**Discover Life’s Next Billion Hits**

08-Oct-11, Athens, GA

Attending: John Pickering and friends: Sam, Michael, Kim, Bobby, Leah, Leo, Nancy, Dan, Karen, Liz, David, Gongying, Tryggvi, Becca, Justin, Annie, Jerry, Melina, Nick, Steve, Mary

**Introductions**

**Initial Ideas**

* Place a “News” link on front page to the new stuff/and or events
* Create a listserv for partners, moderated, with multiple moderators (Nancy, Pick, Justin, e.g.)

**Quick Review**

<http://ocs.ischool.syr.edu/game/luckym_web_sep27/mainscreen.php>

Gongying’s development of an online game interface to crowd source the sorting of species photos.

Includes magnifying glass, will develop online help. Four rounds in all. Suggestion: when there isn’t a match, “no match” should be an option. Because it’s a game, there is no way to ‘go back’ and change a record if you are no longer on the page.

How much do people playing this game get to understand about the complexity of life? They are scoring “knowns” and “unknowns.” Will players be rated? Yes. Potentially, the best players may be hooked up to work at DiscoverLife.

The advanced version of the game should consist of LBMs (Little Brown Moths).

The end of each round should have a link to the appropriate page for the Genus.

Bean Museum has started an iPad program so that students can determine the difference between the species on display. It is very important and useful that both organizations (DL and Bean) are teaching kids to be more observant. Motivation: why do people play games? DL gets moth identification. With a large-scale photo submission operation, gamer participation will be key to successful sorting. High scores table will be visible to all gamers.

It’s important to include at the beginning of the Happy Moth game a note to participants that they are contributing to science.

Future plans for this game include development of an app.

Creating cards about bird migration is also addictive. There is a competition that takes advantage of that.

Genealogists have a similar system for transcribing old records.

“I’m doing something that my government should do and they are failing to do it.”

This app could compete with Angry Birds.

Each person has a login, and they can keep track of their scores. Lucky matches are not verified, but are based on previous performance. Idea: Lucky Matches should only be given on known data.

At the moment, moths are ID’d in this way:

1. Deposit 50K moth photos with DL, including where and when.
2. High school students or lab folks answer the question what shape is that moth or is it even a moth (1 character score). We hope the game will take the place of this step, plus get them a couple more steps.
3. Rough sorting is next, with up to 300 photos per bucket. The game will get the photos beyond this step.
4. Computer checker will throw out the outliers. (this could also be inserted between 2 and 3)
5. Final step is ‘blessing’ where groups of photos are finalized as the same species.

Is the software open source? Can it be used for other species? Yes. There is a partnership between DL guide builders and DL’s 20q software, which assigns them the photos. Some identification is difficult; top and bottom views are required.

What is being done with gamers’ record of ability? Someone with a high score may get their results reported directly to a verifier. Otherwise 100 same scores from beginners gets reported to a verifier.

The wonderful part of this tool is that it is a result of a cooperation of DL with gaming experts.

Suggestion: search for a young lepidopterists’ society. NABA is a modern option.

There are other games that could be used (memory, which is different, etc.), but the results will not be useful to the ID process.

Buckets at bottom of page may be better than boxes.

Name: “Happy Moth” vs. Mothzilla

Summary of points:

* No match option should be included where needed
* High scores/crowd sourcing go to verifiers
* Info about doing science while gaming should be included

Game idea: for every 10, include one unknown. It then can be matched with one of 14. If wrong answer, game says what it should have been and why.

If you want, use the lucky match as a measure of the player’s skill.

So far 17000 photos have been keyed to species. There are 20000 to go.

Game should also include instructions to gamers for how to submit photographs.

Front page of game interface will eventually give players options to sort moths, frogs, local sets of photos, etc.

For the game, moths should be relatively common locally, so there are lots of photos. They should also be relatively easy to ID. There are about 200 species in the east that fit these criteria. We may also see differences from year to year. Did the rare stuff go extinct, common ones stay common? Is there anything new that should be reported?

Mothing is more successful than bee hunt, butterfly hunt, ladybird beetle hunt, goldenrod challenge.

