|  |
| --- |
| **Use Case:** Creation of metadata after the fact and repairing metadata |
| **Point of Contact:** Sarah Ramdeen |
| **Version: 1.0** |
| **Date: 15 March 2017** |

|  |
| --- |
| **Use Case Name**  Metadata Later |
| **Goal**  To continue to improve the metadata of an object (physical or digital) after the initial collecting event or digitization event. The person who improved the metadata and how and when the improvement was made should be recorded. Any user should be able to find this information and use it to assess the actor. |
| **Summary**  Sarah is a metadata curator at a museum. She has found a mineral specimen that has the latitude and longitude inverted on the specimen card. She wants to fix this error in the electronic system the museum uses for managing specimens. Her department gets funding based on how much work they do, so the system needs to record who did the correction, when, and why.  Michael is a scientist accessing specimen at the museum. He discovers an error in the metadata for an item and wants to report it back to the museum. Part of his evaluation of the error is his expertise in the field/domain. Once the record is amended, it is important to capture when the change was made. If someone else previously published findings on this specimen the record needs to reflect that a change happened after the publication. Michael reports the error to the curator (Sarah) for correction. |
| **Actors**   * Metadata Curator, Sarah * Scientists, Michael |
| **Preconditions**   * There already exists a system that contains a digital object (could be a digital representation of a physical object) * The metadata curator logs in and finds the object and metadata of interest. * Sarah has an ORCID * A system of specimen identifiers is in place, such as IGSN |
| **Post Conditions**   * Behind the scenes, the system records the correction, the reason/evidence for the correction, the identity of the curator, and the date/time of the correction. The previous incorrect data are also kept, but deprecated. * Sarah is returned to the splash page for the digital object she was curating. There is a visual evidence that the change was made. * Outsiders viewing the record should be able to see that changes have been made and when they were made. |
| **Triggers**  The user locates incorrect data and clicks “edit” to make the correction |
| **Normal Flow**   1. Sarah is presented with an editable version of the the incorrect data. 2. She makes the correction. Hits “ok”. 3. The editable version of the data disappears 4. She is prompted for the reason for the correction and any related evidence (possibly a controlled vocabulary) 5. She adds the reason for the correction. Hits “ok” |
| **Alternate Flows**   * Every “ok” or “edit” button should also be accompanied by a “cancel” button that takes the user back to the previous page (likely the digital object splash page) |
| **Exception Flows**   * If no evidence or reasoning is given for the change, the system should remind the user that this information is required and give another chance to add the information. |
| **Entities**  Person (Curator) - In this case, Sarah  Action - In this case correcting inverted numbers  Thing that was acted upon - In this case a latitude/longitude pair  Reason for the action - In this case whatever gave Sarah the idea that the lat long was inverted.  DateTime of the Action  Notes - any important notes? |
| **Properties**  PerformedBy and Performed |
| **Diagram**  **RDA_TDWG_diagram_1.jpg** |
| **Priority** |
| **Notes**  The information should be changed by the curator, but there should also be a mechanism for reporting error by the end user to the curator or manager of the system. This is different than evaluating quality. |

|  |
| --- |
| **Use Case:** Curation of digital objects (video, audio, or images) |
| **Point of Contact:** Wim Hugo |
| **Version: 1.0** |
| **Date:15 March 2017** |

