

STReaMS Beta Release

3/9/2015 - 3/10/2015, Grand Junction

Key: **Question**, **Action Item**

MARCH 9 – INTRODUCTION TO STREAMS DATABASE

Introductions:

- Hatcheries need to upload data; lots of folks need to run queries and pull data out of the central database; lots of ways data are collected for various projects and these will all need to be uploaded into STReaMS in batch

STReaMS Overview: (presentation by Kirstin Holfelder)

(Note, Year 1 runs through September, 2015; Year 2 runs from October 2015 – September 2016)

- Kirstin's PowerPoint presentation showing the structure of the database behind the scenes and core tables.
- Year 1, user management roles need to be decided as a group. **Question: Who should have access to edit data?** We will track revisions but roles need to be carefully thought through. Perhaps PI and data managers can edit? **Action Item: We need a working group to develop roles.**
- **Question: Are the data subject to FOIA?** This isn't clear since the database will be physically located at a state facility but is funded with federal dollars. **Action Item: CNHP will discuss this with CSU legal council (as well as CORA, Colorado Open Records Act).**
- **Question: How should we handle sensitivity of the data?** CNHP generalizes locational information for sensitive data. CPW deals with data sensitivity issues as well, especially on private lands. Likely, the solution is to generalize locations and hide certain attribute fields from certain groups of users to protect the resources. This is also discussed later on in the conference. **Action Item: This should be discussed in a Working Group. It can likely be part of the Roles Working Group since it is related to permissions.**
- **Question: What is the minimum that needs to be done to meet the Federal Open Data Policy guidelines?** No one in the group knew enough about the requirements to answer this question. **Action Item: BOR or USFWS need to identify a federal employee that knows the Open Data Policy guidelines and CNHP will discuss this issue with them in more detail.**

- User manual will be part of Year 1 as well as canned downloads. The advanced query builder is scheduled for Year 2. This will allow users to cherry-pick attributes to download through a drag and drop interface. Users will be able to save their custom queries for future use.
- Data manager and quality control tools are Year 2, as well as batch uploads. Batch uploads are too much to do in Year 1, but these are key and will be the highest priority for Year 2. Until batch upload functionality is developed, folks will have to send upload files to CNHP.
- **Question: What should people do to upload data this year?** Dave - business as usual (status quo) for this year until the database is up and running.
- Revision tracking and database dictionary will be Year 2
- Year 3 will include all the stuff we still need that we don't know about yet. Crosswalks from difficult data sources will also be part of Year 3.
- The development cycle will follow a 3wk/1wk cycle - 3 weeks of programming/1 wk of testing and then a new version will be released. **Question: At what point will current data be in the new database?** Data will be in by September 2015, the end of Year 1. The database has a new structure so it does take some formatting in order to import data as comprehensively as possible. Right now, a lot of information is in notes, but it will be parsed into appropriate fields for the first release where possible.
- Kirstin demo's the website
 - How the fish was captured is in the Browse encounters page
 - Information about an individual fish is on the Browse fish page
 - Browse fish includes fields that don't change with each encounter. **Question: What if some of these fields do change (sex and species in particular)? Ex. sex was unknown, and now we know, or only genus was known and now we know the full species. How will this work?** This is important because species ID and sex ID is difficult for some fish, especially when they are young. Kirstin will have to use revision history. Do we need to track this over time in the encounters table or will revision history work? For biologists, this seems to be important, to be able to see the history in the encounters table. The latest record might not necessarily be the most accurate so we need some sort of control here (if species was known at some point, you wouldn't want this to get overwritten by a new encounter where only the genus was identified). Kirstin – we could move this to encounters and assume that each person has a different opinion, but this will fundamentally change how the database is structured and needs to be well thought out (see further discussion below).
 - **Question (from CNHP): For lookup values, are there manuals available?** There is a three spp manual, a San Juan manual, Upper Colorado uses a different set of

codes. So, there is no standard set of codes. Different agencies have to reformat data to share. We can assign aliases to alleviate confusion, but we need to know what codes to store in the database. A spreadsheet with codes to store in the database, cross walked to current codes used by the programs will be helpful. It would be preferable to pick one coding format to use for uploads to speed up the database, but it is not mandatory. **Action Item: Add code spreadsheet to the Data Manager task list.** There are also ISMP Codes (species, habitats, everything you need codes). **Action Item: download and review ISMP codes.**

