publications
  – Can be very costly (eg $500K, 70lb)
  – Stored in basement (eg 4,000 samples)
  – Samples are often only partially exploited
  – Easier to locate samples from cores or boreholes

• No metadata standards to describe samples
  – Some to describe basic things (eg SESAR)
  – Willingness to describe samples moving forward
Cyberinfrastructure Opportunities (IV)

- Samples collected lead to digital assets that are never shared
  - Geochemistry: composition
  - Geochronology: dating
  - Microscope: grain and structure
  - X-ray: texture
DAY 2

- Location: Hind’s Ridge
  - Paleontology
  - Stratigraphy
- Location: Devil’s Gate
  - Deformation structures, faults and folds, and primary up-indicators
  - Sketching
Hind’s Ridge
Freiberger Compass
Magdalena Donahue, UNM, Geomorphologist
Zachary Michels, UW Madison, Structural geologist
Marjorie Chan, U. Utah, Sedimentologist
Stratigraphy Maps
Data vs Interpretation
Cyberinfrastructure Opportunities (V)

• Heterogeneity in collection of data/samples
  – Who is collecting:
    • Structural geologists: strike/dip
    • Sedimentologists: stratigraphy graphs
    • Paleontologists: ii, bbi
    • Petrologists: samples
    • Geomorphologists: landscape observations
  – What questions they have
  – How much time they have
  – How tired or wet they are!
Cyberinfrastructure Opportunities (VI)

• No technology to support sedimentology
  – 2/3 of all surface rocks are sediments
  – They contain a lot of life history
• “Digital stratigraphy graphs”
  – Scale, eg 2cm to 10cm
  – Symbology, eg shale vs limestone
  – Data vs interpretation
• Retrieving, reusing, comparing stratigraphy graphs
  – Would help with model development and testing
Cyberinfrastructure Opportunities (VII)

• Importance of sketching as a conceptualization tool
  – “Sketches is how we think, they have a cognitive function, it is how we piece information together” (M. Mookerjee)

• Need for hands-free data collection
  – Preference for speech recognition over note taking
DAY 3

• Location: Poleta Folds
  • Mapping structures
  • Using ruggedized laptops

• Location: Sage Hen Flat Pluton, Grandview camp ground
  • Distinguishing bedding and cleavage
Poleta Folds
Acid??!?
Joe Andrew, U. Kansas, Structural geologist
Can you see anything??
Finding ourselves in a map
Does anybody need a walking stick?
Walk the “Contact”
Definitely limestone!
Quartzite on one side
Shale on the other side
We have a datapoint!
But which way is up?
A “Rosetta Stone”
Fault vs no fault
Limestone stylolites!
Cleavage
The bedding
Dolomite vs quartzite/sandstone
Protect the Ancients
Please Stay On The Trails.
Collection of Wood is Prohibited
Cloudina fossil, 545 MYA
First skeletons!
There is no dishwasher in the kitchen—it is a sanitizer. Please wash (rinse fully) before putting dishes in the rack. Thank you!

- Turmeric scented ginger brown rice
- Chicken+lamb+pork jambalaya
- Parsley garlic mashed potatoes
- Miso glazed salmon
- Roasted kale
- French bread

Dessert:
- Caramel ice cream
DINNER CLEAN-UP 6:30 PM (or sooner)  
(6 PEOPLE)  
PLEASE WASH HANDS BEFORE TASKS!

