

June 8 CA LCC Informatice Workshop Summary of Recommendations

INFORMATION TRANSFER RECOMMENDATIONS

(Ranked
by Group)

CATALOG

15

Define user and needs. Who is the intended audience?

Use existing structure as foundation (e.g., CalAtlas). Have methods for permission/access, etc. and enhancing
Should LCC develop all encompassing portal?

Use existing "gap inventories to identify data to fund - don't duplicate existing efforts. Awareness of other data
collection partnerships and activities (e.g., JFSP Fire Consortia, GBIF, etc.)

Support existing catalog / portal sites rather than creating a new one

PARTNER CAPABILITIES

6

Identify hosting agencies and convene them to discuss options for federation standards (e.g., maintaining catalog)

MAP SERVICES

5

Develop user requirements for Map Services (e.g., cost structure, technology)

Facilitate data publication through map services to ensure availability to a wide audience

Before developing tools, make sure there is a user.

Provide Map services and downloadable data (e.g., list of services, facilitate access to...)

With metadata, include link to get the data

Interview representative users for requirements and provide systematic way to provide feedback for map services

Provide reliable authoritative consumable map services

Identify potential providers

AGREEMENTS

3

Agreements/MOU needed for whoever hosts web map service. Identify requirements/possible hosts, if second
party hosting what kind of agreement. How to provide the service?

MOU with partners to set foundation for data sharing. (Existing LCC charter does not incl. data sharing)

DATA ACCESS

What can be confidential? Legal issue (e.g., PRA, FOIA). Would be helpful if there was a webinar (targeted to data managers and contributors) explaining realities of PRA/FOIA to agency folks as it relates to data (e.g. URISA?)

ADDITIONAL RECOMMENDATIONS

DATA AMBASSADOR

Outreach to user groups for use of LCC generated data (Data ambassador)

ID existing catalogs, data sets, etc. (data ambassador)

Have ambassador register others data sets, build on existing data sets, building longer term datasets, multi user data sets, more useful metadata

Provide metadata training for LCC funded projects

ID important framework layers for the LCC (esp. authoritative data), find existing inventories for each framework

Develop tagging scheme for LCC-relevant data (registry - global change master directory)

Promote / prioritize filling of data gaps, which may include funding data collection

BEST PRACTICES

Best Practices for data management are made available (LCC provides examples)

Develop / adopt / modify best practices documents. Generic or cross agency to support improving data stewardship

Create or look for best practices for data management / data stewardship and make them available

Find and promote examples of what success looks like for data management

Create standard data delivery language for contracts (metadata, controlled vocabulary, data access)

Develop/adopt/modify best practices document/guidelines for data stewardship to meet LCC partner conservation goals

Require a data management plan for LCC funded projects like NSF / DataOne tool (

<https://bitbucket.org/dmptool/main/wiki/Home>)

PROVIDE LEADERSHIP/COORDINATION

LCC informatics priority is given to joining key datasets (ie GBIF, CalAtlas, Joint Venture project tracking etc.)

Develop communications and implementation strategy for data management for the LCC partners (as part of larger LCC strategy)

Coordination between LCC and CSC data management projects

Will LCC identify and promote authoritative (best available) data sets?

LCC should promote and prioritize filling of data gaps

ID specific objectives for areas of emphasis - prioritize - ID key management questions

Data sharing policy for LCC funded projects

Develop a data sharing policy for LCC and require projects adhere to it

LCC informatics framework across the nation is needed

CSC & LCC need formal collaboration strategy for informatics

Expand existing LCC portal for LCC funded projects

Identify authoritative data sources for different disciplines...make sure key datasets are in the most important places e.g., CalAtlas, GBIF, Joint Venture, etc.

Define authoritative dataset

Adopt upcoming metadata standards and provide training to LCC funded projects. Add controlled vocabularies.

