

Summary of breakout groups as well as general discussion on Institutional Design for GEOSHARE (see below for detailed comments from each breakout group)

Q1: How can a public-goods project like GEOSHARE be sustained over the long run?

While GEOSHARE has drawn inspiration from the GTAP project, which has provided global public goods for the global trade and environmental policy analysis community for the past 20+ years, it differs from GTAP in important ways. Perhaps most importantly, as many participants pointed out, the GTAP data base is sold to non-contributors. Over time, this has become a key part of the revenue stream. Because GEOSHARE data will be freely available, data base sales are not an option for additional revenue. This led to discussions about a number of other potential revenue streams, including: the sale of consultancy services from various communities-of-practice with GEOSHARE tied to the data and tools on the HUB's, specialized courses, the sale of data handling/curation services, establishing an endowment to which private companies and individuals could contribute, and selling spots for logos/marketing. It is difficult to assess the potential for these alternative revenue streams, but clearly experiences from other organizations would be valuable.

There was some concern expressed about emphasizing the public goods nature of the project, as this may repel certain stakeholders – particularly those from the private sector. **An alternative description would be that GEOSHARE seeks to equip communities of practice with endorsed data to support best-available science for a wide range of applications and ensure all scientists and organizations who contribute datasets or who work on designing or sharing workflows get adequate credit.**

The core business model as proposed in the workshop involves formation of an advisory board comprising representatives from key stakeholder institutions. Each board member would contribute an annual fee, and, when combined, these would fund core activities, including maintenance of the HUB and administration of the project. Based on broader conversations at the workshop, it appears that there may be significant incentives for participation on the Board early on, by institutions interested in shaping the project/moving it in their preferred direction. This may prove sufficient for getting the initial group involved. Over time, this incentive will likely diminish. In its place a natural solution may be to have a period of internal release and testing for key data and models. The length of this period could be adjusted to ensure that those organizations which are heavily reliant on the data for decision making have an incentive to join the Board. Others, for whom GEOSHARE is less critical, may prefer to wait and, in so doing 'free ride'. It appears that there may also be some incentive to be on the Board in order to facilitate communication and networking with the nodes and other Board members.

A final point which came up several times is the importance of having some early successes! Being able to demonstrate the value delivered by GEOSHARE will be important to attracting early Board members. All donors, including those from the private sector, will require clear and achievable milestones which demonstrate that this is not simply another open-ended 'research project'.

Q2: What are the appropriate roles for the Advisory Board?

There was considerable debate about whether the Board would be an Advisory Board or a Governing Board. Those with extensive experience with other institutions seemed to lean towards starting this as an Advisory Board while the community takes shape. Thus the initial GEOSHARE leadership would be drawn from the group currently championing the project, as opposed to being hired by the Board. The

Advisory Board would, however, be responsible for funding the core activities and, in so doing, would guide the broad direction, looking ahead to anticipate the future needs of decision makers. (If GEOSHARE doesn't follow their interests, funding will be terminated!) The Board would bring an understanding of GEOSHARE's niche in the broader institutional landscape, identifying gaps to address, and approving new nodes – including those representing important user communities. Finally, the Board could play a role in endorsing particular datasets and tools as being fit for particular purposes.

Q3: What are the incentives for participation in GEOSHARE, as viewed from the perspective of: node leaders, government data generators, private industry, academics, other data consortia and user communities?

Q1 addressed the question of incentives for Board members, and hence long run financial viability, but it did not address incentives for others involved in this activity. Participants discussed this issue extensively. Node leaders would benefit by the value brought by being recognized as a leader in their field, and by the contacts and influence which accompany serving on the GEOSHARE Science Committee. As currently proposed, the Project would only pay for their travel to the annual meeting, along with a modest honorarium for the time devoted to this meeting. (Note: An important point made during the meeting was that the node 'leaders' perhaps should really be representatives of broad communities-of-practice. It would be a mistake to anoint a single individual as the 'top individual' in a given field. This could create unnecessary conflict. Indeed the project would benefit from a rotation on the Science Committee, with node leaders turning over after a fixed period of time.)

For users – easy access to free data is potentially a strong motivator. Contributors are more difficult. A great deal of emphasis was given to citations, and HUBZero is well set-up to facilitate these. Also, by plugging one's own work into an existing workflow on the HUB and making sharing easier within / between communities-of-practice, contributors could also find participation attractive. GEOSHARE could also be an important clearinghouse for hiring and consultancies. The peer-review aspect was an attractive option. The ability to satisfy journal or grant agency requirements to make project data publicly available might be met by "publication" on GEOSHARE and might make GEOSHARE Advisory Board participation attractive to some additional groups/funders. It was suggested that GEOSHARE should hold strategic discussions with the CGIAR to consider having GEOSHARE play a central role in the CGIAR open data initiative and geospatial data effort. This suggests that maybe more groups in this area of geospatial food/agriculture data are currently in need of a "ready-to-roll" open data platform and these should be sought out.

