Guide to Creating SBC Datasets with Morpho

Audience: non-IM personnel

Version 0.9
February 2009
Margaret O'Brien

Abstract: Morpho is a metadata management tool created for ecologists to manage their metadata and data for publication. This document describes how SBC scientific staff should use Morpho to keep data packages in our catalog up to date.

This document may be complemented by these other resources:
Morpho User's Guide: http://knb.ecoinformatics.org/software/morpho/
EML Best Practices for LTER sites: coming soon to http://intranet.lternet.edu/documents/
Attributes and Units in LTER Data Packages (v 0.9) http://sbc.lternet.edu/information_management/documentation/ (in developmeot)

Table of contents:

Morpho was created with support from NSF (with additional support from NCEAS) by the Ecoinformatics group at NCEAS (http://knb.ecoinformatics.org/software/)

Introduction

Data and metadata together comprise a "dataset" or "data package". Our goal as a data provider is to publish products in a state appropriate for their intended use, and we understand that completeness, complexity and quality of data and metadata will vary. From its point of view as a data provider, SBC has established levels for our data products which accommodate a) the sources and timing of metadata additions b) data inclusion, c) quality of both metadata and data, and d) potential uses (Table 1). The LTER network has adopted a similar scheme for metadata from the POV of a data consumer (ref, BP doc).

Defining a dataset

It is generally up to the scientists to decide how to package their data, and often the decision is to publish a product convenient for sharing or analysis. Browse the data catalog to see how existing datasets are defined. In general, SBC is a "lumper;" i.e., as long as processing methods and table descriptions remain consistent we combine data tables from multiple years into a large product which can grow as needed. This practice is not required. Our system can handle data split into any unit that is appropriate and we have tools for querying large data tables.

Table 1. Levels of data products with their potential uses (adapted from SBC IM plan, LTER Best Practices, 2004).

Level	Metadata content	Data content	Potential uses
Information	Unique ~50%: Title, short name, abstract, PIs, identifiers, publication date, keywords SBC "boilerplate" ~50%: project description & contact info, area bounding box, access and use statements	Data optional	Information only Display a summary on a web page
Search	**Information** + Temporal coverage Sampling sites listed taxonomic coverage (optional) Additional keywords (optional)	Data optional	1. Search by time, location, taxon, keyword 2. general mapping
Download	**Search** + Describe at least 1 data table Access statements altered (optional)	Data required, Poor formats are allowed	1. Data can be downloaded, although a user might have questions.
Integration	**Download** + Sampling sites as internal references, if data are to included in SBC's query application	Data required, with clean formats. Data matches metadata	1. Integrate data using metadata 2. Query applications 3. Contribute to Network databases

Datasets

Metadata

The metadata format adopted by the LTER network is called Ecological Metadata Language (EML), written in XML schema. Each dataset's description is contained in one EML document. A desktop application called "Morpho" is available for managing metadata in EML and data in tabular format. It was created by the Ecoinformatics group of NCEAS (located on the 3rd floor at the MSB). Morpho has GUI interfaces for nearly all the metadata needed to create SBC's Download-level data products. You will be entering about 75% of this information, and the IM will add the rest.

Saving

Morpho can save your metadata description in 3 ways:
  1. Saved locally (on your desktop) where it can be viewed only through Morpho. This is a common choice as you start out.
  2. Saved to a network metadata catalog, where the most recent revision can be published in a data catalog if you specify (or it can remain only privately viewable).
  3. Exported to the file system. Exported packages are not in the network catalog but are viewable with a browser. This is useful under some circumstances (see the IM for more information).

Usually, you will choose "Locally" and/or "Network". Morpho can also archive the data table as well as the metadata if this is appropriate.

A data package can be saved to a network catalog at any levels (Table 1). A low-level data package (Information" or "Search" level) might be published in order publicize the fact that we are collecting that data even thought a table may not be available

yet. Once a package has been saved to the network, you can still make changes and add the table description later.

Every data package has an identifier and a revision number: e.g., for a package with id=knb-lter-sbc.10.4, the identifier is "knb-lter-sbc.10" and the revision number is the last digit, "4". As you save your work, you'll notice the revision number increasing, and earlier revisions are still accessible. Note that package Ids will change from the initial value assigned by Morpho. The SBC system reassigns them using our standard nomenclature as datasets are finalized. You will still recognize packages by their titles.