- Most data entry will be done through batch uploads but you can edit or add individual records if you need to.
- **Question: What about hybrids?** They are on the list with a separate species code. This is probably the best way to deal with them. There will probably be the most confusion surrounding hybrids or unknown species.
- If we have antenna data, it will just be a list of tags and time stamps. The species will be unknown if the tag is not already in the database. Otherwise, it will get assigned to the species (and other details specific to the fish) already associated with tag in the database.
- **Question: What happens if the tag is unknown?** There might not be a match in the database (due to error, etc). How will the database handle that? Can these be flagged and reviewed by a data manager?
- **Question: What about multiple tags for one fish?** Do you see all tags? Right now it is the most recent tag that will get auto-populated. Comments: need to see the whole tag history for encounters. This shows that it is an old fish, and shows the whole story. Kirstin says 2 tags is easy, more than that gets complicated.
- **Question: What about duplicates?** Will you get an error message if there is a duplicate tag number? Kirstin: duplicate encounters are a problem. What about a new fish with that tag. This can happen from typos, etc. Kirstin: this would require human intervention. A data manager will have to make this decision. One comment – add “a” and “b” to the tag number as a temporary fix. (AMY: can we flag duplicate records during the upload process so they don’t get uploaded but go into a holding cell and then get approved, or not, by a data manager?). **Action Item: CNHP will figure out a system for flagging records that need to be reviewed by a Data Manager before being uploaded.**
- Comment that the programs need a list of items that need to be addressed by a data manager and how much time that must take. **Action Item: CNHP will start a Data Manager task list.**
- Kirstin: we need a core group of folks that we can confer with monthly (look-up tables , etc.) so we can get feedback throughout the year. **Action Item: Organize**

a core team and set up monthly meetings, with the ability for drop-ins to join when their schedules allow or when certain topics pique their interest.

- **Question: What about wildcard searches?** Yes, this is current functionality through search. Note: this was not working during the hands-on session. **Action Item: CNHP will look at programming and fix wildcard searches.**
- Users would like to search on year type and date ranges on the encounters page. **Action Item: CNHP will add more search items to the Browse encounters page.**
- **Question: What data are in the database now?** So far, it is largely a pit tag database. Non-native fish will be next and then maybe fish that are never tagged eventually. Comment: Biologists go out and sample for a day and there needs to be a way to collect field data (which is applicable to all the captures that day, etc.). This is currently stored in the Sampling Events table. There are datasets with sampling events that we can import to STReAMS now. It could have floy tags, pit tags, water quality, etc. Ex pike minnow – collect site data, river miles, crew, xx amount of time, electro-fishing. Kirstin: thinks the data are in encounters now and she needs to pull it out into sampling events. Comment: There are a few projects that mostly collect habitat data. Comment: it could be specific to an encounter. Kirstin needs to review the encounters table. **Action Item: CNHP will review the database structure and identify where to store additional field data such as habitat, survey info, etc.**
- **Question: Will we be able to filter by hatchery?** Yes. Under Browse encounters, there are a few hatcheries listed in the sampling org dropdown. We will want to add all applicable hatcheries to the organization table.
- **Question: How are records linked?** Sampling events are connected to fish data through a relate table (an additional table in the database). There are many relate tables in the database, connecting records to one another.
- **Question: Can we enter GPS coordinates instead of river miles?** Yes.
- Oftentimes they cannot get GPS coordinate. **Question: Can you relate a coordinate to a river mile?** Yes. Travis has a table. Contact Travis for coordinates. Most people collect UTM's instead of lat/long (both zones 12 and 13). Is it possible to go the other way too, connecting a GPS coordinate to a river mile?
- Comment: encounters are attached to a point but sampling events are not. Some people are tying everything to a point, i.e. work is point focused.
- **Question: What about missing data?** Researchers won't know what is missing when they query data. From a program perspective, it would be good to know who has already submitted data for the year. The committee needs to know that everyone has submitted their data in order to continue funding them. Studies

that are missing: Add a date under studies to last date data uploaded? Perhaps add a way to manage people? CNHP could add a new table to the database to store batch upload details so a data manager could see a list of all uploaded files and when they were uploaded. This would probably be a Year 3 task, maybe Year 2.