In terms of government agencies and data providers, as well as the private sector, it appears that there is a niche available for datasets which are clearly high quality, reproducible (it was pointed out that most geospatial data are actually outputs from other models!) and endorsed by the Science Committee and the Board. Offering something akin to the 'Good Housekeeping Seal of Approval' or the 'Gold Standard' for geospatial data related to food, agriculture and the environment was viewed as a valuable role for GEOSHARE.

Another potential incentive, both for membership on the Science Committee and the Advisory Board, is that being able to represent that one's organization is part of a global community of practice with access to global expertise and international recognition via GEOSHARE may improve funding application respectability.

Q4: How are data and workflow priorities defined?

Participants felt that the Advisory Board should prioritize establishment of new nodes, while the nodes are likely better positioned to establish workflows, which could then be approved by the Board. Clearly coordination between these two entities is required, and this is the role of the Science Coordinator and the Executive Director.

-----Detailed Rapporteur notes follow -----

Institutional Design Breakout: Rapporteur Hermann Lotze-Campen

- Q1: How can a public-goods project like GEOSHARE be sustained over the long run?
 - Geoshare is unlike GTAP w.r.t. financing
 - Examples from other initiatives are needed
 - Role of private sector tbd., e.g. through ads/logos
 - Each board member provides certain budget amount (incl. early access to data)
 - Format and content of data to be defined (very diverse in Geoshare)
 - Various community –of-practice "consultancies" could be run on the side to make use of data/workflows
 - Direct contributors may not be the biggest beneficiaries
 - Provision of "primary datasets" could be interesting
 - Who are the "heads of nodes" on the Board? (potentially very large group)
- Q2: What are the appropriate roles for the Advisory Board?
 - Paying for core activities –
 - Give advice! On relevance, direction, priorities....; Should be separate from executive role
 - An "Executive Director Search Committee" should be installed
- Q3: What are the incentives for participation in GEOSHARE, as viewed from the perspective of: node leaders, government data generators, private industry, academics, other data consortia and user communities?
 - For users: access to free data, with automated documentation, reproducibility, endorsement, easy access, etc.
 - For contributors: less clear; e.g. awards could be given
 - We need to look for incentive-compatibility

- Potentially important: Increase usability and citations of data products which are produced decentrally all the time by many research groups
- Advantages for early movers/contributors should be stressed.
- Q4: How are data and work flow priorities defined?
 - Prioritising new nodes should be done by advisory board
 - Nodes should propose operational work flows etc. to advisory board for approval
 - Coordination between AB and SC is needed

Institutional Design Breakout: Rapporteur Jim Jones

Question 1. Sustainability

- There is a need for good, quick successes (early wins along with success stories being advertised)
- There is a need to make GEOSHARE work well in developing countries, using local data, giving credit and engaging locals throughout the process
- GEOSHARE should clearly define and articulate its Global Public Goods (e.g., the Hub itself, data, etc.)
- Perhaps GEOSHARE could get institutions to pay to support data storage as a fund raising fee. Some discussion on negative aspects of this as well as positive
- GEOSHARE needs an analytical core (e.g., models, visualization tools, statistical tools, ...)
- GEOSHARE should consider giving national agencies “nodes” in different countries with central support; sustainability will be enhanced if countries feel ownership for their part of GEOSHARE.
- GEOSHARE should hold strategic discussions with the CGIAR to consider GEOSHARE playing a key role in the CGIAR open data initiative and geospatial data effort.

Question 2. Roles of Advisory Board.

- It should be an Advisory Board (AB), NOT a Governing Board.
- The AB should help GEOSHARE clearly define and focus on key capabilities (e.g., focus on nutrition)
- The AB should help GEOSHARE be cognizant of other systems “out there”, and develop strategies for partnering when appropriate, but also continually reviewing competitive advantages, niches, and complementarities.

Question 3. Incentives for participation

- Ability of members to shape the focus and evolution of GEOSHARE
- There was some worry about industry competition, but clear opportunities to work with industry to provide data for broad use beyond industry usages
- GEOSHARE would provide a platform that agencies, the public, and researchers would trust relative to the quality of data and capabilities, GEOSHARE should strive to establish high quality standards and aim to build trust from the start. This could give GEOSHARE a very positive brand name (e.g., similar to the Good Housekeeping Seal of Approval)
- Setting standards for data, workflows, interfaces, and tools would provide a platform for participation by all who wish to contribute (in contrast to top down approaches that tend to flounder).