|  |
| --- |
| **Use Case Name**  Digital Objects |
| **Goal**  Keep track of a digital object and any processing that has been performed upon it. The processors should be able to get credit for adding value to the digital object. |
| **Summary**  Kenji is an academic oceanographer who is publishing a paper about primary productivity in the open ocean. He performed analyses on several publicly-available satellite images in order to improve resolution so he could use them in his paper. The improved images were then resubmitted back to the original repository. Maria, the data manager at this repository (of satellite images) notices that these two images are related to each other and wants other users to easily find this out and know that Kenji did the work. She wants to link the images in her repository and add metadata describing who did what to the original image that resulted in the improved image.  Alternative:  Researcher A records a large volume of video with an automated underwater vehicle. The main focus of current research is to determine the abundance of a specific species of pelagic fish that has economic value to the community. To achieve this, the researcher develops automated methods for frame-based image recognition and publishes a paper on the relative abundance of the species.  Some time later, researcher B hypothesizes that certain pelagic fish species alter their schooling behaviour based on proximity of predator species, and analyses the same video sample collection with motion tracking and species recognition software. The only link between the two sets of research in formal terms will be a shared sampling event (digital sample) identifier. |
| **Actors**   * Kenji, oceanographer * Maria, data manager * Researcher A: species abundance expert * Researcher B: species interaction expert |
| **Preconditions**   * There already exists a system that contains a digital object * The data manager logs in and finds the object and metadata of interest. * Kenji already has an ORCID * A system of persistent, unique, dereferenceable IDs for digital objects is in place, like DOI * The repository system is capable of generating citable references for users to copy into their publications for each digital object. * A publicly available resolver exists for citations of digital samples/ sampling events. |
| **Post Conditions**   * A citation is available for users that gives credit to Kenji for the analysis and the repository for stewardship, and enables the unambiguous identification of the image in question by a third party * A citation is available that credits researcher A with the sample and its curation. |
| **Triggers**   * Maria notices the two related images that are not formally connected in her repository * She finds the splash page for the derived image * Maria clicks on a button that says “Make Connection”. * Researcher A needs a citation index on samples and sampling events to demonstrate the value of their sampling programme and automated equipment to funders. |
| **Normal Flow**   1. A pop up window appears with the ID and a thumbnail of the derived image on the left. Moving right there is an arrow pointing to the right and then a text box with two buttons underneath, “Connect” and “Cancel” 2. Maria begins typing the ID of the original image in the text box and it is autofilled 3. Maria clicks “Connect” 4. The ID and a thumbnail of the original image appears on the right. A drop down list appears underneath the arrow. 5. From this dropdown list, Maria chooses the relationship “DerivedFrom” 6. Two buttons appear, “OK” and “Cancel” 7. Maria clicks “OK” 8. The first popup disappears and a second appears asking for metadata about the relationship she just added 9. The exact configuration of the metadata window depends on the relationship chosen previously. For DerivedFrom, the system asks for an actor(s), which is added via ORCID, the action from a controlled vocabulary, and a reference which in this case would be the doi of Kenji’s paper. The system would autoload the date and time of the upload of the derived image, but would allow Maria to edit. 10. There are two buttons at the bottom “OK” and “Cancel”. Maria clicks “OK” 11. She is presented with a pop up review screen and clicks “Confirm” (there is also an “Edit” button) 12. The system automatically adds that Maria is the one who created the “DerivedFrom” relationship on that day.   Alternative:   1. A globally available digital sample/ sample event registry and resolver service is available. The registry supports basic services:    1. Register a digital sample or sampling event by way of a universal metadata record describing the event. This act of registration creates a sample PID (SID).    2. Synchronise SIDs with domain-specific registries (for example GBIF - GBIF does not register digital samples/ events at present, as far as I am aware, but does this for physical specimens)    3. Resolve a reference to a SID by linking the institutional metadata record for the sample, and to the digital object itself. 2. The service mines “citations and mentions” of the SIDs in journal articles and data publication registries (e.g. DataCite) and builds a SID citation index. 3. The service links, where possible, ORCIDs and DOIs to SIDs and de-duplicates the links. |
| **Alternate Flows**   * Maria can click “Cancel” in any of the windows and the pop up disappears * Clicking “Edit” in the final review screen will take the user to a single popup where everything is editable with an “OK” and a “Cancel” button. If the user hits “OK” the review screen reappears with the changes. If the user hits “Cancel” the pop up disappears and no changes are made. |
| **Exception Flows**   * If any of the boxes are left blank, when the user clicks “OK”, the system should remind the user that all information is required and then take he/she back to the editing window. |
| **Entities**  Person doing the action - In this case Maria  Action 1 - In this case connecting two related digital images  Objects that receive the action - In this case the digital images.  Person that performed an action on one of the images - In this case Kenji  Action 2 - In this case the modification of one image into another by Kenji  Publication - In this case the paper Kenji published using the derived image  Digital Sample/Sampling Event  Researcher A, took the video collection and analyzed it  Researcher B, analyzed the same collection later for a different purpose  Equipment - In this case the vehicle and the video equipment  Publication  Video and Video Collection |
| **Properties**  DerivedFrom  Authored hasAuthor  PerformedBy Performed |
| **Diagram**  RDA_TDWG_diagram_2.jpg |
| **Priority** |
| **Notes** |