Action Item: Add reviewing uploaded data/ID missing data to the Database Manager task list.

- **Question: What about antennas?** We need to know when it was on/off, and where it was. When an antenna goes out to a specific creek, we need to know when it was taken out and when it was brought back to the office. You don't want to finish your project with missing data. Kirstin thinks this is a sampling event (Dave thinks of it this way as well). Kirstin: can tell you when it was on and when it was off, but it will be hard to know when it was on but not operating well. How would we check for noise? Perhaps this could be looked at in Year 3. This is icing on the cake...
- **Error: Bonanza Bridge antenna is associated with the Pike River and it should be associated with the White River. Check the database. Amy's note: this has been fixed post-conference.**

BREAK

Discussion

- Changing the characteristics of an individual fish is discussed again. Do we want to track it as a data manager tool (tracking revision history) or do you actually want another record in the table? Some folks want another record, ex, important for population estimates, some folks think we do not need more records in the database and just need to track the field revisions. Another option could be to flag records that have changed and then you can download the changes if you want (or not). Species and sex fall into this category. The question on the table is whether these two fields are stored with an individual fish or with a specific encounter. Most changes are just errors (ID mishap due to immature fish) or more specific information is acquired (genus to species, or sex from unknown to male or female). Think through this more. From a database perspective, it would be more efficient to store sex and species at an individual level. We need to think about how this would affect searches, filters and downloads. **Action Item: Address this topic in a working group.**
- **Question: Can you reassign the pit tag?** Say, a fish is dead and you take it out and reassign it to a different fish. Would you take a pit tag off of a dead fish? Most folks think no, but some say yes(?). Comment: You do want to keep track of a dead fish (say it was consumed by a larger fish), but once the tag has been used, it should never be used

for another fish. Also, there might be issues with sonic tags. Sonic tags can be used for multiple fish, although these fish will also have pit tags. Something to think about in the future.

- Studies are usually similar, but sonic is new technology so some studies might diverge from the norm. This might be a Year 3 topic to explore.
- Look up tables will be in the help manual and accessible online. Some are editable by a data manager. **Action Item: Add to data manager task list.**
- **Question: Under encounter history, can you view details showing the studies information?** Not yet, but this will be added to the database during Year 1.
- **Question: Who will have access to the data?** Right now, users can do anything to the data since we are in the testing phase. Ultimately, we won't use these data but we will need to regulate permissions for viewing and editing data in the final database. Kirstin will approve users for now for the testing phase, but a data manager will likely be approving users and assigning roles in the future. **Action Item: Add user management to Database Manager Task list.** We need to talk seriously about roles. This came up earlier in the conversation. Dave – the recovery program will be hiring a database manager that will have full edit access. Kirstin – you can also make roles by project. The public will be able to view the data at some level and there are data sensitivity issues. Dave said we need to follow open records policy but he is not sure what that is. CO state statute says that data cannot be distributed on private land. This is a huge issue that needs to be discussed in a working group. **Action Item: This should be discussed in a Working Group. It can likely be part of the Roles Working Group since it is related to permissions. (as noted earlier in the workshop)**
- Although there will likely be some general level of information available to the public at all times, there might be situations when data requests need to be fulfilled by a Data Manger. **Action Item: Add to data manager task list.**
- **Question: Can the system contact people that are interested in specific fish or tags?** Kirstin: the system can email you. We can definitely send emails for completion of a batch upload, stocking event, or study and even more if the group thinks this is appropriate. Email announcements for large events (like large stocking events) are definitely of interest.
- Dave: recaptures vs captures. How do we keep track of fish with old tags that get new tags? If you did not implant the tag, then it is considered a recapture. Somewhere in the system we need to know when we deploy new tags. Need to know the old tag history in order to know that it is a recapture. Encounter history will show old tags and new tags (it doesn't now, but it will). An event right now it is split between tag history and encounter. This needs to be revised. Need to know when an encounter is tagged during a deployment.

