Workshop Meeting notes

Notes from the world cafe sessions:

https://drive.google.com/drive/folders/0BzY2TG1VpXRoX3dMMDN0UXRyOU0?usp=sharing

Welcome Presentation by Suzanne Pierce
Lightning Talks

Skinner: Robotics, 3D reconstruction, imaging surveys
Restrepo: Bayesian analytics,
Pennington: geoinformatics and Spatial data mining, team science, physical geography
Lie: Robotics, perceptual capabilities, 2-D forward scanning sonar, ship hull analysis, underwater surveys, active path planning, color correction and reconstruction
Babaie: structural geology, hydrogeology of fractured rock, ontology development for polycrystalline rock plastic deformation, data attributes
Thompson: robotics and machine learning, remote sensing with airborne imagery and spectroscopy,
Knoblock: Building knowledge graphs, semantic data and integration,
Cabral-Cano: Digital cartography, geology and remote sensing, urban ground subsidence, geodetic infrastructure
Arora: HPC, scalable and user-friendly infrastructure for datasets, analytics on heterogeneous datatypes, metadata use for provenance within data collection, science gateways which are special portals for user groups
Krum: immersive experiences (Virtual and augmented), VR tools for geoscientists, low cost VR tools, drone and photogrammetry pipeline brings in terrain data into VR, imaging pipeline for fine detailed objects (e.g. feathers)
Ebert-Uphoff: Knowledge discovery from data in geosciences, new applications of machine learning methods
Gel: Statistician, machine learning for uncertainty quantification for climate-induced risks from natural non-catastrophic events, data fusion of multi-source information, visualization of peril maps and vulnerability
Hill: INFEWS decision tool, federal incentives for agriculture, energy, and water management, interactive web-based,
Christophersen: Automatic image recognition for Mars dataset. Machine learning and automated detection using images.
Eftelioglu: Context aware spatial data mining (e.g. clustering), remove chance patterns (statistical significance). Eliminate the iid assumptions in traditional machine learning methods.
Guereque: Satellite image processing, dust concentration/aerosol models, temporal and spatial data, large and big dataset that is heterogeneous, Geo-computational analysis in outreach and STEM education
Garijo: Data Narratives, workflow provenance, and automating human readability using automated narrative generation from scientific workflow results

Samet: Spatial representations, mobile computing and location-based analyses, anonymized location brokers,
Fuka: Genetic by genomic applications (e.g. taking EarthCube applications in other directorates), coupling atmospheric and surface systems
Ravela: Machine learning with atmospheric data, adaptive sampling, autonomous mapping, mixed machine learning and computer vision systems,
Pankratius: Geo-informatics, Computer-aided discovery, mobile data collection and processing, parallel / cloud computing
Walling: Software development and statistics, R development and programming with non-traditional users, tool development,
Plale: Persistent identifiers with handles and propogation, Universal data object provenance
Pierce: Interactive data dashboards and reusable decision support tools
Gil: Capturing scientific knowledge and using to communicate, semantic wikis, metadata, ontology and collaborative data management/provenance, automated discovery using worksflows and lines of inquiry approacheds, Discovery process documentation and automation

Notes from the world cafe sessions:

https://drive.google.com/drive/folders/0BzY2TG1VpXRoX3dMMDN0UXRyOU0?usp=sharing

Collaboration Interests:
Interactive analytics, education, knowledge maps, spatial data mining, complex non-linear and dynamic systems, find general aspects of Earth Sciences, bridging physically-based models with learned components, next generation remote sensing, automatic indexing geoscience information, distributed access, expert systems, Machine learning and pattern detection in data streams real-time, techniques for 4D problems, Skills development for non-traditional HPC and advanced computing users, large-scale dataset analyses techniques, science gateways, VR tools for geoscientists, drones with geomorphology, machine learning applications to big datasets, spacetime risk visualization, integration of data, multi-resolution and multi-scale data, Geocomputational education, crowdsourced identification techniques, climate risk, natural hazards, extreme events, and decisions with uncertain properties, Use data mining to identify real-world phenomenon, case study analysis and applications to geoscience problems, Reusable modules for education, shared datasets for education, testing with machine learning and/or provenance techniques, training modules