Metadata content

In addition to the metadata that is identical in all SBC data packages ("boilerplate"), there is content that can be stored centrally so that you don't have to retype it or copy-paste. At "Information" and "Search" levels, SBC has standard centralized content for our sampling sites and for people, and additional content will become available from centralized sources in the future (e.g, keywords).

When you import a data table (i.e., at the "Download" level), Morpho will be able to deduce some information (e.g., data format), but you will add information about the column names, their descriptions and units (as appropriate), since we don't have a database of measurements yet. So some of your metadata entries will be straightforward; for some you will want to confer with or get approval from the dataset's owners (usually PIs), and for some parts you may need additional help from the IM staff.

Table 2 shows the metadata elements that belong at each completeness level and how they are filled in: either by you with Morpho or by the IM scripts. Table 3 is an empty grid to use as a template as you start to gather information.

Figure 1. Scheme for integrated creation of datasets. Laboratory personnel gather metadata and create packages with Morpho and informration management personnel add standardized information and manage contributions to the data catalog. Arrows indicate areas of the fileserver or catalog affected by the actions.
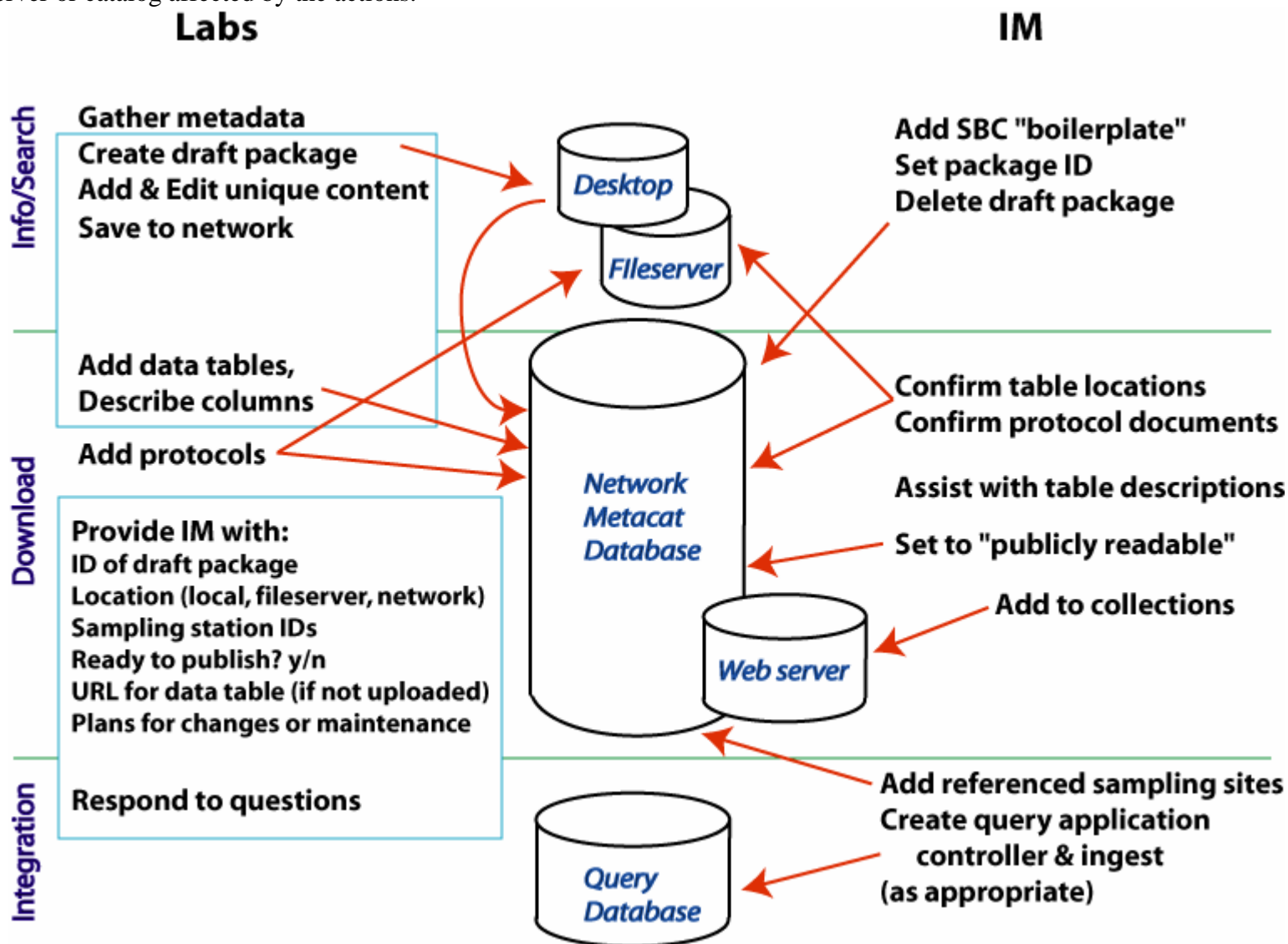
Table 2. EML Metadata elements associated with data package levels