- **Question: What about relocation?** Example, a fish captured in ephemeral wetland and released in a new spot, or pike minnow trucked upstream. Would it be an encounter, or a stocking event and then an encounter? If it is a stocking event, it will get lumped with hatchery data. Perhaps pick brains with PNW folks on how they deal with relocated fish? Could we add two fields for captured and released? Or do we have a new field for translocation? Could this be considered a release instead of a stocking event? There is some talk that it should be a capture and a release. We need to be able to calculate distance traveled by a fish, so we must have a way to subtract out capture/release points from any distance traveled. Kirstin: we can add a new type of encounter for transplants. **Action Item: Add a new encounter type to the database for transplants. CNHP will contact folks in the PNW and see how they handle this in their database to make sure travel distance is accurate.**

BREAK

Hands-on session in www.streamsystem.org

Comments from hands-on session:

- Strip punctuation out of pit tag codes for searching.
- Wild card does not seem to be working for pit tag codes.
- Total number of records is not showing up for searches (it does for filters, but not searches).
- Folks want to see dates added to filters and others mention projects. People like the idea of filtering data before downloading files, which lends towards adding more filter options.

MARCH 10 – HATCHERY SOFTWARE AND DATA LOGGERS

Morning follow-up

- Dave opens the floor to first impressions about the database. Open floor. General consensus is that it is set-up intuitively, user friendly and was easy to jump in during the hands-on session.
- **Question: How do we report errors?** For now, email errors to Kirstin (her info is on the “Contact Us” page), but we will set up a better system for error reporting in the future.
- **Question: What format should data come in?** This has not been decided. **Question: Should we use templates from each program or some combination?** It depends on how much standardization we want as a group and really depends on what the recovery

programs want. We assume there are valid reasons the groups do things differently. We will need to store data in a standard way in the database, but we can accommodate different formats during the upload process.

- **Question: Are we going to have work groups to answer questions or solve problems?** Biologists may have different opinions on descriptions and definitions. Now is the time to visit this. Working groups might be the best way to answer these questions. **Action Item: We will need working groups for various issues. Write up a list of issues that need to be addressed by working groups.**
- Between now and September Kirstin would like to hold monthly webinars to discuss development issues, address questions/needs that arise, and share release notes. Who will be on this call? There is a core group: Dave, Kevin, Scott, Travis and Koreen. Dave would like agencies to select a point person that could provide technical input on these calls. If anyone is interested, Dave is soliciting volunteers. Kirstin: usually there is a core group, and then certain experts participate now and again based on what items are on the agenda. Dave will send a message out this week, reminding folks to participate if there is interest. CNHP will send out the materials before each call so folks can decide if they want to jump in. Dave: it would be nice to pick a certain day of the week that is consistent and regular. We will solicit opinions on when that might be. **Action Item: Create Core Group consisting of Dave, Kevin, Scott, Travis and Koreen. Identify a consistent time that works for everyone. Set up monthly meetings, with an open door policy for anyone to attend.**
- **Question: Is it possible to track calculated fields like growth, movement, days in the field, etc.?** Kirstin: it is easy to calculate fields. Just let her know what you need.
- In September, there will be the first draft of the database. It might be ready for reports this fall.

Hatchery Software: Biomark software and PitTagIS (P3)

- Dave – we'll review tagging software routines used by hatcheries and we need to make the connections between the tagging and the database. How can we make this process efficient? Two software packages are used now: Biomark and PitTagIS (P3).

Biomark (presentation by Mike Gross)

- Hatchery has been using it for two years. It did cost money, not 100% sure how much, but Mike thinks it was less than \$100. The old software was problematic and had a lot of down time. Biomark's tag manager has fewer errors, less duplicates, and has worked out well for them.