Glossary terms:

Subsidence -

Science gateway - portal provisioned for computational and analytical resources accessible for end user community

Notes from the world cafe sessions:

https://drive.google.com/drive/folders/0BzY2TG1VpXRoX3dMMDN0UXRyOU0?usp=sharing

Useful links and reads:

The NAS report:
https://www.nap.edu/catalog/9942/bridging-disciplines-in-the-brain-behavioral-and-clinical-sciences
Related reports (from Deana) on fostering collaboration between communities:
https://www.nap.edu/catalog/19007/enhancing-the-effectiveness-of-team-science
https://www.nap.edu/catalog/11153/facilitating-interdisciplinary-research

Papers for fostering **interdisciplinary research**:

The EMBeRS project: employing model-based reasoning
in socio-environmental synthesis:
https://www.researchgate.net/publication/282589272_The_EMBeRS_project_employing_model-based_reasoning_in_socio-environmental_synthesis

A conceptual model for knowledge integration in interdisciplinary teams: orchestrating individual learning and group processes:
https://www.researchgate.net/publication/286764020_A_conceptual_model_for_knowledge_integration_in_interdisciplinary_teams_orchestrating_individual_learning_and_group_processes

Transdisciplinary Research, Transformative Learning, and Transformative Science:
https://academic.oup.com/bioscience/article/63/7/564/289183/Transdisciplinary-Research-Transformative-Learning

Bridging the Disciplinary Divide: Co-Creating Research Ideas in eScience Teams:
http://link.springer.com/article/10.1007/s10606-011-9134-2

DRF: Notes

NOTES FROM THURSDAY

TOPIC: Takeaways from Monday

Imme:

- IS people often consider that data is data and you apply a method to it. The geoscience community views their datasets through the lens of hypothesis and purpose, and pre-processing matters a lot. Pre-processing strongly depends on the question you want to answer.
- IS sees it as data, and data is equally.
    - Comments from Juan...regarding importance of aiding IS and Geo researchers' ability to understand the different research interests and stances
    - Vipin… important point because there is a dichotomy between data driven in CS and hypothesis driven in geosciences (not always, but usually).Risk is that the IS output isn't relevant/useful because it's not well situated with GEO data context...have to take extra effort to find a mutually useful approach
    - Vipin..there are various physical sciences communities and then there are virtual sciences. There are artificial boundaries viewed within these domains, but the IS-GEO researcher can be both an effective scientist in either physical/virtual
    - Juan...doesn't matter what other groups 'call' us, we just practice and pursue strong research
    - Transcending the boundaries...we are trying to develop more generalizable and robust approaches that can be applied to other areas
    - Dan...therefore IS-GEO is not a new 'core' discipline it is a core focus on problem solving
    - Sai....in many respects it's similar to applied math (e.g. Operations Research), because they can apply generalizable approaches across problem cases, but as approaches have become more complex, they are beginning to specialize and focus on specific data types and problems...we are expanding **Computational Thinking within Domains**
    -
    -
- Vipin: How can we expedite the process of transferring knowledge from IS to geosciences - and vice versa?

Notes from the world cafe sessions:

https://drive.google.com/drive/folders/0BzY2TG1VpXRoX3dMMDN0UXRyOU0?usp=sharing

Interdisciplinary Teamwork - Deanna Pennington. This talk is available in Github: https://github.com/IS-GEO/Materials/blob/master/Talks/Pennington%20Interdisciplinary%20teamwork%20issues.pdf

- Team science - collaboration sometimes across disciplines to research a topic
- Current IS-GEO is multidisciplinary (individual goals under thematic umbrella) vs. interdisciplinary (single shared goal) and transdisciplinary (e.g. bioinformatics)
- Differences between multidisciplinary (starting point), interdisciplinary, transdisciplinary. We are looking to be interdisciplinary. To reach that point, we must identify relevant research problems.