This is a very good tool for museums, too. Maybe there won’t be enough pictures? No way. There are way too many. Larry from TN started a backlight porch project, and 20 people in TN, plus 50 from elsewhere.

Need a stable URL, plus a developers’ URL that is hidden.

Part of the appeal is that this is an accessible way to learn about moths. That should also be further developed.

Is there an easy way to upload a photo anonymously to DL? Yes, if it has data. It goes to second tier. DL, however, is not really interested in single photos, because it adds so little to the monitoring effort. Unless, it of course ‘takes off’ and singles add to millions. Normally single photos should be uploaded to Bug World, because they don’t contribute to inventory and monitoring. Project NOAH may be a DL partner in the future.

Jane McGonagall is an advocate of games for good. Appeared on the Colbert Report.

Idea: contact the Grains of Rice game folks to see if a charity benefit reward system could be set up with a non-profit. You will need a donor/benefactor to do this.

We are concerned about receiving inappropriate photos from anonymous accounts.

Once gamers are involved there will be a higher risk of hacker attack.

Previous problems have included discovering that some features are too powerful and sorter indiscretions have damaged/erased previous data entries. Undo Vic is a new program that can undo errors, which is very useful.

**Moth Lists (Becca is mastermind of the moth list organization)**

The building of these local lists will eventually enable the creation of moths lists for every county.

We anticipate that there will be 20 different types of tables, for use by students for school projects but not for publications.

Each table has a column for each moth-er in that location. The moth idea is being promoted at high schools and colleges. There are complex table options for effects of climate on moth flying, etc, and descriptions of possible analyses.

Spark lines is a way of making excel data into skinny graphs for seasonal abundance.

Ruler discussion: best thing would be to have a gridded board for posting under the black light. Color/grayscale adjustment option In the ruler in the shot is also a good idea.

Documenting that moths are not there is important. And very hard to get, because observers don’t take pictures if there aren’t many moths there.

Protocol description is at <http://pick14.pick.edu/moth/methods.html>

Mite data is very noticeable sometimes, too.

**What are the groups for which we want to push photo taking?**

* Moths
* Bees
* Plants (wildflowers, ferns, trees)
* Pollinators in general (syrphids are generally lacking; LA Museum experts and Chris Thompson’s Diptera site)
* Other species groups as partners’ expertise dictates (e.g. lichens)

Idea: There should be a ‘we want photos of…’ page that links from the home page. It may be species level photos, it may be photos of diagnostic parts of organisms.

**Hand held access**

Charlie Elliott Wildlife Center

20 participating teachers that don’t know all of the species they find

For the site, there was a bioblitz of common species and individualized guides were created for the major groups. It needs to be more clearly stated that the groups’ species list is for common species, not all species.

What to do about the issues that the guides are never complete? Every time someone sends us an additional species, DL will include the new occurrence to the guide. One component of what DL does is similar to a Wiki. There needs to be user input for improvement.

Lists of species are very valuable after keying. If the name is on the list, the ID is likely to be true. The issue is where the name is not on the list but really does occur there.

Suggestion: there needs to be more metadata about each of the DL guides so that the user can tell which guides resolve, which are beta, which are abandoned, which are actively being improved.

One thing that is being done for bees is that occurrences at X and Z state are interpreted so that state Y’s list has the species listed as ‘potential’ and it appears on state Y’s guide.

With lichens, Sean is building county lists. People relate to their county, and they have an allegiance to the county. They could be recruited as list contributors. County level data is the way to go.

Georgia’s state botanical garden has narrowed the lists of wildflowers to the trail level. This can build up to the county level lists. Seasonal variation will also be very important for many species types. This is a future feature in many of DLs guides.

This is not GBIF. We honor each data providers’ desires for where the data is generated. If they don’t want the data released it won’t be. And when there is a (time or provider) series of data for the same lat/lon, the point that is displayed at the deepest zoom level is randomly generated, so any one data provider’s info can’t be scraped completely. GBIF releases all the data (which could be a subset of data) that is contributed.