- Business incentives are needed and could focus on sustainability, food and nutrition security, climate change, water, etc., thus giving participants capabilities that are difficult to achieve when one has to assemble all of the data, quality check it, and build work flows from the ground up (e.g., more efficient and high quality assessments of major issues)
- For company support, GEOSHARE needs to articulate benefits, timetables over short and long term, payoffs on investments. To do this, GEOSHARE should seek guidance from the private sector
- GEOSHARE could provide a source for hiring and consultancies for public and private institutions and projects
- Benefits could include access to training for members, and early access to datasets and new capabilities
- There needs to be a formal review of GEOSHARE to report to donors to give them confidence in the overall effort and its nodes (provide high level of accountability).

Question 4. Design of workflows, databases

- GEOSHARE needs to identify key geospatial data needed for geospatial analyses of food and nutrition security, climate change, sustainability, etc., and work with partners (e.g., in CIMSANS, AgMIP, CGIAR, and others)
- Geoshare should also establish an “open source” process for version control of its data, models, and analytical tools to help participants find easy ways to contribute to its high standard evolution.

Institutional Design Breakout: Rapporteur Ron Sands

Q1: How can a public-goods project like GEOSHARE be sustained over the long run?

- GEOSHARE provides a community to address shared needs
- Value Added
 - Convert raw data into something useful
 - Provide a product that is in demand by user community and donors
- Improve data product over time, especially with updates
- Home institution ensures data are up-to-date and have robust quality checks (Gold Standard)
- Provide a unique data product
- Flexibility to shift direction as needed
- Concerns
 - Not likely to find a central source of funds
 - How do nodes get funded?
 - Governments may demand open access

Q2: What are the appropriate roles for the Advisory Board?

- Ensure financial stability of GEOSHARE
- Ensure transparency of process
- Ensure a process to create the “Gold Standard”
- Re-evaluate governance structure, especially during first few years
- Be forward-looking – anticipate needs in advance

Q3: What are the incentives for participation in GEOSHARE?

- Node leaders
 - Interaction with other scientists
 - Participation in a global community
 - GEOSHARE provides a mechanism to coordinate multi-disciplinary research
- Government data generators (e.g., ERS and NASS)
 - Already have a goal to see wide use of data
 - Incentives line up well
- Private industry
 - Public data can help with private research
 - Mechanism to share at least some data

Q4: (ran out of time)

Institutional design Break out: Paul Hendley’s notes (Ron Sands was rapporteur)

Q1 - GEOSHARE sustainability

Strength of GEOSHARE is the endorsement /data quality thinking

While there are more external funds available for food related programs - they all require an advocacy element.

All proposals must include elements of data updates and highlight how the underlying data will fit against the silver/gold/platinum dataset concept

Models need to be forward looking and suggest how needs may transition due to likely future thinking

Need better benefits for contributors of datasets

Communities of practice are critical and they need to work as a team to endorse on another's proposals and thus add value by credibility by consensus - this nodal funding issue needs to be addressed in CIMSANS scoping document

Are we using the medicine analogy enough - the analogy of aggregating clinical trial data to produce a public good of more value than the individual company trials

Wariness over the terminology about public goods since this may scare several types of stakeholders - maybe we need to stress **this is about equipping communities of practice with endorsed data to support best-available science for a wide range of applications and** ensure all scientists who contribute datasets or who work on designing or sharing workflows get adequate credit

Q2 - Advisory board role

Needs to be highly proactive. Financial responsibility - ensuring total transparency and gold standard approach for the science committee

Ensure best learning from GTAP is brought in

Q3 - Incentives for participation

Not money really - much more important is collaboration

Some saw that the advantages were in having a multi-disciplinary community available. Also recognition and linkages to global expertise and regularization of approaches

Private companies have the various motivations described by CIMSANS as well as the increasing driver of meeting sustainability standards (e.g. Walmart)

The CGIAR-CSI needs a system like GEOSHARE to achieve its recent commitment to open data and more access to the global community - wants some quick wins in terms of datasets and modeling and feels the case could be compelling.. Useful links for CGIAR who are not great at disseminating models

John I. - he would like to see more attention on temporal dimension, livestock, dairy, fish and coastal zone issues as well as support to better understand land tenure arrangements around the global regions

He mentioned mapping food-sheds and informal food transport as well as financial access - a key will be identifying information needs and spatial scales of different workflows.