- Everybody likes interdisciplinary research when it goes well. However, when it goes bad, it might be a disaster.
- It may take up to 3 years to figure out how to be productive. How can we make this process faster?
- Why is interdisciplinary collaboration difficult? Vocabulary, finding shared goals, time consuming, bad research problems, difficulties getting funding/publications, etc. There is evidence about this. It is not anecdotal.
- Strong leadership, communication, flexibility/adaptability and ability to learn are crucial for success.
- Research indicates someone's ability to learn is linked to mental models, which are mechanisms generated and built by every individual to handle knowledge.
- Experts generally have very rigid mental models that are difficult to modify, therefore intermediate experts such as postdoc and graduate students may do better connections between different areas, due to the flexibility of their mental models.
- Lateral learning problem: Iterate->Grasp->Reflect->Transform, connect, synthesize.
- It's important to reflect on the ideas after the discussion. The "Aha" moments. Try to capture the moments that sparked your interest.
- To facilitate the understanding of new knowledge, we must be able to communicate and teach our science in a way that requires the least minimal effort from our team collaborators.
- Pattern for success: Use examples that are close to the science that you want to make the connection to. Building these connections creates patterns that lead to emergence of synergistic research ideas, a shared vision and problem model.
- Yolanda: Myers-Briggs Personality type, see https://en.wikipedia.org/wiki/Myers%E2%80%93Briggs_Type_Indicator#Differences_from_Jung , helps to understand personalities and understand interactions in meetings, etc., that they cause.
- Science moves forward by model-based reasoning. Someone proposes a model and someone other proposes another model against it.
- Do the links with the people sitting on the table, not something that could be useful for someone that is not there.
- Phobias in collaboration.
- Short term vs long term aims.

Notes from the world cafe sessions:

https://drive.google.com/drive/folders/0BzY2TG1VpXRoX3dMMDN0UXRyOU0?usp=sharing

CONFERENCE DISCUSSION

Or possibly a workshop series that moves conference to conference?

Observations:

- This group is delving into deep discussions
- We do not see each other at meetings and it is rare to find a shared venue for the two groups
- Bringing about change will require engaging with younger researchers
- Finding opportunities to share approaches and perspectives

We will need to consider the different cultures and credit for research activities

- For example GEOs need 'credit' for peer-reviewed publications if possible: Could do a hybrid model that the publications for GEOs go to journals and the computer scientists submit to the conference
- Handling differences in IS people reviewing GEO papers or vice versa
- Another option would be to submit conference papers to a special issue in an existing journal (e.g. Earth Science Informatics or Computers and Geosciences) - comment that special issue once per year isn't enough to propel a community forward
    - Also there is the Dynamic Data Driven Environmental Sciences (DYDES??? - Sai?) - We could possibly model our own conference after this one or host a workshop at different conferences
    - One possible concern is that various venues have different traction, acceptance rates need to be about 20%....only goes to quality venues
- Journal first - doesn't offer the opportunity for interaction among participants as much as a conference/workshop
- For computer scientists the conference model with peer reviews is valuable, done in combination with a journal

Sponsorships

- Frequently conferences supported by professional societies, is there a sponsoring body

We will continue to discuss this topic. For now, it seems that we could pursue:

- A journal special issue
- A workshop at an IS conference - eg AAAI

Invitation to submit papers to conference:
https://sites.google.com/a/umn.edu/mbdce-2017/home

Notes from the world cafe sessions:

https://drive.google.com/drive/folders/0BzY2TG1VpXRoX3dMMDN0UXRyOU0?usp=sharing

POSSIBLE CHALLENGE AREA WORKING GROUPS:

Model-driven sensing:

1. Augmenting models with adaptive sensing of microscale processes.
    a. Katie Skinner, Suzanne, Sai Ravela
2. Optimal simulation for model reduction and model abstraction
    a. Jie Li, Suzanne, Sai Ravela, Katie Skinner, David R. Thompson
3. Reducing parameterizations of (chemical) transport and/or reaction models.
    a. Katie Skinner
4. Resilient disaster response policies based on best available information.