Note: It may be possible to

* Extend the DataCite metadata schema to account for samples and sampling events
* Convince DataCite to allow registration of samples and sample events as pseudo-data objects. I do not see any reason why this will not be possible.

## DataCite Mapping

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Metadata Element** | **Cardinality** | **Mandatory** | **Purpose** | **Notes** |
| Identifier | 1 | Yes | Citation | DOI |
| AlternateIdentifier | 0-1 | Yes | Re-Usability | Do you have an identifier for the object? |
| Title | 1 | Yes | Citation | Title |
| Creator   * Affiliation * Identifier | 1-n | Yes | Citation | Type is ‘Author’ by default, and can be an institution instead of an individual. Maybe use institution here by convention for samples |
| Publisher | 1 | Yes | Citation |  |
| PublicationYear | 1 | Yes | Citation |  |
| Contributor   * Affiliation * Identifier | 0-n | No | Citation | Type=  DataCollector  DataCurator  HostingInstitution  ... |
| Subject   * Scheme * Authority | 0-n | No | Discoverability  Coverage | Scheme and Authority can be used to accommodate discipline-specific vocabularies for sample types, event types, protocol types, instrument types, etc. |
| Date   * Type | 0-n | Yes for Samples | Discoverability  Coverage | Date of sampling event is critically important  Date types= |
| GeoLocation   * Type | 0-n | Yes for Samples | Discoverability  Coverage | Type=  Box  Point  *May need extension to account for polygons, lines or polylines (transects or voyages), 3D volumes.* |
| ResourceType   * Type | 0-n | Yes for Samples | Discoverability | Open, but can also be a type.  Applicable Types for (Digital) Samples=  Audiovisual  Collection  Event  Image  PhysicalObject  Sound |
| Description   * Type | 0-n | Recommended | Reusability | Types=  Abstract  Methods  SeriesInformation  TableOfContents  Other |
| RelatedIdentifier   * Type | 0-n | Recommended | Reusability | Use to link to other registries. Supports the following and should be extended for existing and future physical sample registries such as GBIF, iSamples, …  ARK arXiv bibcode DOI EAN13 EISSN Handle ISBN ISSN ISTC LISSN LSID PMID PURL UPC URL URN |

|  |
| --- |
| **Use Case:** Performance and collections review |
| **Point of Contact:** Jason and Vince |
| **Version: 1.0** |
| **Date:15 March 2017** |