Level	General description	Unique content -- Morpho	Standardized content – IM scripts
Information	Unique content, ~50% (e.g., title, abstract, PIs, keywords) Standardized content, ~50% (e.g., SBC's "boilerplate")	*Required*: title shortName abstract creator(s) [PIs] keywords contact (see notes) *Optional*: associatedParty metadataProvider additionalInfo	packageId datasetId creator[1] keywordSet[1] intellectualRights distribution (information) geographicCoverage[1] contact project access publisher pubDate
Search	**Information** + Temporal coverage Sampling sites listed taxonomic coverage (optional) Additional keywords (optional)	*Required*: temporalCoverage *Optional*: taxonomicCoverage keywordSet	geographicCoverage(s)
Download	**Search** + Describe and include at least 1 dataTable	*Required*: tableName attributeName(s) attributeDescription(s) unit(s) dataFormat URL for distribution (or included table) Protocol(s) *Optional*: additional access rules	from IM scripts if not uploaded with Morpho download link, file physical description
Integration	**Download** + Sampling sites added (for inclusion in SBC's query application)	None	Sampling sites (attribute level), with references to resource level.

Table 3. Empty template of Metadata elements associated with data package levels.

INFORMATION Level						SEARCH Level			DOWNLOAD Level		
Short Name	Id (will change)	title	PIs	abstract	Keywords (gcmd)	Sampling sites	Temporal coverage	Taxa (opt)	Protocol docs (pdf or word doc)	Link to data table	Measurements, described, units

Software Set Up

1. Download and install Morpho
http://knb.ecoinformatics.org/software

2. Set up the software
Create a profile using your LTER network login (not your SBC login).

Go to http://intranet.lternet.edu and look yourself up in the personnel directory. Or see Margaret

Choose File-> Preferences to set the catalog you will use for network saves to http://sbc.lternet.edu/catalog/metacat (although you don't need to save to the network right away.)

3. Get a copy of the "SBCLTER PEOPLE" data package.
      a. Login as yourself
      b. Search for Title "SBCLTER PEOPLE"
      c. Double click to open, and save locally
      d. log out

4. Open up the Morpho Guide:
http://knb.ecoinformatics.org/software/morpho
or click "Help"
It has specific instructions for creating data packages, plus screenshots and examples. The Morpho guide is also available in your local installation directory, e.g.,
C:\Program Files\morpho-1.6.1\docs\user\index.html

Create a Data Package

Click "Create a new data package" from Morpho's main page. There is no need to log in yet; you can keep everything locally at this point.

Morpho will tell you that it is going to request information for these things to start:
-Title and abstract
-Keywords
-People and organizations
-Usage rights
-Research project information
-Coverage details
-Methods and sampling
-Access information

Only a few metadata elements are required to get started. Table 4 shows the screens in Morpho's "New Data Package Wizard" with a list of elements that you should fill in. Also use your filled-in Table 3 as a guide. Required form fields are labeled by Morpho in red, and must be filled in. Any others can be skipped for now, and you can come back to them later. Table 5 covers the screens for describing tables and their columns.

Tables 6-9 (at the end of the document) have more specific guidelines for what SBC data-describers should put into Morpho forms. They are organized by level, and the EML element concerned is given in (parenthesis), although Morpho's dialog boxes do not always show these. Tables 4 and 5 will refer you to a section of these tables for more information.

Table 4. Screens and forms you will encounter in Morpho's "New Data Package" wizard and whether (or how) you should fill them in for SBC datasets. Starred* items can be skipped for now and added later from the "Documentation" menu. Detailed instructions for content can be found in Tables 6 and 7 (for Information and Search levels).