- To use the software, plug the scanner into the computer to download (the scanner has a battery and will save all of the data), select ports and verify. Then, perform the memory download. You need to checkmark the box to filter for duplicates. Build an Excel file manually and save it before scanning the first fish. Mike names the file by species, date and pond. Scan the fish, and the software automatically puts the tag number and time stamp into Excel. The time stamp (date and time) is left as-is and not parsed into two fields. The length and weight fields are entered manually. The Biomark software won't scan the same pit tag twice. If you try to scan the same pit tag twice, it does not add the code to Excel. Once you are done in Excel, stop scanning in Biomark tag manager, clear the memory so it doesn't fill the hard drive (eventually the scanner will bog down and get slower). **You have to clear the memory off of the all flex reader in order to clear it** (this is a bug in the Biomark software – the clear memory does not actually clear it). You can also clear it from the wand (it doesn't have to be cleared from the computer). This process works well for big batches of tags (especially if a length or weight was missed – it is easy to go back and fill in the missing data).
- **Question: How is this submitted to the database?** Later on, they add some standard attributes to the Excel file, like hatchery, pond, and other stocking information, etc. before submitting the data to the recovery program. Every row will end up with about 15 fields of information. Kirstin does have samples of Excel files from the hatcheries. Kirstin: all of the stocking attribute data can be added when the file is uploaded into the database (so you enter it once, similar to the profile created in PitTagIS) and it can be uploaded the day it is done. We will make sure the hatchery can export their upload to a file, so they can archive an Excel file with the full suite of attributes like they do now.
- **Question: Do we use periods or no periods on tags?** San Juan doesn't use periods, Upper Colorado does. The Biomark software doesn't, but PitTagIS does. The old scanners use the period, new ones don't. Kirstin: how do we want to store it in the database? The decimal point visually breaks up the code. How will storing it one way or the other affect functionality or the speed of queries, downloads, etc.? Dave: likes the period. The prefix used to mean the manufacturer, but maybe that is changing? This is a good topic for a small working group. **Action Item: review this topic in a working group.**

PitTagIS (presentation by Matt Fry)

- This is freeware and was designed for PNW, so some of the fields are not applicable but it is still useful. Also, Biomark software has to be licensed. They can only have it on 2 computers, so if a tagging station goes down, you are out of luck. PitTagIS can be on any computer since it is freeware.
- To use it, the first thing you do is set up your profile (essentially your canned attributes). Since this was designed for the PNW, some of the fields are not applicable, but you do

have to fill all of them out in order for the software to work properly. One major limitation is that you can only add 3 digits to the file name (the rest of the name is auto-generated which not very user friendly and can lead to confusion).

- You can use the attributes to identify handling stress (ex do fish have better survival rates if they are stocked at certain temperatures?). The new database should make these sorts of analyses easier so they can identify the ideal stocking conditions. The canned attributes can be changed after the stocking event takes place (they usually let tagged fish rest 3 days before they are stocked so if anything needs to change in the profile before the stocking day that is not a problem). Matt's hatchery checks the temperature and pH. When fish are transported, they need to add bicarbonate to the water to keep the pH up (otherwise CO₂ from fish will the lower the pH to undesirable levels). For the most part, Matt likes the program (but there are some quirks and disadvantages).
- For brood stock, you can scan a fish and the genetic details come up. This is light years better than what they had before. Biomark software is quicker for tagging, and it can remove duplicates, but P3 can manage information important to brooding and Biomark cannot. Ideally, combining both software packages would be the best of both worlds (Biomark tagging with PitTagIS brood stock info).
- **Question (from CNHP): What is a production fish?** They are fish raised from larvae at the hatchery and stocked in the river.
- To use P3 you have to use their database standards. Dave: do we need a standard form from hatcheries? When you export to Excel, it has to be reformatted. The hatchery usually tracks height and weight for first 100 fish, about 10%, and then only length for the rest. It is possible for fish to go out in the river without a tag, but it is unlikely. The user does have to add a bunch of columns to the spreadsheet before sending it to a data manager.
- P3 and Biomark are the only two software packages used in the hatcheries.
- There is debate on the size of the fish you can tag, for the health and safety of the fish. Location is important too.
- P4 is coming out in the summer of 2015. Training webinars will be available. There is also a mobile app version in P4. The interface looks completely different, more similar to an app.

BREAK

Data Loggers (presentation by Matt Breen and John Hawkins with help from Tildon Jones)

- Several people do use data loggers. Some people use both loggers and notes (not at the same time – it depends on the project). Dave – data loggers should save time and

decrease errors. This presentation is informational. We aren't forcing people to use data loggers. We know many people still use paper.