Data Stewardship: LCC should encourage longer term data sets including what those datasets might cover

LCC should require funded projects to identify data life cycle and maintenance

Add functionality of existing LCC website that gives it more of a portal role

Make archiving retention of data part of the data sharing / management policy for LCC

Develop a database of different partners on the ground

California Landscape Conservation Cooperative (CA LCC) Informatics Workshop

June 8, 2011

Modoc Hall, Sacramento State University

Summary Notes

Informatics: Of or relating to information, facilitating information acquisition, interpretation, translation, exchange and availability

Objectives for workshop in regard to informatics:

- Issues clarified, classified, prioritized, recommendations (long and short term) made.
- Identify existing resources
- Create short term and long term recommendations
- Discuss options for how LCC can move forward in this area
- Identify a few key next steps in the short term, e.g. work for one person for one year.
- Provide a list of actionable items.

What is going on with CA LCC- Rebecca Fris

LCCs were envisioned to address the following issues

The lack of coordination of science managers

Lack communication network to coordinate and exchange.

The California LCC includes part of Baja California in Mexico. There are 21 LCCs in the country – 4 in CA – California LCC, North Pacific LCC, Desert LCC and Great Basin LCC

The primary goals for the CA LCC have been developed by the CA LCC steering committee. Needs have also been heard through a series of workshops in various geographic locations within the LCC.

The CA LCC is a science / management partnership. The 5 goals are:

- Fosters collaboration and integration of science and management
- Supports research, monitoring, and the development of technical products to inform and enhance conservation decisions and actions
- Facilitates information acquisition, interpretation, translation, exchange and availability
- Promotes adaptive management strategies
- Communicates information, findings, activities, and opportunities within and outside the LCC community

Question: Do we have partners in Mexico as part of the LCC? There are informal connections at this point, trying to formalize connections.

There are 11 multi-partner projects funded in 2010 that are ongoing and 13 projects for 2011. Additional information can be found under Projects at CaliforniaLCC.org

The CA LCC Informatics working group is envisioned to:

- Identify existing resources,
- Bring together existing data
- Provide guidance for future data collection and data use
- Provide users with greater access

Comment: LCC not intended to compete with or replace current agency programs or partner activities. The CA LCC is here to help current agencies by coming together to be mutually supporting. Many partnerships exist already.

Question: Are LCCs working individually or are they interconnected?

There is a national network – national meetings are being held. LCC Working Groups are supposed to do work in between meetings; there is a national data group. Rebecca Fris is tracking it. CA LCC is one of the more developed LCCs and may come up with the guidelines for the rest of the county to follow. There are data standards for Climate Science Centers that may be useful to LCCs. They are based on the NSF standards.

CA LCC Emphasis areas (not all inclusive):

1. Climate Change
2. Habitat Connectivity
3. Water Resource Issues
4. Species and habitat information
5. Scenario Planning (identify a future scenario under climate change and the impact on a species or habitat in that scenario).

How can the LCC add value and support data managers and their organizations and whatever their missions are without duplicating what already exists? The LCC would like to see where there are opportunities to share data and accomplish tasks that are greater than the sum of its parts.

Description of workshop: Tom Lupo

Three goals were identified for discussion from the outcomes of the previous workshops.

Goal 1 – Provide guidance and standards for collecting, documenting, and sharing data (Pat Lineback/Kaylene Keller)

Goal 2: Synthesize existing data, assess data gaps, and identify priorities for filling gaps (Tom Lupo/David Harris)

Goal 3: Facilitating data transfer and presentation (Steve Goldman/ Jennifer Carlino)

For each goal there will be facilitated discussion: What data sets, tools, experience, examples do participants have related to the three priority areas. Review and define what each priority means.

Goal 1: Provide guidance and standards for collecting, documenting, and sharing data

The Life Cycle for Data Management: Federal Geographic Data Committee (FGDC)

(The following provided by Pat Lineback)

A. Define User Requirements

1. *Identify business needs*
2. *Identify agency or program requirements*
3. *Define data & analytic needs*
4. *Define existing or new data standards*
5. *Evaluate past data uses for programs and projects*
6. *Any privacy act or sensitive data?*
7. *Define & develop QA/QC measures*

B. Inventory & evaluate data assets and data gaps as they relate to business-driven user needs

1. *Maintain inventory list of data*
2. *Develop AOI*
3. *Locate viable resources for required data*
4. *ID internal resources that may need conversion*
5. *Review dataset descriptions, metadata records, domains, attributes, etc.*
6. *Review accuracy issues and tiers of accuracy*
7. *Review date ranges and data currency*
8. *Review other data limitations (e.g. public domain, contractual, sensitive)*
9. *ID closely matched datasets*
10. *Recommend datasets that meet the specified requirements*