Knowledge maps:

1. Assisted authoring of papers based on the workflows and metadata of a project
    a. Daniel G.
2. Automatically coming up with representations of the data resulting from a computational analysis
    a. Daniel G., Yolanda G., Craig K., Suzanne Samet
3. Accelerating model integration by assisting with model selection, variable mapping, and data ingestion
    a. Yolanda G., Suzanne
4. Integrating data from sensors in time and space
    a. Craig K., Yolanda G.
5. Automating processes for analyzing the data
    a. Jie Li

Theory-guided learning:

1. Data with a story that includes question and validation as GEO-driven case studies for IS researchers
    a. Victor Pankratius, Sai Ravela, Emre, Mariana G., Imme
2. Develop methodology for different climate etc model evaluation and comparison using complex networks and random graphs
    a. Sai Ravela, Imme, Suzanne, Yulia

Trusted Threads

1. Real time knowledge extraction from a well curated sensor data collection
2. Curation of multi-sensor multi-instrument data
    a. Beth, Mariana G
3. Track evolution and history of data
    a. Daniel G., Emre
4. Improving secondary uses/re-purposing of data by capturing pre-processing steps
    a. Yolanda G., Beth, Daniel G., Craig K.

Atmospheric/Ocean/Polar Science

1. Prediction Challenge Problem: Combine Machine vision/learning model with numerical model to do better than either source alone for storm prediction, hurricane intensity prediction or tornado early warning
    a. Yulia R. Gel, Sai Ravela
2. Representation challenge problem: Produce a model data-driven approach to parameterize subgrid scale effects in deep convection.
    a. Sai Ravela, Katie Skinner
3. Understanding Challenge Problem: Classify and cluster numerical model output for anomalous, rare, extreme or transient events and assess the predictability of features.
    a. Sai Ravela, Emre
4. Downscaling Challenge: Produce a statistical-dynamical downscaling of significant rainfall in a changing climate with uncertain probabilities.
    a. Sai Ravela
5. Standardizing remote sensing data with an application to carbon concentration
6. Detection of methane point sources emitted from infrastructure through image analysis and machine learning
    a. David R. Thompson

Geospace

1. Create larger resolution maps of ionosphere to do new kinds of science (eg hazard monitoring)
    a. Enrique Cabral-Cano, Suzanne

Earth Science

1. Coupling of near surface models to the interfacing surface and atmospheric systems, as well as to the deep surface and other sinks.
    a. Daniel Fuka
2. Early warning or planning systems: enhance predictions through data integration, model integration, ontology/relations, graphical models, dashboards
    a. Enrique Cabral-Cano, Suzanne, Yolanda, Daniel G., Craig K. Jie Li, Daniel Fuka, Samet, Emre

POSSIBLE GROUPS: (Color coded to reference back to the proposed Challenge Area working groups from Thursday)

- Data Collection & Integration
    - Craig K., Yolanda G., Beth, Mariana G., Enrique Cabral-Cano, Suzanne, Yolanda, Daniel G., Craig K. Jie Li, Daniel Fuka, Samet, Emre
- Geoscience Case Studies, Benchmarks, Feature Detection / Discovery
    - Data with a story that includes question and validation as GEO-driven case studies for IS researchers
    - Victor Pankratius, Sai Ravela, Emre, Mariana G., Imme, Gulia, David R. Thompson
- Geo-Simulations (parametrizations, scales, integration of different models, etc.)
    - Jie Li, Suzanne, Sai Ravela, Katie Skinner, Sai Ravela, Emre

Imme's presentation

- Intellectual contribution model (GEO):
    - Necessary: must address important questions
    - Necessary: Well documented methods
- PLUS one of the following
    - New insights? YES
    - Better prediction? YES
    - New pattern with no insights?

  Sai, David disagree on the last point. Enrique added that it may happen in physics, but no geophysics.

IS:

- Must contain a new method/algorithm. (Daniel: not in some conferences/journals)
- Evaluation: should be done quantitatively. What if there is no ground truth?

PLUS

- Better prediction (YES)
- Existing algorithm applied to new application, yielding to better results (NO)

Yulia: in statistics there is a replication problem.
Publishing: for geo people, conferences don't count. Journals take less in GEO than IS.
Cross-disciplinary outlets? [SAI TO ADD SOME OF THE ONES HE MENTIONED HERE]
Yulia: we need a better outreach to improve understanding. Why do taxpayers need to fund us?