He has suggested an alternative admin structure (slight changes in emphasis) and he STRONGLY recommends a TEMPERATE region node

General points

What's in it for CSI - what would a user case look like - GEOSHARE has to make that one work

GEOSHARE needs to keep stressing the importance of the data curation and endorsement center role

Who are the communities of practice - how can we develop them / add to their number - I see this as pretty key - GEOSHARE needs to discuss with the nodes how to help the new approach [Paul Hendley] to getting funding directly at Node level make their proposals more powerful

THOUGHT - should the nodes be communities of practice - appoint a leader but do not stress the "individuals as world leaders" - this might be a strategic bonus given the change in fund seeking approach since it might induce more cooperative action with each CoP

Who could GEOSHARE identify as a source of funding for Navin to fund a Masters student to write a workflow in Hub Zero to allow frequent rerunning of his famous (1999) dataset with the latest data available.

Similarly we need to consider how we could propose multi-purpose spatial aggregation/disaggregation tools to support the harmonogram concept. My feeling is that if we design a tool approach to disaggregation /aggregation at the same time others can be agreeing what data need to be harmonized – would involve GEOSHARE/IFPRI/Navin

GEOSHARE will need to discuss the issue of "lowering entry barriers " which caused some concerns among some potential stakeholders i.e. the concern that "idiots will misinterpret my data "

GEOSHARE needs to press the "providing a new approach to peer review" - community credibility by consensus for datasets being used in new publications

Encourage GEOSHARE to work a citation tracking scheme into auto-metadata approach

Explore how to bring the GEOSHARE thinking alongside data.gov approach - how to fund??

Workflows thinking makes the inclusion of access to latest versions of PRIMARY datasets key - e.g. who owns the updating in Admin boundaries (IFPRI with open access via GEOSHARE) – does member ship of GEOSHARE actually save them FTE's for making data available and tools for aggregating /disaggregating and dissemination

Challenge Carol and Hub Zero team to cannibalize other Hubs for key tools - e.g. temporal data handling tools

Institutional Design Breakout: Rapporteur Kate Schneider

Key takeaways:

- Sustainability: initial influx of money may be necessary to prove the value, but advocacy/marketing would be helpful in finding the eventual business model
- Advisory board: dependent on the funding model
- Incentives: professional prestige, advancement – but this is more challenging for some disciplines than others
- Priorities: needs input from the scientists, but also connection at the higher level to big priorities/questions relevant to policy

Detailed notes:

1. Long-term sustainability
 - a. Business, institutional, funding models – all of this is what needs to be determined
 - b. Core funding – cyber-infrastructure, people – this is the minimum needed, what does it cost? \$400-500k/year
 - c. GTAP comparison?
 - i. Pay model isn't on the table for this, open source is different so is this a relevant comparison?
 - ii. GTAP has a diversity of funding streams, but that's where the similarities end
 - d. How do you convince someone (e.g. a university) to pay if they can get access to everything for free?
 - e. An influx of initial core funding recommended – and then the long-term sustainability becomes a question out of that
 - f. What about private sector funding? They might see a lot more utility
 - g. How do you create a mix across types of sponsors (e.g. universities vs private sector or private donor)
 - h. What about organizations with open data policies? Is this an incentive to contribute?
 - i. Dual-licensing option could be an option
 - j. Another example could be the google/weather underground model – a certain amount of data is free but heavy users have to pay
 - k. There's always a cost to open – maybe a membership structure is an option
 - l. What about the “holy grail” of an endowment? In the climate area there are lots of private companies that might be interested in funding, just to “do something good”
2. Roles of the advisory board
 - a. Guide the scientific course – initial proposal
 - b. Responsibilities are related to the funding model
 - c. AgMIP teams interested in geospatial analysis is only about 20% of AgMIP
 - i. These kind of users would be interested in guiding the scientific directions
 - ii. These people would have a role in the notes – but then certain institutions would want to be members or influence discussions bilaterally as well (e.g. IIASA, PIK)
 - d. Appoint personnel
 - e. Scientific committee – should recommend new notes to the advisory board
 - f. What about advisory board recommending needed analysis?
 - g. Identifying sources of funding – board needs to back up the notes in finding funding
3. What are the incentives to participate?
 - a. Will provide a common set of tools to validate, aggregate data – output analysis
 - b. Others using your data enhances your reputation
 - c. Opportunity to show leadership in a particular area, leading a part of a community – professional prestige
 - d. Building your network
 - e. Node leader – is this label explicit enough for professional advancement? (e.g. would promotion committees recognize its importance?)
 - f. Weight that alternative (non-journal publication) citations carry depends on the discipline – evolving rapidly but for some publication of a dataset in a journal article is still the only option to get a citation for data
 - g. Does Geoshare need to include a suggested citation? Hubzero is very particular about it