For volunteer monitoring data, creating the schedule and habit makes the task easy.

Idea: approach weather monitoring personnel and add a task like an insect trap or simple protocol to take moth photos, and they may be very amenable to cooperating with DL.

**How can we put DL out into the public sphere? How can we institutionalize DL?**

We all need to ‘own’ what DL is doing next.

FYI on death of main DL players: If Pick dies, Maria Pickering will decide who continues in his place. If Kevin Weick dies, his wife Claire will decide who takes his place. Justin is in charge of the hardware/software decisions of DL. Pick is actively seeking an endowment to continue DL into perpetuity, by licensing the 20q software to a commercial enterprise.

Idea: Donate button on home page.

Idea: Set up specific things people can help purchase through the donate button.

Idea: Go on NPR to talk about the Happy Moth Game.

Giving the data away only a point at a time should be more clearly announced for reference by potential data providers.

Data with errors is corrected and shown in the right place. Zoo specimens and cultivated specimens need to be tagged and preferably eliminated. The only way to accomplish the tagging and elimination from the default view will be by crowd sourcing. DL can accomplish this once things are reported.

LUNCH

**Technical Discussion**

Nancy already set up a google group for news.

Currently the bandwidth is provided by UGA as long as Pick is a professor. After that, a different option will need to be sought. IP address ownership would be a battle as well. There will be a need to institutionalize the site, with, e.g., EOL. Need to avoid gov’t as much as possible. Within 5-10 years, tops, another solution will need to be sought. UGA server housing costs 160 dollars per machine per month. There are six now.

Jerry Castes from ALA has also offered to host, but the ideal is to keep the site where your main user base is located.

When STRI offered to run a mirror, they couldn’t get the machine to function behind their firewall. MOBOT also mirrored for a while, but shipped it back as well. USDA PLANTS has described a chain of security and firewall requirements that makes mirroring there impossible.

A mirror would only require a single large box. Ideally, there should be a chain of mirrors that back up to each other.

Sun solaris is being phased out and replace with Apple server boxes. 11 TB and 12 cores and 24GB RAM for $7K.

It would be great if EOL/Woods Hole would be willing to host a mirror immediately, to provide more server security.

Problem with mirrors is that it adds a layer of technical complexity to the updating process. Work stations vs. rack mount units would be an easy adaptation.

Morph Bank collaboration model may be worth investigating (with the commercial sector).

Already host datasets that are no longer being served. Are also able and willing to preserve other Web sites’ look and feel.

In general, the applications that go belly up first are the commercial solutions. (Kodak, TopoZone)

Look into getting detailed map layers from the USGS National Map. Terraserver has given problems.

Georgia Natural History Survey could end up as the umbrella for DiscoverLife, and be owned by the state of Georgia.

Please everyone should be included as technical advisors to DiscoverLife. There is a DL Board that are famous people, but they are not dedicated to DL; they are doing their money raising for their own causes.

Pick has met with billionaires to discuss selling 20q to them in exchange for endowing DL in perpetuity.

Mark Becker, CIESIN, may be approachable for hosting a mirror of DL. They are well-funded by NASA.

Another option to investigate is UGA at Tifton. Talk to Chuck Bargeron and his deans while pointing out the importance of the newly forming Georgia Natural History Survey.

At least one outside mirror is important, in case of a natural disaster. Eventually it would be great if the mirror became true server to receive info and trickle it back to GA.

Current hardware can handle the anticipated growth for the next five years.

Australians are going to come in with $20K worth of hardware in the near future.

Other large organizational interest? Penn State

When MOBOT competed for the mirror of DL, they didn’t say that their bandwidth was small, so the mirror eventually didn’t work out.

There is a need to send data rescuers out to work with naturalists and other data holders to get their data onto DL before it’s too late.

How do we get photos from Flickr? EOL does it. DL actually does it, too.

Idea: tell listserv that Flickr photos can be downloaded to DL and how.

The issue is the ongoing reconciliation process between Flickr and DL. EOL has a tagging functionality in Flickr where photographers can advise what pix to upload and when.