Screen title	action
Title and Abstract*	You must add a title now but an abstract can be added later if you wish. Table 6, #1, #2
Keywords*	To be filled in, but you can skip for now and come back later, Table 6, #5
People and organizations	Owner: at least one is required now. Table 6, #3 Contact: add anything, it will be overwritten (Table 6, #4) Associated Parties: not necessary, but you can add these in the same way as for owners. (Table 6, #4)
Research Project Information	Skip. Added by IM scripts
Usage rights	Skip. Added by IM scripts
Geographic coverage	Skip. Added by IM scripts
Temporal coverage*	Add start and stop dates. If the dataset is ongoing, put the current year in the end date field (Table 7, #2)
Taxonomic coverage*	Skip for most datasets, or see Table 7, #3
Methods*	Add now if you wish, or later when you describe the data table itself. See Table 8 #6
Access information	Choose whether or not the package should be publicly readable. See Table 6, #6 Add the "SBC" user and yourself (Table 6, #6)

Table 5a, 5b. Screens and forms you will encounter in Morpho's "New Data Table" wizard and how you should fill them in for SBC datasets. Find the wizard by choose 'Click here to finish this wizard and add a new data table' or by opening an existing data package and pulling down the menu for 'Data-> create/import new data table'. Table 5a is general information, and 5b contains examples for choosing unit types and units. Detailed information is in Table 8.

Screen title	action
Data Location	you will probably choose either IMPORT: Bring the data file into Morpho, where you'll see it in a spreadsheet-like layout. The data will be archived along with the metadata.This is a good choice for data that is not expected to change or be added to. You'll add documentation for the columns using the wizard, and some will be automatic. or DESCRIBE: The data are NOT included in the data package, but you create the documentation for the data with the wizards just as if you chose import. This is a good choice for ongoing data. We'll create a URL to the data table itself (talk to the IM about the location).
Text import wizard #1	Clicking the box "column names are in the first line" works like it does in spreadsheet imports. The "number of lines to skip" means lines ABOVE the line with column headers.

Text import wizard #2	Table name: the name of the table when it was on the fileserver
	Delimiters: as with a spreadsheet import, choose one or more, and how to treat successive delimiters
Define attribute or column	Name: the column's name. if these were in the first row, Morpho imported it (required) Table 8, #1
	Label: a prettier version of the name Table 8, #2
	Definition: a short description of the column (required) Table 8, #3
	Category: See details in Table 8, #4 and in the document "Attributes and Units in LTER Data Packages".
	If the instructions here, in Table 8, or the "Attributes and Units" document don't make these choices clear for you, then fill in the following information as a work-around for that column, and we'll come back to this table later. These choices will be allowed, but someone who downloads these data will be confused because this description is not really correct for a column that is numeric and needs to have a unit defined.
	Category: "Unordered"
	and in the bottom box,
	Choose (from the dropdown): "Text values (free-form or matching a pattern)"
	Definition (type in the box): "text"

example observations	Unit Type	example units
number, count	dimensionless	number
length	length	meter
biomass	mass	grams, kilograms
matlab datenumber, elapsed time	time	nominalDay, second
temperature	temperature	Celsius
density of kelp or benthic critter	areal density	numberPerMeterSquared
standing crop	areal mass density	gramsPerMeterSquared
phytoplankton density	volumetric density	numberPerMeterCubed
concentration, dissolved nutrients	Amount of substance concentration	molesPerLiter
particulate CHN	mass density	milligramsPerMeterCubed
water currents	speed	metersPerSecond
stream discharge	volumetricRate	litersPerSecond, cubicMetersPerDay
production rate	areal mass density rate	gramsPerMeterSquaredPerYear

Editing and Viewing

Editing in Morpho

When you first enter data into Morpho, you will use wizards. When you edit an existing package, Morpho has two edit modes: wizards and the tree editor, but nearly all your editing should be with the wizards. The tree editor is there only as a last resort, and is really meant for people who understand the underlying structure of EML. If you find yourself in the tree editor and it is not clear what to fill in, feel free to cancel out and contact your IM. The wizards are all available from the menus at "Documentation -> pick-a-section
[note: don't click the "Edit" button on the right. It almost always goes to the tree editor. IM is working with the programmers to fix this behavior]

Viewing the package in the catalog

Probably before or just after you submit your package, a PI or dataset owner will need to approve the package. Or you may need some input from experts on how to describe one of your table's columns. You can have the owner view the dataset in Morpho, or the same metadata can be viewed with a browser after the package has been saved to the network. The advantage of viewing in Morpho is that you can make changes right away. The advantage of a using a browser is that you can send the URL in an email, and the metadata are summarized more concisely.