NOTES FROM FRIDAY

TOPIC: Takeaways from Overall Workshop

Summary of suggestions for activities to do as we move forward:

Future Webinars:

1. Open to the Public/Recorded for dissemination: To present peer-reviewed research by current IS-GEO members
2. Closed off and interactive: To allow members to present active research not yet published

Forum: Where members can pose open-ended questions to members of the IS-GEO community (Chair: Emre Eftelioglu)
More opportunities to recruit/involve students

SESSION: Combining IS and GEO Themes

3 Topics proposed for Working Groups (Color coded to reference back to the proposed Challenge Area working groups from Thursday):

- Topics:
    a. Sensor-based data Collection & Integration, possible application for early warning systems
        - From IS: Craig K., Beth, Daniel G., Samet, Emre
        - From GEO: Mariana G., Enrique Cabral-Cano, Suzanne, Jie Li
    b. Geoscience Case Studies, Benchmarks, Feature Detection / Discovery: Data with a story that includes question and validation as GEO-driven case studies for IS researchers
        - From IS: Victor Pankratius, Sai Ravela, Emre, Imme
        - From GEO: Mariana G., Enrique Cabral-Cano., Daniel Fuka,
    c. Geo-Simulations (parametrizations, scales, integration of different models, etc.)
        - From IS: Jie Li, Sai Ravela, Katie Skinner, David R. Thompson, Emre, Yolanda G.
        - From GEO: Suzanne, Daniel Fuka

Goals of this session:

- Each group discusses common topics of interest and defines one or more of these based on individual's interest:
    - Short term prototyping efforts
    - Mid-term proposal-size activities
    - Long-term driving vision
- Each group should consider setting up an IS-GEO Working Group:
    - Identify lead(s)
    - Report to the larger group
    - Prepare a brief description to publicise on IS-GEO web site
    - Set up group telecons/webinars and involve students and others signed up to IS-GEO

SUGGESTION: Each of the three groups should develop a write-up of the general topic [addressing the issue(s) to tackle and the datasets and tools that may be incorporated, to be distributed amongst the IS/GEO community, and facilitate recruitment of potential new members.

Case Studies and Benchmarks Group:

Geoscience Case Studies, Benchmarks, Feature Detection / Discovery: Data with a story that includes question and validation as GEO-driven case studies for IS researchers

- From IS: Victor Pankratius, Sai Ravela, Imme, Beth, Dave T., Yulia, Hassan
- From GEO: Mariana G., Daniel Fuka,

Notes from Imme (longer version) available here:
https://docs.google.com/document/d/1QSUjhgjAVYHw3Zl516eOmzh-MbhLY9QPhW_B-HwHyuU/edit?usp=sharing

Notes from Mariana (shorter version):
We must find a situation and data set that can function as both a case study and as a benchmark.

- Methane source detection - Dave T (ground truth)
- Vegetation response to onset of CA drought - Dave T (no ground truth)

- Detection of storms tied to climate change- Sai Ravela (ground truth available)

- Partitioning different processes that were responsible for the values it collected

- Simplified models for quantifying uncertainty and risk of surge storms to plan better for the future - Sai Ravela (ground truth available)
- Downscaling with uncertainty of resolution by using a circulation model at two resolutions fine and coarse), or use a specific event, and evaluate the distribution of change in error between both resolutions (given the relationship of variables at the specific scales) - Sai
- Simulation data for representative problems from the IS community (soft inverse problems to test inverse models): Imme
    - the Methane dataset, also unmixing of vegetation spectra - both proposed by Dave T
    - TOY problems in convective diffusion, 1. standard inverse problems 2. stochastic problems with probabilities and probability distribution, 3. Assistive modeling problems - Sai Ravela

Four main areas encompassed by the proposed case studies:
Inference from data,
inference from models output,
inference from data and models,
Learning models from data/Identifying processes.