Idea: Pick meet with Liz and/or EOL person to make DL Flickr upload more flexible/powerful.

**Web Page Layout discussion (send suggestions to Nancy)**

* Life lists need global header
* Breadcrumbs design at the top
* Help icon no one uses can be helped by creating a glossary of terms. Do it with simple links to a glossary page. A student may need to be assigned to this task in order to fully implement it. Another alternative is to make the glossary a Wiki that album owners can modify.
* Guide builders will use a symbol [TBD by Pick] before/after terms that should be displayed in the glossary.
* Becca will manage the glossary(ies). Determine if we have separate glossaries, one per guide, or one integrated one with various definitions for overlapping terms.
* The glossary needs to be a tool for guide builders that is constantly updated and built upon.
* Each of the ‘buckets’ on the home page needs a champion or a very small team to organize and take the lead on the content development.

**Themes for discussion**

* How can DL better cooperate with EOL? E.g. discussion of photo uploads. How can they be better tied together?
* What additional partners can be included and how?
* How can the sources for the information rich species pages of DL be built out? Is this still important? Or are we now emphasizing occurrences through photos?
* Species interactions
* Quality control issues
* Species pages
* Images are servable from any other Website as an image forever; what are the copyright issues
* Data generators recruitment and support
* Georgia Natural History Survey: how do we push it forward? Include in the plan the ability to broaden taxonomic scope, or franchise the operation. Don’t step on the toes of existing NHSs, however. Build on principles of competition and pride. Hook up also with weather stations.
* Education: Don’t forget the need for ‘prepackaged and easy’ things for teachers. Write grants for teaching the teachers. Remember to give tables to teachers for student projects.
* Organizational structure including Wiki content? How to keep standards high and volume, too. Going the social networking site direction has potential pitfalls, but there are tasks for amateurs that are appropriate to amateurs, and the volume of the work can be very helpful. Making DL the biodiversity monitoring database Web site. “All taxa” is daunting. “Focal taxa” is doable.

Sunday, 10-Oct-11

Topics to discuss:

* Happy Moth
  + Subtle sounds are appealing to many
  + Sound option must be easily turned off
  + Pictures can be scattered over the desktop, but should not be overlapping, and one should be selected for user
  + The offered selection can by changed by user by clicking on different photo
  + “Clicking the destination” will cause picture to be sucked into the destination bucket
  + Statistics are very important, and could even be more detailed.
  + NPR options are many. Leo’s cousin works on Radio Lab.
* Big Picture: how do we get the next billion hits?
  + Tools for the public to do natural history
  + Biggest emphasis is on inventory and monitoring, not individual reports
  + GA has 137 counties and is the biggest state east of the Mississippi; physiographic diversity from mountains to coasts
  + Low in science education among states, and this can change that
  + 89 (threatened) weather stations could use DL monitoring to justify their existence
  + butterfly raising for teaching science morphology and pre-established teaching standards.
  + DL’s role is to provide the best technologies for biodiversity inventories and monitoring
  + Three main tools to help with issues of pollination and herbivory:

1. Localized identification guides [each guide needs a description about its completeness, scope, resolvability, and a contact for improvement suggestions who can tweak them]
   * + There is the ability to support people who want to create their own local guides (Nancy is the contact for this)
     + There is a demand for the creation of an ‘insert this guide in my page’ functionality. Btext burl and bicon can bundled put together and DL delivers it so that the guide appears to be on the creator’s web site, but it’s not. More advanced option: insert onto the page like EOL inserts the global mapper on species pages, and it would run like a widget within a page. Eg. Similar to the LA Museum i-frame setup.
     + Taxonomic groups being emphasized by ID Nature Guides:
       1. **Moths** (this may be the only group we should emphasize for inventory and monitoring)
       2. **Plants** (**wildflowers**, trees, ferns)
       3. Bees (scientific audience)
       4. Lichens (?) easy to photograph and around all year and has NSF support. Only Sean, Sam, Bobby. However, not an exciting topic to the public
       5. Pollen: need to scrape [www.paldat.org](http://www.paldat.org) (palynological database) web site. Easy for students to do with setting out sticky cards, but identification is only do-able to Genus.
       6. Possible others:
          - Dragonflies
          - Ladybugs
          - Fungi (associated with Mushroom Observer)
2. Global Mapper
3. Photo Albums