4. How are the data and workflow priorities defined?
 - a. Need to be able to fit with general procedures around the world but also need to have flexibility
 - b. Would we need mandates? Or could individuals find money to do the work they are interested in?
 - c. What about someone with money already who wants to use Geoshare to implement?
 - d. If there's broad demand
 - e. Money for the software side is one part of it, but also need to fund researchers for some things
 - f. What about deciding where to put money that is already on the table
-

Institutional Design suggestions by John Ingram: (see also schematic on following page)

Advisory Board

- To give strategic advice on priorities and broad-level user needs
- Comprises broad stakeholder community, including donors (in their personal capacity) and Executive Director (ex officio)
- Chaired by independent stakeholder
- Managing Director to be Secretary
- Meets annually

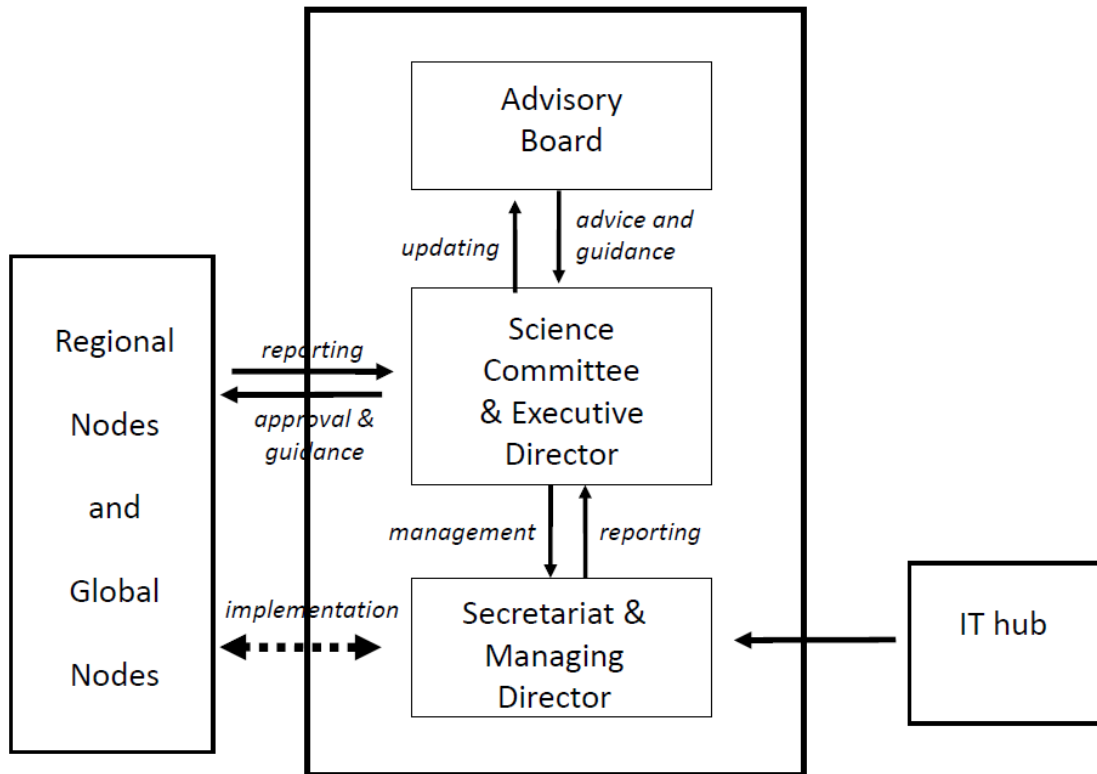
Science Committee (not 'scientific' Committee to make the point is discusses science and hence can have 'non-scientists' on it)

- To agree science direction, priorities, activities and viability, based-on (but not dictated by) advice from Advisory Board
- Comprises Rep from each Node and IT Hub (to ensure all Nodes are in touch and communally agree agenda) and some stakeholders
- Chaired by Executive Director (at least half-time)
- Managing Director to be Secretary
- Updates Advisory Board (members attending as Invitees)
- Meets twice a year (once immediately following each Advisory Board meeting, and six months later)

Secretariat

- Ensure the implementation of plans agreed by the Science Committee.
- Comprises full-time Managing Director and at least 1 FTE Science Officer and 1 FTE Admin
- Run by Managing Director

Proposed GEOSHARE management and reporting structure



Provided courtesy of John Ingram