C. Collect and convert geospatial data

1. *Establish a plan to acquire data identified in requirements gathering*
2. *Collect, modify, or develop data through various efforts including vendor or agency developed*
3. *Ensure that requirements & standards are being met (e.g. metadata, standard schema)*

D. Make data produced known and retrievable to the community through documentation and discovery mechanisms

1. *Publish – determine where to publish the data or offer services*
2. *Disseminate – consider data organization, perform records management review, evaluate data steward completeness, 508 compliance, user access levels, etc.*

E. Maintain – implement processes and procedures to ensure that the data meet business requirements.

1. *Develop and implement a maintenance plan for data and services*
2. *Determine and implement records requirements and retention schedule*

3. *Create a data QA/QC plan for the dataset*
4. *Develop a strategy for updating the process of maintaining the data*

F. Complete an ongoing assessment, validation, and potential enhancement of data to meet user needs and business requirements

1. *Review requirements to see that use of the asset meets the requirements and intended outcomes.*
2. *Validate and verify that the data continues to meet business requirements and user needs.*
3. *Evaluate the data as a part of common services across lines of business.*
4. *Evaluate how resources are being utilized and tracked as the data are used.*
5. *Determine the availability or development of user groups to share tools, id and develop services, encourage partnerships and resource sharing, encourage feedback between user and producer of the data*

G. Archive – Required retention of data and the data’s retirement into long-term storage

1. *Identify business needs for archiving*
2. *Develop and implement a written policy for data disposal in accordance with requirements*

Pat Lineback: What does good data management look like, from all agencies’ viewpoint? Let’s lay down what we need in a data management process and let the managers ensure there will be buy in. It is important that whatever is developed is useful and used.

Meeting participants identified issues and recommendations for each category.

(**** Denotes possible Action Item for this year)

A. Define user requirements (needs, standards, what is available?, sensitive?, QA/QC)

Issues:

1. User requirements are not clear.
2. Program data needs and data specification attributes needed to create requirements (technical person, requirement person) need two different skill sets that may not cohabitate in same person. (DH)
3. Who are the users and what is their access to the funded projects and data?
4. Scope is huge, many users many needs. (DC)
5. Data needs can change when researcher gets to survey location, e.g. problems with methods or they find additional data to collect. (DH)
6. Threatened and endangered species locations need to remain confidential.

Recommendations:

1. Develop a data sharing policy for LCC and require projects adhere to it. A data sharing tool. **** [The National Climate Change and Wildlife Science Center is leading an effort to develop a data sharing policy. Greg Gollberg / Sean Finn involved – get a copy of draft version.]
2. Plan on data sets being combined and re-combined. (GB)
3. Find and promote examples of what success looks like. (GB) Projects that used and managed data well. Create or look for best practices for data management / data stewardship and make them available. **** includes data stewardship (more of a long term provider).
4. Catalog of projects and data (use existing systems and structures) (KK)
5. Establish a continuum of data access levels / rights. (GB) Addressing private and sensitive data. Need 7 different sharing levels. 1 – one person access to 7 – totally public. See **Attachment A** for a definition of the 7 levels from Grant Ballard.
6. Focus on goals – for all of scope. Goals get lost in the mix of trying to get other stuff done. Focus on end project
7. Budget time and dollars at end of survey season to ensure data cleanup and to complete metadata. [This should be an allowable item in the LCC RFP process and clearly stated.]
8. Require and budget for rigorous accuracy assessment.
9. Fuzz and remove location data during reporting in single consistent way (JQ) (species locations) Stick to single polygon size like square miles because if it isn't consistent people can parse out specific locations of sensitive data.
10. Require data management plan for LCC projects like NSF/DataOne Data Management Plan (DMPTool) Tool (<https://bitbucket.org/dmptool/main/wiki/Home>) **** (JC) A very simple plan, like data should be public within a certain number of years and should look like X.