To view the link:
1. Save the package to the network.
2. Note the package Id (i.e., jdugan.22.5). You don't need to keep track of the last number of the id (the revision) unless you want to view an earlier reversion. Metacat defaults to the most recent revision.

3. If you have chosen to disallow public viewing for this package, you will need to log in before you can view it. Go to http://sbc.lternet.edu/style/skins/dev/login.html

Use this URL as a template, and fill in your package ID in place of "jdugan.22":
http://sbc.lternet.edu/catalog/metacat/jdugan.22/sbclter

If you need to see an earlier version, then add the revision number to the URL:
http://sbc.lternet.edu/catalog/metacat/jdugan.22.2/sbclter

Export

An alternative to saving to the network is to export the package to the fileserver. This is a good alternative when you need to send a link to someone who can't use an LTER network login for some reason. You (or the person you are sending the link to to) will still need to use an SBC-local login however, because you're using a URL to our internal fileserver. Contact the IM for a location to export to. You'll browse to the data package from here. https://sbcdata.lternet.edu/internal/metadata/

Tips, Tricks and Caveats

Speed: Don't log in if you don't have to. When Morpho goes to the network, it can be painfully slow, and it will default to looking for packages at the network if you log in. If you are editing packages located on the network, open all the packages you intend to edit before you save any of them. The first time you click "Open an existing dataset" (from the main screen), Morpho go to the network and the list will take a long time to appear. But when you go back to that screen and click "Open…" again, it will pull the list from memory and be much faster. Then log out and save locally while you work. At the end of your session, log in again and save to the network.

Freezing: If you try to cancel out of a network search, Morpho will probably freeze up. While you wait, go get coffee. Or better yet, go complain to your IM (who can make developers aware of these sorts of problems). It Morpho does freeze, your data still might have been cached and your IM can help you look for it.

Editing: Don't click the "Edit" button on the far right. It goes to the "tree editor" which is not designed for non-IM personnel. This behavior should be changed in the next version.

Submitting the Package to SBC

Talk to your IM

You can submit your package at any level. The IM will add other content, depending on the level and your plans. Information you will need to provide to the IM:

1. The id that Morpho has given the package

2. A short name for the data package. This can be 1 - 3 words, like "kelp NPP", "cruise lter06", or "stream discharge AB00". This could be the name you give the dataset in your own lab. It's a useful way to refer to the data package, since no one remembers the id, and the title is too long.

3. Where it is (your desktop, the network or exported)

4. The names or ids (ids are better) of sampling stations to be added (for integration level)

5. Whether the package is ready for publication (the IM can change these settings as needed.)

6. The URL for the data table if it is not uploaded (you may have to discuss this)

7. Your plans for updates to the data and metadata (i.e., frequency, potential format changes, etc…)

Table 6 Information level fields: Datasets at low levels can be used for general purposes only, since they do not include the actual data tables. At this level, only the dataset titles and people are required. You must add at least one person now, but you can always come back and add more.

Content	Detailed instructions
1. title	Start the title with "SBCLTER:" The rest should be short, but complete, like "Moored CTD and ADCP Data from Arroyo Quemado". In the past we added other parts to the title to distinguish between ocean, reef, etc, but this practice creates confusion since these overlap, or were too general.
2. abstract	A paragraph containing the what, when, where, why. You can put general terms describing the measurements here. PIs generally like to review the abstracts. Some examples: http://sbc.lternet.edu/catalog/metacat/knb-lter-sbc.2005/sbclter http://sbc.lternet.edu/catalog/metacat/knb-lter-sbc.19/sbclter Abstracts are routinely searched and queried, and so are good places to include general terms.
3. PI, dataset owner or creator(s)	These are PIs, or sometimes postdocs. You can 'grab' someone from another data package. At the upper right, choose to "Add a person from another dataset". When the wizard appears, choose the dataset labeled "**SBCLTER PEOPLE**" (if it doesn't show up, bug Margaret). If the person you need is not there, use one of the address templates and edit the name, email and phone number. Templates exist for MSI, BREN, EEMB, ICESS.
4. contact	A "contact" is required by Morpho for each dataset, so you have to put something in the form in one of the red-highlighted boxes. It doesn't matter what you put - it will be overwritten by the IM scripts.