|  |
| --- |
| **Use Case Name**  Performance Review |
| **Goal**  Make the impact of curatorial activities quantitative for the purpose of curatorial staff performance review and assessing the resources needed to curate a collection |
| **Summary**  Ben is a collections manager in a natural history museum. He has a staff of five curators whose performance he must review annually. He has his own performance reviewed by a superior who is particularly interested in the overall state of the collections he oversees. He knows he doesn’t have enough staff to quickly curate all of the collections in the museum, so there is much curatorial work to be done that Ben thinks is important for his own review. Unfortunately, the only metric widely available for review is publication record. While publications are important, Ben wants his curators to feel comfortable spending time on curation activities instead of only publications. He would like to keep track of the curation activities of his employees and generate a “curation metric” that can be used in a review or on a CV. Sergey, one of the curators, has spent all year sorting and identifying (including DNA barcoding, entering each specimen into the museum’s specimen management software, and placing each specimen in its own container with a new label) small specimens in a large jar that were collected during a remote expedition over 100 years ago. He discovered a new species and was able to publish it, but Ben would like Sergey’s performance evaluation to reflect the large and important task of sorting through the entire jar, not just publishing the one new species. During his own performance review, he would like to be able to ask for the resources he needs to properly maintain the collections he oversees and have data to back up that request. |
| **Actors**   * Ben, collections manager * Sergey, curator |
| **Preconditions**   * The museum uses specimen management software that powers an online portal where third parties can search for specimens the museum holds * Sergey has logged in to this system during the course of his work and logged his activities * Sergey has an ORCID (ISNI - dead people) * Every specimen in the system has a persistent, unique, and dereferenceable ID, such as IGSN * Ben logs in to museum specimen management software as a manager |
| **Post Conditions**   * When Ben is done reviewing his curators’ work, he can exit the system and log out. |
| **Triggers**   * He can see a menu item “Review” under which he chooses “curators” |
| **Normal Flow**   1. Ben is given the option to search or browse the curators in the system. The best option will depend on how many curators are present 2. Ben chooses the curator of interest by clicking on the person’s name. 3. Ben can see a high level summary of the person’s curatorial activities 4. Ben can click through to see a more detailed version of the person’s curatorial activities |
| **Alternate Flows**   1. Under the “Review” menu item, Ben can choose “specimens” instead of “curators” 2. Ben is given the option to search or browse the specimens in the system. The best option will depend on how many specimens are present 3. Ben chooses the specimen of interest by clicking on its ID. Metadata about the specimen are present to help Ben find the specimen he is looking for. 4. Ben is shown a high level summary of how much the specimen is used |
| **Exception Flows**   1. If Ben searches for something that isn’t present, he is presented with a message saying “no results”. The search box is still available for him to try again. |
| **Entities**  Person - In this case Sergey  Action - In this case a variety of specific curatorial tasks  Publication - In this case the publication describing the new species  Objects that receive the action - In this case the specimens  DateTime - |
| **Properties** |
| **Diagram** |
| **Existing Terms** |
| **Priority** |
| **Notes** |

|  |
| --- |
| **Use Case:** Linking samples/specimens to other things that are related to them |
| **Point of Contact:** Sean and Sarah and Denise |
| **Version:** |
| **Date:** |

|  |
| --- |
| **Use Case Name**  Linking specimens |
| **Goal**  Make it easy to find data and research products related to a specimen/sample and keep track of a specimen/sample’s history and impact. |
| **Summary**  Victoria studies the history of climate change on Earth. An important source of data for her are sediment cores collected from old lakes. Many of these cores are stored in a special facility. They are expensive to collect and preserve and many analyses require destruction of a small part of the core. Thus, it is very important to be efficient with analyses. Victoria needs data from a specific lake to finish her work and has found a few publications about cores from that lake. She is able to find a core in the storage facility that would have the information she needs, but it is difficult to tell what kind of analyses have been run on it already, if any. She can’t be sure if this is the core that is discussed in the publications she found. Victoria would like to be able to discover information about the core through the storage facility web site and add her contribution to the analysis for others to find. |
| **Actors**   * Victoria, scientist |
| **Preconditions**   * There is a core storage facility that is searchable online. All the cores have a unique identifier, IGSN or similar. * Victoria has an ORCID * Victoria has searched the core facility and found a core she is interested in * Victoria is on the splash page for that core |
| **Post Conditions**   * Once Victoria has seen the metadata and analytical history of the core, she can trigger an action (clicking?) to request use of the core * If Victoria requests use of the core a message is sent to the core facility. If the request is granted evidence is visible in the analytical history of the core. |
| **Triggers**   * Triggers an action (clicking?) to view the entire analytical history of the core |
| **Normal Flow**   1. Victoria sees that there have been three analyses performed on parts of this core and two of those were included in a publication. 2. Victoria clicks on each analysis to see the specific metadata for each. This includes the people involved, the methods, the funding agency, and any publications or other data products. 3. She sees that one analysis she needs has not been performed on this core 4. She requests use of the core from the storage facility. This process proceeds in accordance with the facility’s policies. 5. After her analysis is done, she goes back to the splash page for the core and clicks on a button to add information. A pop-up form appears for her to add information about the analysis and any products. |
| **Alternate Flows**   * The core Victoria inspects is not what she needs. She clicks on a button to go back to search |
| **Exception Flows** |
| **Entities**  Actor(s) - In this case Victoria and any other researcher that has analyzed the core  Core  Analysis  DateTime  Publication |
| **Properties** |
| **Diagram** |
| **Priority** |
| **Notes** |