B. Inventory & evaluate data assets and data gaps as they relate to business-driven user needs

Issues:

1. Lack of understanding by user communities of data availability and understanding use and limitations of that data. (PL)
2. Lack of methodology to evaluate comparative value of different data sets to allow data portfolio management. (DH)
3. Need clear agency roles & responsibilities articulated for each aspect: discovery / portal / clearinghouse. Make sure this is funded and has continuity.
4. Identify a clear method on how to make data searchable publishing through Google or arcgis.com.
5. Need a formal way or place to record field protocols.(JQ)

Recommendations:

1. Use existing “gap” inventories identify data to fund (KK) – don't duplicate existing efforts. (JC) (***)
2. Identify authoritative data sources for different disciplines. (PL)

People have some sense where to go to find these data. Look into other efforts that are out there, existing systems. Which are most important and worth contributing to? What will become the repository for LCC data? Identify core data sets and ensure that we partner with those people. What are best locations, repositories to put data? Make sure key datasets they should be put in most important places. CalAtlas, GBIF etc. Joint venture (***)

3. “Data in a box”, e.g. water for CA or roads for broad distribution to certain user communities. (PL)
4. Awareness of other data collection partnerships & activities. (JC). Make sure all the pieces are connected, like fire consortiums, GBIF Broadening the scope to include different systems, geographies, etc. Awareness of what is out there that we want to contribute to and use. Add partnership component to data inventory action item. ***
5. Support development at standard data schemas for common data collection efforts instead of having every researcher reinventing the wheel wasting time and eliminating data interoperability.
6. Develop or find data management best practices docs as LCC recommendation .****

C. Collect and convert (geo-spatial & non geo-spatial) data. Establish plan to gather data; convert to standards (metadata, schema, etc.)

Issues:

1. Reluctance to share data before publication. (DC)
2. How to enforce metadata publication? (RL)
3. Force a complete data and metadata set.
4. Expand existing SFBJV and CVJV projects tracking system to provide historical data of conservation activity in support of landscape level conservation planning. (XW)
5. Length of time before releasing data. (RL)
6. Scientists can't agree on protocols / field methods.
7. Project tracking system so there is a historic timeline of conservation data.
8. Can't guarantee data confidentiality due to Public Records Act (PRA) and FOIA (Freedom of Information Act) (maybe a dot org is immune?).
9. Record controlled vocabularies used if not standard. (JQ)

Recommendations:

1. Adopt upcoming metadata standards and promote training within the LCC. Add controlled vocabularies (***) [USGS Biological Informatics Program does metadata training, provides a clearinghouse and offers other metadata services such as metadata creation. Might be an opportunity to collaborate in the metadata arena.] – metadata is relatively easy to do. Provide metadata training for LCC funded projects. [Each year the FGDC supports a grant program to provide seed funding to promote metadata training. Usually opens in the Fall. Might be something to consider for additional funding. Has to be non-feds applying for grants.]
2. Encourage / require use of existing protocols

3. Make existing protocols available.
4. Create standard data delivery language for projects / contracts *** (metadata, controlled vocabulary, data required as deliverables)

Example: 1. Data will be delivered and it will be the property of LCC. 2. Data comes with minimum level of metadata. Language will be created for project managers in the hopes of raising the quality of data delivered e.g. Bonneville Power Administration / Pacific Northwest have this language. UC will not sign away data rights, they will share, but researchers can't be precluded from publishing research. They will share, but not give away. [Language "Upon completion of contracted work, **metadata** information should be updated with data sets at the final data repository/ies." Additional guidance: "All Program funded RME (Research Monitoring and Evaluation) data need to be readily accessible and in an agreed-upon electronic format. RME data will consist of either the raw data, derived data, or summarized performance measure as agreed upon by the project proponent, the funding agency, and Council. RME data, **its metadata** and relevant reports should be available annually, as well as within six months of completing a significant phase of any research project or within six months of project completion."]

5. Establish data sharing policies. (DH)
6. Improve the centralized geodatabase to enable multi-year tracking. (XW)
Ability to handle multi-year conservation data on same location. Improve current data to make it available to the LCC.
7. Improve efficiencies for data collection and dissemination. (PL)
8. Develop standards and system to crosswalk data across different models. (PL)
9. Monitoring network to feed data into other systems and incorporate local and regional monitoring efforts. (PL)

D. Make data produced known and retrievable to the community through documentation and discovery mechanisms

Discussion of this item under Goal 3 on Page 11.