Lead for this group: Imme

Geo-Simulations (parametrizations, scales, integration of different models, etc.)

- From IS: Jie Li, Sai Ravela, Katie Skinner, David R. Thompson, , Yolanda G.
- From GEO: Suzanne, Daniel Fuka

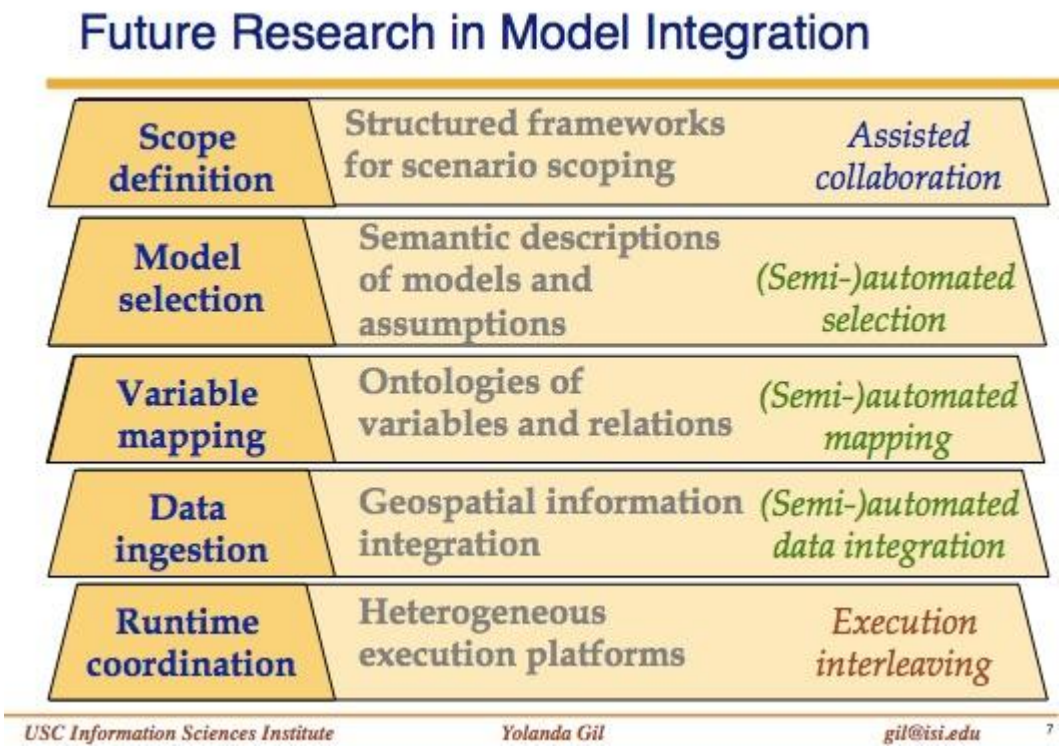
- Each group discusses common topics of interest and defines one or more of these based on individual's interest:
    - Short term prototyping efforts
    - Mid-term proposal-size activities
    - Long-term driving vision

- Each group should consider setting up an IS-GEO Working Group:
    - Identify lead(s)
    - Report to the larger group
    - Prepare a brief description to publicise on IS-GEO web site
    - Set up group telecons/webinars and involve students and others signed up to IS-GEO

What is each of us interested in:

- Emre: The effects of different scales on the output, e.g. not only spatial but also at temporal scales.

- Yolanda: Work with hydrologists, model integration, 1) need to define the scope, 2) then select the models, 3) perform the variable mapping, 4) data ingestion, 5) runtime coordination. Breakdowns fail at all stages.



- Daniel Fuka: Currently applied modelling in several realms. Primarily coupling three atmospheric domains (historical measurements, future climate forecasting, and real time and short term weather forecasts) to surface eco-hydrological modeling systems. Coupling these coupled models to ocean and estuary systems models, including coral health, estuary health, etc.
- Suzanne: 10 dimensions for model integration described in this paper (open access): http://www.sciencedirect.com/science/article/pii/S1364815214003600
- Decision support session, AGU chaired by National Labs fellow.