5. keywords	1. At a minimum, we must use the "LTER Core Research Area" keywords in our datasets if they apply. It is fine if no "core area" applies, or if more than one does. Please use these terms exactly as they appear here. Use these keywords only if the dataset actually contains that product, and not if the data are merely related. For example, one might use PAR in computations of primary production, but a PAR dataset should NOT be labeled as "primary production". Spelling counts, but capitalization does not:
	primary production
	populations
	organic matter
	inorganic nutrients
	disturbance
	At "Keywords" click "Add" - you are adding a "Keyword Set"
	Click "Add" again to each keyword in the set
	If you used Core Area keywords, then click the button at the bottom labeled "These keywords are chosen from a pre-defined list", and in the form, paste this text: "**LTER Core Research Areas**"
	2. We have also been adding keywords out of the "Global Change Master Directory" (GCMD) to our datasets. You can see many of these in related datasets (and copy-paste as well). If you use these, put them into a new "keyword Set" and use "GMCD" as the name of the pre-defined list. The entire set of GCMD keywords can be found at: http://gcmd.nasa.gov/Resources/valids/archives/GCMD_Science_Keywords.pdf
	Note: the GMCD keywords are under revision. However, all our current datasets use version 6.0 of the keywords, so we will deal with the switch to a newer version at a later date.
	3. You can also add any other keywords that are necessary. Some glossaries work best with certain data types, or a PI may have his/her own preferences. If you know that a word comes from a particular glossary, put that in as the name of the Thesaurus. If not, then leave that group checked as not belonging to any list (the default). One widely used thesaurus is from NBII (National Biological Information Infrastructure), which can be browsed here http://thesaurus.nbii.gov/portal/server.pt
	4. The LTER Network is considering developing its own set of keywords for searching. If so, these will be available in the future.
	5. You may have noticed that our existing datasets have our sites (places) listed as keywords. The IM scripts will put those in, as necessary.

| 6. Access | 1. If you want the public to be able to view this dataset as soon as it is saved to the network, leaved the default checked.

2. Add 2 users to your data package, and give them "all" permissions. From the LTER tree in the wizard, choose
    1. "**SBC Site Metacat User (SBC)**"
    2. **yourself**

3. You can choose to give specific access right to individuals too, so that someone else can edit this dataset. |

Table 7. Search level details: Datasets at this level have everything from "Information" plus content that allows them to be searched by time and place (and taxon, if appropriate). SBC's bounding box will be added as "boilerplate" and individual stations added later with scripts. You should add keywords, or taxonomy (if appropriate).

Content	Detailed instructions
1. Geographic Coverage (geographicCoverage)	SBC's bounding box will be added as Information-level boilerplate by the IM. You will need to provide a list of sampling stations to be included and these will be added by the IM with scripts. [Note: users can add the SBC bounding box to their local copies of Morpho, although the description will be just the name of the site, instead of the full description. In the menu list, choose "Santa Barbara Coastal LTER (SBC)" and delete it. Then put in these coordinates for north, west, east, south boundaries, respectively: 34.531 N, 120.475 W, 119.255 W, 34.000 N]
2. Temporal Coverage (temporalCoverage)	The start and end dates of the dataset. The start is usually known. If the end date is known and more data are not expected to be added, then put the actual end date in. There are no provisions (yet) for designating that a dataset is ongoing, i.e., that a data table will be appended in the future. In this case, the end can simply be the year, and make sure there is a sentence to this effect in the abstract, such as "this data collection program is ongoing".
3. Taxonomic Coverage (taxonomicCoverage)	Use this section only if you have just a few very important taxa to list – not for all our taxa. For example, the NPP dataset has Macrocystis pyrifera listed as its taxon.
4. Keywords, additonal (keywordSet)	You can add other keyword sets as necessary. See this entry in the "Information" table.