E. Maintain – implement processes and procedures to ensure that the data meet business requirements.

Issues:

1. Few long term datasets exist (long term data is better than one-off things that can't be compared).

Recommendations:

1. Data stewardship: LCC should encourage longer term data sets. (TL) Repeat cycle data like breeding birds survey, gives sense of how things change over time. Including identifying what those would be. ****
2. Define the roles and responsibilities of different types of data stewardship to assure data sustainability. (DH) What does it mean to steward data? What to expect from others and what is expected.
3. Promote use of data / DST (decision support tools) and adaptive conservation framework. (GB) Make sure your data meet the requirements. Go back and ask "did we deliver what was expected?"
4. Develop/ adopt/ modify best practices documents. Generic or cross-agency to support improving data stewardship. **** Then require it for funding. Provide training? Support? So that the data are most useful. We want the data to be well organized so we can use them to meet our goals. Not only to get projects, but show Natural Resources Managers what they are and how to use them.

5. What can LCC do in the next year to encourage buy in? Develop best practices and then sell them. LCC really wants that science to be used and to assist in conservation outcomes. Enforcement of best practices and standards. Marketing. Provide guidance. We need data criteria up front to save both sides time. “Develop best practices documents / guidelines for data stewardship to meet LCC partner conservation goals ***.” Trying to influence beyond LCC. Develop a communications implementation strategy for data management for the LCC partner groups .**** Should nest in overarching communications strategy. Lot of the opportunity for data stewardship to be done well and probably in partnerships. Who does what best?
6. Collaborate with LCCs to avoid redundancy. Action: initiate conversation. Coordination between LCCs and CSCs.

F. Complete an ongoing assessment, validation, and potential enhancement of data to meet user needs and business requirements

Issues:

1. Lack of follow through and resources to keep data current. (PL)

Recommendations:

1. LCC should require funded projects to identify data life cycle and maintenance (PL) ****
2. Promote use of data in adaptive conservation framework. (GB) Promote the use of the data period. There is a lack of awareness of what is available. Need to do outreach, an LCC data steward that knows all about available data and doing outreach to user groups. Easy: Users walked through. Harder: Getting managers aware, also need leadership outreach. Help user communities see the value. Data Ambassador **** [one example may be position in Pacific Northwest]. In process of hiring positions when they are better defined so that the correct expertise is present.
3. Promote development of multi-purpose data sets rather than single use projects.(TL) Should be acknowledged in best practices document, perhaps with example, under category of data collection. Multi-use datasets costs more and are harder to develop. Could possibly be a criteria for selecting project for funding.
4. ACTION: Add functionality to existing LCC website that gives it more of a portal role. ****

G. Archive – Required retention of data and the data’s retirement into long-term storage

Issues:

1. Broad lack of knowledge of records, management and retention schedules. (PL) No written guidelines.
2. Acknowledge human demographics (age / retirements). Causing mass data losses. (TL)
3. If data isn’t in use it, it has little value in context of LCC priorities. (GB)

Recommendations:

1. Develop best practices for archiving / retention and communicate it. (PL)
2. Make this part of data sharing / management policy for LCC (GB)
3. Get data in play or unlock data. (GB)
4. Archival system needs to be tested to make sure data can be released.

Goal 2: Synthesize existing data, assess data gaps, identify priorities for filling gaps

First do analysis of gaps.

Early in first year – look at what exists and what gaps have already been filled.

Need good elevation, good imagery, good parcels, etc. all foundation GIS data.

LCC needs all this too, but a lot is already being worked on by other entities.

What can LCC do to assist ongoing efforts ?

*What efforts are underway and what data is being developed right now? What partners / collaborators can we point to? Climate Science Centers will do foundational science but it will be a very broad scale.

What do we not know? Need outreach from the LCC to advertise the need for data or the goal of the LCC that others may share – how do we get together with the groups who may have this that we don't know about?

We should drill down further in the emphasis areas and define what we are looking for in these areas. As we drill down, some datasets may show up over and over and point the way to what we find most important and then we can prioritize and focus. ACTION ITEM ***** Identify specific objectives for the areas of emphasis – prioritize – ID key management questions.