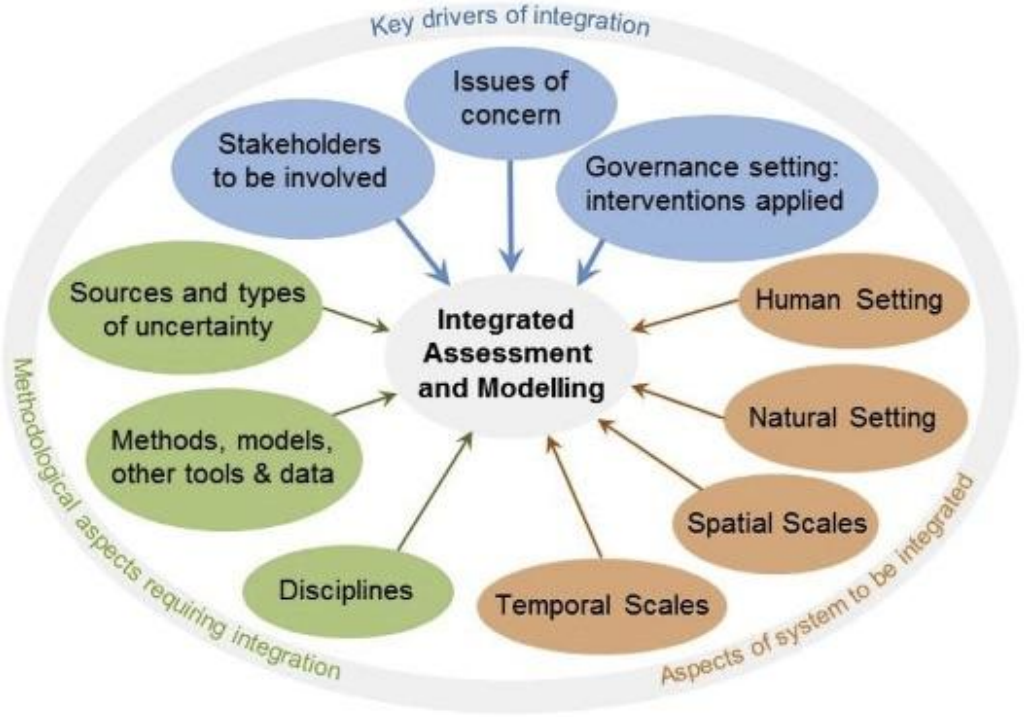


Fig. 1. The ten dimensions of integration in integrated assessment and modelling (IAM).

Katie: Interested in targeted sampling not necessarily to get observations where model is uncertain but to target areas where model actually fails. For example, going from coarse scales to microscales. Can we take target samples at this intersection to aid integration of different models?

Common areas of interest:

1. Actively seeking data to improve the modeling activity (linking to the sensor-based data collection and integration group, but not just being passive consumers of their output)
2. Accelerating data ingestion, calibration, and integration of models- tools to assist scientists with these processes (eg rescaling, finding data, mapping/coupling models)
3. Managing simulation results - capturing workflow, tracking provenance of many runs (eg 40K runs), characterizing simulations, repositories of select simulation products

Target GEO area mid- and long-term: human-natural model coupling

- Lots of reports describing this, it is also in the NSG GEO AC document

- Funding opportunities: FEWS programs, (eg NSF INFEWS), etc.
- Incredible levels of interest in the community
    - By GEO people not at this meeting
    - Texas Mesonet - statewide sensor networks (e.g. humidity and temperature) - meeting in a few days

Short-term GEO focus for initial prototyping/demonstrations: soil moisture and water runoff

- Understanding recharge of subsurface water systems
- Involves combining long-term weather (link with atmospheric) and short-term weather forecasts
- Merging models and observed data, which is of great interest
- The human side: human responses to flooding forecastings, or real-time farm planning, waste management

Short-term IS focus: TBD, but soil moisture and water problem covers all 3 aspects of simulations identified above by the group