* Partnerships & Organizational structure & Non-competition (30 min. brainstorming)
  + Species pages: Information rich species pages of DL (10 min.)
    - EOL is doing species pages, DL should get out of that business
    - One exception: Bee pages
    - Note: historically, DL was first in species mashup pages
    - Need: hold a symposium at SPNHC meeting on species monitoring, inventory, and species pages. 2012 meeting is on technology, at Yale University, Newhaven, Connecticut on June 11-16th, 2012
    - DL can give users what, where, when, and phenology data. DL’s scope is “anything anywhere any time.” However, unless resources and expertise become available (e.g. Sam and bees), ***the default emphasis is currently inventory and monitoring of moths and butterflies, flowers, and trees in North America.***
    - DL future partnerships that could make a lot of sense
      * BAMONA
      * NABA
      * The other lep group
      * SPHNHC
      * NH collections seeking applications for their newly digitized information
  + Images copyright issues (10 min)
    - What does DL want to do about copyright of images?
    - Current policy: DL images, unless specified otherwise are freely available for non-commercial use if credited.
    - EOL has a strict policy of images being Creative Commons (CC)
    - Range of CC licenses with different conditions
      * Don’t give up rights, but adds relatively easy conditions for use
      * EOL can’t use licenses with a ‘no derivatives’ restriction (two of the CC licenses are not acceptable to EOL)
      * Because of the legal complexity of ownership and use, the sharing of images between DL and EOL is not a priority at this time.
* Services that DL will provide:
  + Where = web service of maps
  + When = time lines
  + Species interactions, as in what eats what. Controlled vocabularies.
  + Diagnostic characters for identification
    - should include ‘look-alikes’ information
    - can be used to find ‘shortest path’ to species ID by accumulating unique characters.
  + Special literature
  + Local checklists
  + Unique identifier prefixes for records
  + Image storage
* Recruitment:
  + Education (before lunch)
  + Georgia Natural History Survey
    - Georgia DNR would be supportive of this effort, says Bobby
    - Potential concern about specific location of T&E species being publicly available, but we don’t anticipate that being an issue
    - BioBlitzes at the county level will hopefully get all 80 coastal plain counties involved.
    - Geographical barriers. Mountain region under represented and 0 participation in SW GA. Bobby has contacts in SW, Tom Patrick and Lisa Cruise can help.
    - There is a misunderstanding (Tom’s) in the DNR about what we are trying to do on species distribution that needs to be resolved.
    - How is excitement being generated without money? Grants are being written, donations have been received from this group.
    - Priorities will be plants, moths, lichens, possibly salamanders and frogs.
    - How to motivate large numbers?
      * To motivate the right kind of participation, talk about a county natural history atlas.
      * Associate names with the county records.
      * Set up a competition.
      * Be sure to set up quick display of contributed records.
      * Approach preserve personnel to help them inventory their species for inclusion in management plans.
      * Enable easy viewing of county level maps with their data points.
      * Biggest impediment is 159 counties that GA has.
      * Political boundaries need to show up better.
      * Zooming in to a specific place with one click is needed.
      * “First county records” need to be publicly announced, put on FaceBook, etc.
      * Announce top monthly contributors.
    - How will outside participants contribute to the GANHS?
      * By starting one in their own states
  + Data generators & support
    - We have the tools and the interest in their data
    - Small sources of inventory and monitoring data are very important
    - Will provide the service and support for all data fields that are provided.
* Species interactions (10 min. plus 10 min. discussion)
* ~~Quality control issues~~

Note: the Guides have been redesigned to remove the frames to make them handheld friendly.