Climate Change

We can identify many groups that are doing climate change work. A Data Ambassador would coordinate the work at a broader scale to make sure things don't fall through the cracks.

*What do we not know? How do we find the science?

* Find assessments already taking place

Habitat Connectivity

*DFG has a grant to establish standards for connectivity mapping. Will feds comply? Need to have input from a diverse group.

*This varies species to species. Will the LCC be prioritizing species for study? Other people who have done this work have prioritized species on regional basis, want to take a similar approach, but in our sphere what would it take to focus on framework layers e.g. climate, connectivity – LCC wants to be more habitat / ecosystem/ suite of species based. Projects often identify their own target species.

Framework layers

*How is this framework going to be used? What broad categories are important to LCC work (headings) of types of data are important under different headings, like biological, climate, etc. What data categories are important to LCC emphasis areas? Add

a tag like “LCC climate change” to authoritative datasets so they can be searched and define official search terms. ID important framework layers for the LCC (esp. authoritative data) and find existing inventories for each framework..

*ID existing catalogs, data sets, etc. (data ambassador)

*Develop tagging scheme for LCC-relevant data. Have tagging in metadata with a keyword dictionary. Use controlled vocabulary. (Registry – global change master directory)

*Encourage submitted datasets to all clearinghouses involved to use controlled vocabulary, keywords.

Authoritative data sets – defined as data set that can be agreed upon for being the truth. There is not a lot of controversy about these data. Agreed as a single source of truth by rough consensus.

- This is hard with biology. Does LCC want to take any role in deciding which datasets are authoritative? [Species names might be a start. Integrated Taxonomic Information System (IT IS) very active now. If need updates, process much more straightforward than in the past.] Or take them to their steering committee? If you get out of phase with the science how will you know? *Maybe identify “best available”? Authoritative indicates data management, but “best available” doesn’t indicate management in the same way. Does LCC want to have any role in promulgating datasets for authoritative use? Phrasing “Authoritative vs. Best Available” – How does LCC decide what is best available or authoritative?
- Some things lend themselves better to having a single “authoritative dataset” like ‘water rights.’ Some are difficult to do, some can be done. Does LCC want to have a role in that? Is that an important role for LCC? Or is it beyond the scope?
- Authoritative datasets must be freely available, public domain and have a data steward. Have agreed upon standards, available for immediate use, is maintained, steward is contactable, Define what authoritative data set means. What type of data, who is in charge? ****
- Make a list of conflicting data sets (range maps) available for comparison.

Make funding available for filling data gaps. (data development), When we can figure out what the most central data sets are, Funding or coordinating for shared funding, also promoting / prioritizing data gaps. ACTION ITEM: Data ambassador to promote / prioritize filling of data gaps, which may include funding data collection. This would be difficult but there will be some wins, probably especially with the highest priority items.

Develop a database of different partners on the ground

Question: Can LCC do something to get a hold of data that isn't already clean and discoverable? How does LCC get this data, - Data outreach, time to find data, clean it up, figure out if they even have data that is worthwhile.

Gray Data

*Clean up gray datasets to make them valuable.

*Support staff by giving time to be data managers, metadata, cleanup.

Some people are reluctant to share data, do not know how. - * "If you are not sharing your data, its not science" A leadership issue. Also help provide training and sharing tools.

- a. How to get people to share their data? Agencies have real institutional barriers. Not just time, but institutional reluctance, investigators are preparing a document and don't want to release data until internal review is complete. Data may not be complete, or data is pending in scientific research, data not digital (need data rescue), maybe data originator won't get credit, doesn't want data criticized.
- b. Provide profession recognition to people (not just scientists but supervisors) who help get their data out. Also helpful in letting people know what data is out there.

ACTION ITEM – have Data Ambassador register other's data sets. – Not a first year task.

* Build on existing data sets, building longer term datasets, multi-user data sets, more useful metadata.

Some gray data is not worth having. Perhaps the majority of what is worth having is organized?

-Global Change Master Directory – practical use?

-There's a lot of info that's on paper and not digitized in the first place. Lots of data that isn't discoverable.