Lead for the group: Daniel Fuka

Sensor-based data collection & integration Group

Daniel G., Jie, Enrique, Hanan, Craig

- Each group discusses common topics of interest and defines one or more of these based on individual's interest:
    - Short term prototyping efforts
        - Rapid alert system: Now-cast: Short-term prediction of heavy rain/flooding within the next 30 minutes and send a notification to the city manager for the area
            - Pressure, temperature, humidity, wind,
            - Water vapor
                - (radio sondes)
                - GPS (calculated from the error in the GPS)
            - Apply some logic to these values to predict green, yellow, red for severe thunderstorm
            - Start with a retrospective analysis
            - Explanation of prediction
    - Midterm proposal-size activities
        - Develop a general framework for rapidly integrating a variety of sensors and pull them together on the fly for various types of rapid alerts
            - Severe thunderstorms/flooding
            - Tsunamis due to earthquake
                - Network of buoys

        - Rapid integration of multiple sensors, dealing with noise & uncertainty
        - Workflows that trigger certain types of analysis based on features
        - Learn to update the model from past event predictions
        - Explanation of data/model of a specific prediction
    - Long-term driving vision
        - Learn betters models of these types of efforts
        - Allow citizens to build their own warning systems
        - Allow scientists to understand the underlying mechanisms and models
- Each group should consider setting up an IS-GEO Working Group:
    - Identify lead(s): Craig & Enrique
    - Report to the larger group
    - Prepare a brief description to publicise on IS-GEO web site
    - Set up group telecons/webinars and involve students and others signed up to IS-GEO

Each of the 3 WGs will follow up with regular telecons.

We will use the wiki for documenting datasets, people's relevant papers, etc

- In people's individual pages in the wiki
- New pages for each of the WGs
- A page for each dataset
    - Daniel will add guidelines about how to cite/share data

Monthly IS-GEO calls as follows:

- 20 min technical presentation
- 10 min student presentation
- 10 min report/discussion from each of the 3 WGs, exploring points of interest

Education:

- Proposal (Yulia)

EDUCATION SESSION DISCUSSION NOTES

Various courses are being taught by the IS-GEO community members (list from Imme)

Levels:
Undergraduate and Graduate students

Pahdraic - various opportunities, including:

- NRT training grant ML and Physical sciences, ~20 PhD students with co-advisors, team science component in proposal,

- NASA minority with Cal-State LA, encouraging students to go on to advanced research/study (running in August)
- Issues and problems - basic skills in computing not a given, different levels need to be careful not to lose people in math concepts, prerequisites, asymmetry in the student disciplines (e.g. computer science students go to internships, geos very interesting).
- PhD students teach short courses, grad students tend to engage first and then faculty get interested
- Use python and R and jupyter notebooks
- Sustainability - hard to maintain interest at department levels, not easy to resolve because happening on margins in departments and often depends on individual people
- Question: Can these courses, once designed, be taught later by others either lecturers or doctoral students? ...Answer: Probably possible to have other people teach...requires development of course curricula and materials by upper level researcher and then can be taught by others pretty easily because pretty intro to intermediate level material
- Question: are you happy to share classroom resources? Answer: Yes most of it is up on Github….Pahdraic - willing to share resources and materials

Sai - Two Courses at MIT

- Lots of good infromation about the two courses (see slides)
- These slides will be available online and can be linked to by the IS-GEO community
- Lecture level is aimed at enabling students to use techniques and helping them identify links between knowledge domains where ML draws on concepts from other areas.

Yolanda - Course on “Introduction to Computational Thinking in Data Science” (datascience4all.org)

- Slides:https://www.dropbox.com/s/3ncw9fq82l7mfua/Gil-IntroDataScienceCourse.pptx?dl=0
- Papers about the design of the course: http://datascience4all.org/publications/
- Syllabus: https://www.dropbox.com/s/bfwyuax6j6csl3b/Course-DataScienceForNonProgrammers-Syllabus.pdf?dl=0

NEEDS/OPPORTUNITIES:

- Listing with descriptions, example syllabi, course materials,
- Online video/webinar courses for Prerequisites in IS-GEO (e.g. Introduction to R, python, Jupyter notebooks)

- Finding sustainable models for the courses and materials