|  |
| --- |
| **Use Case:** citizen science, students contribution |
| **Point of Contact:** |
| **Version:** |
| **Date:** |

|  |
| --- |
| **Use Case Name** |
| **Goal**  *Citizen scientists and volunteers can get credit for their contributions* |
| **Summary** |
| **Actors** |
| **Preconditions** |
| **Post Conditions** |
| **Triggers** |
| **Normal Flow** |
| **Alternate Flows** |
| **Exception Flows** |
| **Entities** |
| **Properties** |
| **Diagram** |
| **Priority** |
| **Notes** |

|  |
| --- |
| **Use Case:** Monitor professional contribution |
| **Point of Contact:** Jon Norenburg |
| **Version:** |
| **Date:** |

|  |
| --- |
| **Use Case Name**  Professional Taxonomy |
| **Goal**  Professional taxonomist wants to keep track of the impact his specimens have on science |
| **Summary**  Joe is an invertebrate taxonomist at a natural history museum. He has spent decades collecting specimens and publishing new and revised taxa. The current system of professional rewards requires metrics that describe his contribution to science, including publication records. Joe has a page in ImpactStory that displays the impact of his publications, but nothing about his specimens. He knows that other taxonomists have viewed his collections on multiple occasions, but is not sure how his specimens have been used. |
| **Actors**   * Joe, a taxonomist |
| **Preconditions**   * Joe has an ORCID * A system of unique, persistent, dereferenceable identifiers for specimens is in place |
| **Post Conditions**   * A JSON output is available for use by ImpactStory |
| **Triggers** |
| **Normal Flow** |
| **Alternate Flows** |
| **Exception Flows** |
| **Entities**  Joe  Specimens  Publications  Other Taxonomists  Taxa |
| **Properties** |
| **Diagram** |
| **Priority** |
| **Notes** |

|  |
| --- |
| **Use Case:** Mass curatorial actions on databases |
| **Point of Contact:** Vince Smith |
| **Version:** |
| **Date:** |

|  |
| --- |
| **Use Case Name**  Mass actions on databases |
| **Goal**  Maintain the provenance and attribution of curatorial actions on whole databases |
| **Summary**  Sophie is a database manager at a museum. She is regularly given large data files by the scientists who work at the museum for upload into the database. These files often require a high degree of manipulation before they can be uploaded. She is also responsible for guaranteeing the quality and availability of the database. That means she needs to run error-finding algorithms and occasionally needs to update the database software. These processes are important and need to be attributed to her. |
| **Actors**   * Sophie, database manager |
| **Preconditions** |
| **Post Conditions** |
| **Triggers** |
| **Normal Flow** |
| **Alternate Flows** |
| **Exception Flows** |
| **Entities** |
| **Properties** |
| **Diagram** |
| **Priority**  Identified as low priority at TDWG 2016, RDA P9 - worth including |
| **Notes** |

|  |
| --- |
| **Use Case:** Digitization |
| **Point of Contact:** Vince Smith |
| **Version:** |
| **Date:** |

|  |
| --- |
| **Use Case Name**  Digitization |
| **Goal**  Keep track of which specimens were digitized by whom |
| **Summary**  Employ teams of people to digitize a physical collection and create records about objects. Scanning herbarium sheets. Photographing pinned insects. 3D scanning of fossils. They are the point of contact for applying the identifier i.e. adding the barcode. Especially for the legacy specimens. Transcribing information about specimens from specimen labels. |
| **Actors** |
| **Preconditions** |
| **Post Conditions** |
| **Triggers** |
| **Normal Flow** |
| **Alternate Flows** |
| **Exception Flows** |
| **Entities**  Digitizer  Physical Object Being Digitized  Digital Object derived from physical object |
| **Properties** |
| **Diagram** |
| **Priority** |
| **Notes** |