Goal 3: Facilitating data transfer and presentation (Steve Goldman/ Jennifer Carlino)

Technical – publishing and sharing tools and technologies

Policy – sharing constraints and agreements, crediting sources

- **How should the LCC attempt to catalogue the data?**
- **What are the minimum tools of data exploration the LCC wants to provide?**
- **What data sharing agreements / MOUs would be helpful?**

D. (from Data Life Cycle above) Make data produced known and retrievable to the community through documentation and discovery mechanisms

1. Publish – determine where to publish the data or offer services.
2. Disseminate – consider data organization, perform records management review, evaluate data steward completeness, 508 compliance, user access levels, etc.

Issues:

1. Does a single portal and / or mapping platform meet the needs of the individual projects or whatever kind of collective analysis needs to be done?
2. No one portal will solve all needs.(GD)
3. Keeping data current and authoritative.
4. Provide professional recognition for publishing data.

Recommendations:

1. A reliable authoritative consumable map services ***. PLOS for data? The simpler we can make it for people to consume it and use it the better. Have available data services for multiple users. Keep data with data owners / data creators. Some places have good data but not the staff for upkeep or servers to host data. In that case others should host, ideally with a map service. Good thing about map services is that the data isn't replicated. To increase response time the service can be replicated. Solves the issue of things becoming out of date.
2. Get used to distributed applications / databases. (GB) Try to get people used to the concept that there is not one place to get things. Web 2.0. Not everything lives in one place.
3. Look for crowd sourced solutions. (GB) Creative ways to use social media (iNaturalist) to guide people to where the data are.
4. Identify users' current methods of data access. Give us a good idea of how best to reach them (profile user). (WA)
5. Establish standard protocol for search across federated data repositories. (DH) Tagging. Separating the catalogues from the repositories.
6. Explore connections with data one model / architecture for data dissemination. (JC)
7. Look at other LCC activities others are doing, like the Great Northern LCC Landscape Conservation Mapping and Analysis Program- LCMAP. (Webinar 6/15)
8. Support existing catalog / portal sites rather than creating a new one. ****

Break out groups discussed what they felt were the most important issues from this discussion.

PINK TEAM

- Map services, first having someone write up requirements, including cost structure. Defining technology and cost associated with providing map services.
- An MOU would be needed with whomever would host a web service, probably not a year one activity. Best case scenario, year 2. To define how the service is provided, identify requirements and identify who is suitable to provide this service. Identify potential providers.

BLUE TEAM

- Look at existing structures for hosting (CalAtlas) with different security.
- Having a link to data from metadata, where can it be downloaded.

- MOU with the partners to set the foundation for data sharing and making sure everyone is on the same page. There is a LCC charter which has been approved, but not a MOU at this point.
- Who is the target audience for the map services.
- Provide map Services AND downloadable data.

LAVENDAR TEAM

- Develop / promote a shared subject taxonomy allowing search across existing catalogs.
- Identify hosting agencies and convene them to discuss options for federation (LCC & partners) standards.
- Identify & interview representative users for requirements, provide them with a systematic way to provide feedback on data services.

GOLD TEAM

- Facilitate data publication to map services to provide data to a wide audience.
- Should the LCC provide an all-encompassing data portal?
- What is confidential and what is not? Put together a webinar explaining the realities of PRA and FOIA to data managers. People cannot really shield their data. Educate ourselves and then make it clear to the people who contribute to us. Find, organize and publish existing data on these acts and what they mean.

YELLOW TEAM

- Make sure we are capturing the LCC funded projects and getting them fed into existing structures.
- Before developing tools, make sure a user and their needs are defined.

GREEN TEAM

- Facilitate access to map services on LCC website. Several resources exist for this.
- Build a catalog of relative data layers. Data ambassador finds what is already in the catalogs and adds new content to LCC specific catalog.
- Network of nodes that are all agreeing to standards for data sharing, set up web services and maintaining catalog.
- Think about not differentiating between spatial data and other data. Most places that can do one can do the other pretty well.

The group went through a ranking process for this section. See **Excel file** of recommendations.