Table 8. Download level details: Datasets at this level have data tables and columns described. Also see the companion document: "**Attributes and Units in LTER Data Packages**"

Content – each column	Detailed instructions
1. Column Name (attributeName)	In database vocabulary, a table column is called an "attribute". If data were arranged in rows instead, then a row name could be also be called a table attribute. If you used Morpho's automatic table description feature and the names are in line 1, then you'll see them already named. If the columns were not imported with names, then you will add them here. For example, you might have a column called temp_c for "temperature in degrees C", or npp_c for "net primary production of Carbon"
2. Column Label (attributeLabel)	If your column name is vague you can include an appropriate term to be used as a label. In this example, the label might be "Water Temperature". Don't worry about the unit that you've attached to the name – Morpho and EML have another place for units. See these links for how labels can make columns prettier: http://sbc.lternet.edu/catalog/metacat?action=read&qformat=sbclter&docid=knb-lter-sbc.10&displaymodule=entity&entitytype=dataTable&entityindex=1
3. Column Description (attributeDescription)	A short description of the column. It is required, even if the name is self-explanatory. If you have not given any indication of a method used, this would be a good place to put that, for example "water temperature with seabird CTD". See the data link under 'Column Label' for some examples of descriptions.
4. Measurement Units (standardUnit or customUnit)	Units in LTER have an entire working group devoted to them. See the document titled: "Attributes and Units in LTER Data Packages".
5. Missing value code (missingValueCode)	What the system should treat as an empty value. You may use more than one, and they do not need to be of the same type as the data (ie, you can have a missing value that is 'NA' even if your data are numeric.

Content – ancillary information	Detailed instructions
6. Protocol(s)	SBC keeps its protocol documents in PDF form, and adds a URL to the file's location in the EML metadata for the entire dataset. The protocol's link belongs is entered under "methods"
7. Additional access rules (optional)	For the most part, our data is viewable by the public and can be downloaded right away. It is conceivable that there might be some cases where we will want to include a table, but not allow it to be downloaded. EML can handle this sort of requirement, and ff this is the case for any of your datasets, see Margaret.

Table 9. Integration level details: You will stop here. Since Morpho does not handle the additions required for SBC's integration level, those additions are left up to the IM. The "Integration" level is required for data to be slurped into SBC's query application (http://sbclter.msi.ucsb.edu/sbc_data_browser), and for data to be reliably handled by scripts that send it to network databases.

Content	Detailed instructions
1. dataTable	If all of the data table's column information is complete and accurate, then you are done. The IM takes it from here, and may contact you with specific questions. Remember that if you have left fields labeled as "Unordered", or "text" when they really should have had a unit defined, then you aren't finished yet. Talk to the IM if you have questions or aren't sure how to fill these fields out.

Quality Control

THIS SECTION IN DEVELOPMENT: We aren't talking about the data values themselves here, only if the data and metadata are correct and well-formatted. Morpho is very capable of finding inconsistencies, but the data and metadata can become unsynchronized. Some of the ways data can get screwed up:

Metadata

Can be wrong (ie, goes with another table, or column
Wrong number of columns

It doesn't match (data gets updated, but not metadata), or you fudged something and planned to fix it later.

Data

Wrong number of elements in each data row (common with excel tables).
Data aren't what you think they are; it's supposed to be a number but a letter got in.
Missing value flag changes from row to row (you can list as many as you need)
.

Table _. Data quality control tasks associated with data completeness levels. In this context, "Quality Assurance" refers to the goal or expectation and "Quality Control" defines the steps required to assure the goal.

Level	Metadata content	Metadata	Data
Information	Boilerplate 50% Unique 50%	*QA*: required fields complete and approved *QC*: 1. minimum EML content required by forms 2. examination and approval by PIs or their staff (via HTML or Morpho) 3. script run to alert missing elements (IM staff)	NA
Search	**Information** + Temporal, taxonomic coverage Sampling sites listed Additional keywords (optional)	*QA*: required fields complete and approved *QC*: examination and approval by PIs or their staff (via HTML or Morpho)	NA
Download	**Search** + Describe and include at least 1 dataTable	*QA*: URL returns data *QC*: If data not uploaded by Morpho, IM script checks location	*QA*: Data resides at URL *QC*: 1. If not uploaded by Morpho, a. IM script checks location b. data table audited for changes and IM staff alerted
Integration	**Download** + Ready for machine reading	*QA*: Metadata matches data with respect to table's physical description, data typing, etc.	*QA*: Clean data formats *QC*: IM script attempts loading data into DB using metadata