|  |
| --- |
| **Use Case:** Funder |
| **Point of Contact:** Alex (iDigBio) |
| **Version:** |
| **Date:** |

|  |
| --- |
| **Use Case Name**  Funding Impact |
| **Goal**  Demonstrate the impact of research funding |
| **Summary**  Research funding agencies and foundations must show that they are spending money effectively and having an impact. The products of a research grant must be identifiable and their use must be documented. Martin is an intern at a foundation that funds collecting expeditions. He has been asked to work on the annual report to the board that demonstrates how their funds were used. Martin’s boss wants him to locate all the research products of a specific expedition and find all instances of those products contributing to the increase in scientific knowledge. |
| **Actors**   * Martin, intern |
| **Preconditions** |
| **Post Conditions** |
| **Triggers** |
| **Normal Flow** |
| **Alternate Flows** |
| **Exception Flows** |
| **Entities**  Funder (Agency, organization, etc)  Expedition  Specimens collected during expedition  Publications  Really we want the impact of those specimens (See Professional Taxonomist use case) |
| **Properties** |
| **Diagram** |
| **Priority**  *<out of scope - maybe rewrite* |
| **Notes** |

|  |
| --- |
| **Use Case:** Transforming a digital or physical object - two use cases (annotating and transforming are two different things) Artl@s |
| **Point of Contact:** Sarah Ramdeen |
| **Version:** |
| **Date:** |

|  |
| --- |
| **Use Case Name**  Object Transformation |
| **Goal**  Keep track of the transformation of an object in a collection. |
| **Summary**  Drawing on a map. Restuffing. Applying preservative. Making decisions about - do we get DNA or do we preserve the anatomy? Could involve destructive sampling  Ruth is curator for a collection of materials related to snow and ice. One of the items in her collection is a scientific map. She discussed the map with members of the indigenous community and they were able to give her additional information from their historical knowledge of the area. She has the new information as an overlay - hand written annotations on a copy of the map. THe transformed object needs to be added to the collection records with the connection to the original documented. Ruth must consider: does the record just represent the annotations or both the annotations and the map? How is the record structured? Thankfully there are guidelines for how to create a record for such an instance. The system allows for the connection of the two records as related/associated items. |
| **Actors**   * Ruth, curator of the collection |
| **Preconditions**   * The system allows for records to be connected with a defined relationship |
| **Post Conditions** |
| **Triggers**  *Something changed with an item in the collection. This may be the addition of annotations but includes such variations as re-housing samples which may result in a secondary related object which will need a record in teh database* |
| **Normal Flow** |
| **Alternate Flows** |
| **Exception Flows** |
| **Entities**  Person doing the transformation  DateTime - Start time and end time? Restuffing can take a while.  Reason for the transformation  Transformation action |
| **Properties** |
| **Diagram** |
| **Priority** |
| **Notes** |

|  |
| --- |
| **Use Case:** Nomenclature Curation |
| **Point of Contact:** Nicky Nicholson |
| **Version:** |
| **Date:** |

|  |
| --- |
| **Use Case Name**  Nomenclature |
| **Goal**  Give attribution for the maintenance and research of nomenclatural acts that support authoritative classifications |
| **Summary**  Sharon is a curator at an herbarium who works as part of a team to maintain an authoritative plant classification. This classification is used by many other institutions and projects as essential infrastructure. The impact of her work and the work of her team is huge, but no one knows that she is behind the scenes. This can be a problem when funding is up for renewal. How can Sharon really know her contribution and communicate that to her funder? |
| **Actors**   * Sharon, curator |
| **Preconditions** |
| **Post Conditions** |
| **Triggers** |
| **Normal Flow** |
| **Alternate Flows** |
| **Exception Flows** |
| **Entities**  Person doing the action, in this case Sharon  Curatorial act  DateTime  Reason |
| **Properties** |
| **Diagram** |
| **Priority** |
| **Notes** |

Add a use case from the institution point of view so a department or museum can get credit and assess their impact

iDigBio has a quasi-list of attribution roles related to work flows

IGSN roles - Anytime there is mentioned a system of identifiers for specimens, I am thinking IGSN