DISH DAWG - 2 PEOPLE (6:20pm)  
- Scrape & rinse all dishes, pots, pans  
- Run through washer/sanitizer  
- Put away all pots, pans - Dishes will air-dry  
- When dishes are done, clean sink with cleanser & empty strainers  

FOOD CLEANUP / STORAGE - 2 PEOPLE (6:30pm)  
- Place leftover food in containers  
- Label and DATE containers (use masking tape) & take to Researcher Kitchen Fridge  
- Throw out any mystery food & food older than one week that is in the Researcher Kitchen fridge  
- Clean all kitchen surfaces  
  - stainless countertops  
  - wood counters  
  - stovetop - remove grates & clean w/degreaser  
  - mixing table

BROOM & MOP - 2 PEOPLE (6:40pm)  
- Fold chairs and place against wall  
- Clean all tables & serving areas and coffee station w/ towel & sanitizer  
- Take out trash (goes in dumpster out back) & replace bags  
- Sweep & Mop Kitchen & dining room  
- Sweep & mop walk-in fridge & pantry as needed
Cyberinfrastructure Opportunities (VIII)

• Scouting requires hours of investment
  – Before data collection: looking for perspective, the right outcrop, the right exposure
  – Would help to know what other samples have been collected in the area and where

• Dealing with obstruction and inaccessibility
Cyberinfrastructure Opportunities (IX)

• Trip preparation is time consuming and never complete
  – Prior publications about the area
    • Incorporate geochemistry and geochronology results
  – Maps and other data

• Integrating what is known while doing field work could be a game changer
  – Currently this is done in the lab after the trip
  – Could enable dynamic adjustments to data collection the field
Cyberinfrastructure Opportunities (X)

• Multimedia could be very helpful
  – Visualizing formations, lithostratigraphic units
  – Visualizing evolution over time
  – Augmented reality
DAY 4

- Locations: Bishop Tuff, Owens River Gorge
  - Vulcanic eruptions
  - Columnar joints
- Locations in Mammoth Mountain: Rosy-Finch Shear Zone
  - Metamorphic fabrics and shear zones
The Long Valley Caldera
Magnetic polarity: S/N
Columnar joints
Building Bridges Across Disciplines: How is Field Science Done?
Building Bridges Across Disciplines: How is Field Science Done? (v2)
What We Did Not Learn

• What is the actual process of fieldwork?
  – A whole day, a whole week, a whole month

• What do they do the other 10 months back at the office?
  – How do they use the data?
  – What are they missing from the field?

• How could we capture their models?
  – Stratigraphs, sketches, maps, ...
Was a Field Trip Necessary?

• Immersive learning
• Get our full attention
• Forge collaborations
• Get us field credentials

• Field science is fascinating!
Summary:
Major Needs in Field Science

• Field scientists have very little technology to support their research
  – In the field
  – In the lab

• Field scientists lack cyberinfrastructure to share their data and models
  – Lack of standards and community repositories

• Field scientists generate significant digital assets that are not currently exploited
Conclusions

• Lots of opportunities for cyberinfrastructure, in particular for AI research:
  – Data/knowledge modeling and integration
  – Workflows and metadata
    • Raw data + derived digital products
  – Geospatial reasoning
  – Voice activated assistants
  – Conceptual sketching
  – Document retrieval and data integration for trip preparation
  – 3D reconstruction from pictures
  – Robotic scouts
(Did I Deliver on My Promise?)

- 13 geologists,
- 12 computer scientists,
- 11 compasses,
- 10 bottles of hydrochloric acid,
- 9 clean-up crews,
- 8 ruggedized laptops,
- 7 tablets,
- 6-course dinners,
- 5 go-pro cameras,
- 4 vans,
- 3 high-school students,
- 2 linguists, and
- 1 psychologist
(Did I Deliver on My Promise?)

✓ 13 geologists,
✓ 12 computer scientists,
✓ 11 compasses,
✓ 10 bottles of hydrochloric acid,
✓ 9 clean-up crews,
✓ 8 ruggedized laptops,
✓ 7 tablets,
✓ 6-course dinners,
✓ 5 go-pro cameras,
✓ 4 vans,
✓ 3 high-school students,
✓ 2 linguists, and
✓ 1 psychologist
What Next?

• 2015 EarthCube EC3 Field Trip
  – Details will be posted at:
    http://workspace.earthcube.org/ec3
Thanks!