LUNCH

**Wrap up comments**

Bobby: Whether or not I get the grant I’ve applied for, I’m going to do my best to work through the educational system and the counties to get people on board for the GANHS. I hesitate to go to garden clubs, because they may not produce data, but they could talk up things. I will go to Chambers of Commerce to set up BioBlitzes (with expected species lists and experts participating along side of novices). Check out the “Citizen Scientists’ Bill of Rights” on the Web. Will also work through the GA Botanical Society because I’m on the BOD. Linda Chapin (?) Elaine Nash has a lot to say about the Marie Mellinger grants. I have friends in GA Plant Conservation Alliance I will also recruit.

Leah: I’m in.

Steve: I’m encouraged that in the age of dwindling funding we are trying to work out ways that we will not compete and will cooperate. I’m also glad to see that we have better defined our focus, to GA for the moment. It is a great place to start because it has half of the species of the eastern US.

Pick: How can we get the managers of the different field stations excited and involved? Steve and Liz could help us a lot with this.

Liz: There are two retired botanists involved in Banshee Reeks, and that has provided us with focus. Also there is a training program affiliated with the reserve that lends us more expertise and credibility, plus the Loudoun Wildlife Conservancy. Where the interest and capability is there, the locals should build the guides. There needs to be a tried and true curriculum to get the guide builders going. Giving a workshop on how to build a DL guide at ESA may be an option. Also it seems like it would be useful to have an annual DL blitz, so that volunteers could ‘bang away’ at what needs to be fixed, verbally and literally. Create a business card sized photographic ruler for use with mothing. Make it downloadable from the DL web site. E.g., like the one cyberflora produces.

Nancy: I could go up and provide training in guide building, if there are enough students, and Sam could also train on this and give support to the group. I want to keep my eyes on the big picture, as in our five year plan, possible franchising to other states of the NHS model. We are applying for a state level grant to do another teacher training workshop next summer (like the one that brought us Bobby) for 1 week and 20 people. Bobby can help any time BUT May-June. NSF proposal in the works to Informal Science Education, to build a network of study sites in GA and CA. Third, Macrosystems Biology group of NSF may be something we will apply to.

Pick: I think we’ll have more success with feedback from guide users than recruiting guide builders. I think the botanists should give us feedback and tell us how to improve the guides. In our experience, learning guide building takes a week or more.

Bobby: I might get 4 out of 50 that could be interested in building the guides.

David: Bee names issue – DL keeps a very complete list. Pick says he’d like to work with ITIS to get the list updated. ITIS does monthly updates, but I would like to be able to submit a bee name updates as they occur. ITIS is too overloaded at the moment to get the work done. They trained a potential volunteer, but they didn’t work out. ITIS needs a DwC Archive document in order to be able to respond to the updates quickly. Recently, ITIS has uploaded 75K plant names. Grasses of NA, Gesneriads of the world. We will be close to done by the end of the year with plants. Sarah Alexander is the main ITIS botanist contact who sits at the Smithsonian. Alan Weakley (formerly TNC) has provided the most definitive plant names lists for the southeast, says Steve. Pick says Becka is the contact for moth names for DL to ITIS. Micromoths are difficult in to resolve especially in CA.

Pick: We would like to be more involved with ITIS for moths, plants and other names.

Annie: DL is a virtual network that needs to work with and get buy in from brick and mortar organizations in order to achieve its goals. There is a whole set of studies that have been made about the difficulty and benefit of virtual networks. Study them to understand the pitfalls before they occur. I also suggest you set up a more definite structure that defines roles and provides concrete projects/deliverables that each volunteer is in charge of. This will help Nancy immeasurably, because it will take the load off. But there needs to be more organizational structure and individual rewards for useful deliverables to the network. Monthly teleconferences where volunteer leads of projects report on their progress to the whole group will help maintain focus and direction, and create synergies among the different projects. Nancy can provide the hub of the direction to these projects, and benefit by successfully delegating more tasks in a coordinated way.

Pick: I either want to moth or code, so I am very, very grateful to Nancy for taking a lot of the organizational load off of my back. I think we should hold an annual meeting for the GANHS, in order to advance the organization and plan for the coming year. Let’s plan for Columbus Day weekend, as long as there is no home game. We are going to give it one more year pushing mothing. That will be our primary focus.