Global Issues

Issues:

1. LCC needs to join disparate data sets to succeed (GB). Bring existing data sets in ways that we haven't done before. Data that live in different shops.
2. How to get buy in? (DC)
3. Cost and staff time required to implement recommendations. (DC)

4. LCC and Climate Science Centers need clear information connections and well defined roles and responsibilities. (PL)
5. Climate change and habitat and species relations are required to provide guidelines for data integration. (XW)
6. Informatics groups needs to hear from PIs and other agencies about what is needed.
7. Continue to work closely with Joint Ventures.

Recommendations:

1. Focus on middlewears / schemas / web services. (GB)
2. Develop a plan for selling to management of each agency / organization.
3. Interagency data sharing agreements and MOUs. (LW)
4. National LCC needs a framework for informatics collaboration (PL)
**** Try to influence the national LCC. Get to the larger picture. We don't want to duplicate and work at cross purposes.
5. LCC needs a portal to point to key maps services and clearinghouse. (PL)
6. How can LCC save taxpayers money? (PL)
7. LCC should facilitate best practices for geographic areas to improve collaboration. (PL) CSC & LCC need formal collaboration strategy for informatics. ****
8. USGS should provide more climate model options. (XW) - To be determined is the Climate Service Center role.
9. Close collaboration between Climate Science Centers and LCC in terms of data management and technical infrastructure. (JC) e.g. leverage resources, don't duplicate efforts. Use the Great Northern LCC and Northwest CSC as example. Note: this is related to the email from Greg Gollberg (UIdaho-CSC) and Sean Finn (GNLCC)._
10. **** Also expanded description of LCC projects on LCC website.

Integrate input of the day, draft set of recommendations to the CA LCC, next steps (Tom Lupo) – see Excel spreadsheet

Wrap-up

There is a long list of work for the Data Ambassador to do in the next year. Notes will be to attendees by end of next week (6/17). Then there will be a smaller group discussion to develop a duty statement, then that will be circulated by end of June. The goal to get someone hired which could take 2 – 18 months depending on whose hiring process. It will probably will be six months before someone is sitting at a desk.

Are there things that were captured today that can be contracted or disseminated to another agency to get done before we have a Data Ambassador in place? What can we move forward on?

To move forward, we need approval by the CA LCC Steering Committee. One of our first requests to them can be to approve the hiring of a Data Ambassador. We can start the paperwork now.

Given the long list of recommendations, a smaller group can prioritize and chronicle the initial steps for the Data Ambassador. It may make sense for us to start on these steps with people we already have, and/or the funding doesn't have to come entirely from the CA LCC.

Attachment A - Description of the CADC data access levels (there are 9, technically, not 7), in order from most restricted to least restricted:

Status	Description
RAW	Raw Input – Unreviewed – do not use outside of project
CLEAN	Biologist Reviewed – awaiting project leader approval – do not use outside of project
RESTRICTED	Restricted access - clean and project leader verified data, but must contact PRBO or contributing institution project leader or division director before using. Data is not distributed to other Avian Knowledge Network (AKN) partners automatically.
APPROVED	Available for analysis - internal PRBO or contributing institution use only
AVAILABLE Level 1	Same as APPROVED but also shared with AKN: Data are stored in the AKN's primary data warehouse. The warehouse serves as the primary archive of all AKN data, and no applications connect directly to the warehouse. Instead, data from the warehouse are ported to separate data views created specifically to optimize the performance of an application that connects to it. Data owners can specify how their data can be used in the data views, with the option that their data are not exposed to the public at all.
AVAILABLE Level 2	Same as AVAILABLE Level 1 with the following addition: data can be used in certain publicly available, predefined visualizations (i.e. maps and graphs), but direct access to the data is restricted.
AVAILABLE Level 3	Data are used in publicly available, predefined visualizations (i.e. maps and graphs). Additionally, the complete BMDE data set is available upon request, subject to approval from the original data provider.
AVAILABLE Level 4	Data can be used in publicly available, predefined visualizations (i.e. maps and graphs) and also may be available upon request. Additionally, some components of the data are made available to existing bioinformatic efforts (GBIF and ORNIS). These bioinformatic efforts only provide the data "marked-up" to Darwin Core, used to describe primary occurrence (location, date and species for example).
AVAILABLE Level 5	Data are used in publicly available, predefined visualizations (i.e. maps and graphs), and are available to existing bioinformatic efforts. Additionally, the complete BMDE data set is available for download directly via download tools.