- Preview of the Juniper systems. The units are water proof and last a long time (10+ years). Units have a numeric key pad. It also comes with a stylus or you can use your finger. Units cost about \$3,000. Matt and John paid an extra \$2,000 to have custom applications (or data sheets) built. Their data sheets have site data and fish data (press F5 to switch between sheets). There are QC controls built into the forms so it is easier to deal with downloads (less error). Forms are set up for different project types (these are different applications on the unit). You can preload values by pressing F2. You can also add values to the look up tables and modify QC controls (change limits, etc.). GPS coordinates auto-populate (UTM) and it auto-saves. They back it up at night, but it would be possible to drop it during the day and lose the unit. That would lose an entire day of data, but data sheets can blow away, etc. so either method is fallible. Dave: likes the field Pit Tag Batch number. The tag is populated automatically when you scan the fish so you don't have to manually type in the tag. Place the cursor in the tag field and it will automatically populate. The tag check field is an extra QC measure (this signifies it has been verified by field tech). QA/QC has been most important feature, that, and not typing in tag codes. If you pick native or non-native fish, you will get a different form.
- They can share their applications with you, but you will need a license to download your data. Each unit has a license and you need the license in order to download the data from a specific software program. Comment: there is a software program that creates an app from an MS Access database (does anyone remember the name?). P4 will also have a mobile version for tough pads. No one knows if there will be a monthly fee for mobile P4.
- They have been running the Juniper system for 3 years now, but it paid for itself in the first year. **Question: Will the Juniper units work with the new scanners?** Yes, and now they have Bluetooth receivers on the new models (any of the new scanners will have blue tooth). Maybe you could backup your data to a computer or phone via Bluetooth? Matt and John could not handle the capacity of data if the logger was not being used. Otherwise, they could not meet deadlines.
- When you are done, plug unit into computer to download using licensed software. You have to use the file converter after you download your data and then you will end up with an Excel file. For their applications, you end up with 3 tabs of data (native fish, non-native fish, and site data).
- The units are handheld computers so you can use Word or Excel files on the unit and you don't need proprietary software to download these. The Excel or Word files are not auto-saved. You must perform as "Save As" on these files. The auto-save function must be built in to the custom applications Juniper developed. It runs on a Windows Mobile

operating system. The battery lasts a week before you need to recharge or you can pop in a spare battery or a box where you can put in three AA batteries. The Moab group uses the Juniper unit but doesn't use the program. They just use Excel and they have to type in tags. Auto-populating GPS and Tag info is attractive to all users. It has automatic save in the application. If it hangs up, just turn it on and off. You can also use an SD card in the unit. Dave: there is a new unit available on the Juniper site. Tildon: there is also a tablet version.

Discussion

- Oregon RFID: Dave gave John Hawkins a unit to test. It is a pit tag reader. The download from John is that it has some data entry options but they are limited. John thinks it is clunky for what they do, but it could be good for some basic information/limited use. CPW comments that their Plains researchers use them. It has blue tooth capabilities. Units are about \$600. **Question: What about having folks outside of the program use them, like rafters?** Interesting thought. John is going to test it more this summer.
- Dave – Biomark is not servicing cheese blocks anymore. They are buying them back and giving you a \$400 credit towards a new scanner. Dave thinks Biomark can help them out with parts and scanners since they do so much work with them.
- **Question: What are our options for the field now? Cheese block is the standard but it won't be serviced anymore.** Dave – we can buy what we need to buy. There shouldn't be any issues not buying cheese blocks. The program needs to make a decision on when to cut the cord. For now, you can write new units in as line items on your budgets.
- **Question: Do the Juniper units have cameras?** Not sure. The Yuma tablet does have a camera. It is also ruggedized and has a Windows 7 operating system. These cost \$4,000-\$5,000.
- Dave – Juniper website has a tablet with RFID.
- Dave – encourages folks not using data loggers to check them out.
- **Question: how much do we need to worry about these different ways of collecting data?** Kirstin – it will get standardized as we bring it in. We might need to develop a template. After Year 1, when we get data, this can probably be hashed out more. **Question: can we set uploads to auto-match fields and have user manually link unmatched fields?** Kirstin – this seems doable.
- **Question: Can you filter on year, hatchery and river?** Yes.
- **Question: What day of the week might work for monthly meetings?** Hard to get everyone on the same schedule, but Thursday afternoon might work (field crews out during the middle of the week, back at beginning and end of the week).

Thanks Everyone!

ADJOURN