



Guidelines on the Implementation of the Traditional Knowledge and Biocultural Labels and Notices in the European Reference Genome Atlas for Biodiversity Researchers

Version 1

Updated: 31st August, 2022

Authors: Ann M. Mc Cartney, Astrid Böhne, Janette Hamilton-Pearce, Ashley Rojas, Libby Liggins, Corrie Roe, Rob Davey, Felix Shaw, Stephanie Russell-Carroll, Maui Hudson, Jane Anderson

Section 1: Recognizing Indigenous Peoples and Local Communities in The European Reference Genome Atlas	3
Section 2: Disclosing IPLC Rights and Interests in Biodiversity Data	4
Section 3: The Traditional Knowledge and Biocultural Labels and Notices	5
Section 4: Researcher Workflow within the Local Contexts Hub	7
Box: Which Account type is Right For Me?	7
Section 5: What is the Right Scope or Scale For My Project?	9
Section 6: Which Communities Should I Notify?	10
Section 7: What Else Should I Do After Applying a Notice?	11
Section 8: How should I implement a Notice/Label Into My ERGA Metadata Manifest?	11
8.1. Guidelines for ERGA Managers and Sample Ambassadors	12
Acknowledgements	13



Section 1: Recognizing Indigenous Peoples and Local Communities in The European Reference Genome Atlas

ERGA, the European Reference Genome Atlas (<https://www.erga-biodiversity.eu/>), the European hub of the Earth BioGenome Project (<https://earthbiogenome.org/>), is a moonshot proposal to generate high quality reference genomes for the entire eukaryotic biodiversity across Europe. ERGA is a fast-growing consortium with an expanding member base spanning EU member states, EU Associated Countries, and other countries within the European bioregion. A major goal of the ERGA is to develop a pan-European, inclusive, genomics, informatics and bioinformatics network structure that enables the equal and equitable participation of each country, institution and Peoples throughout each step of the reference genome production pipeline. ERGA is committed to responsible data sharing and management practices amongst participating researchers and aims to informatically support robust, and streamlined metadata collection processes to ensure that all genomic data produced by the project is associated with comprehensive, and appropriate metadata. Responsible metadata 1) maximizes scientific value, 2) minimizes sampling redundancy by connecting genomics information to samples in ex-situ collections, 3) supports ethical and legal access and utilization, and 4) promotes fair and equitable benefit-sharing and attribution.

From the outset of ERGA, significant efforts have been placed on streamlining the reproducible collection and safe/documented storage of comprehensive, ethical and legally compliant metadata for all genomes and transcriptomes generated by ERGA - including from the lands of Europe's Indigenous Peoples and Local Communities (IPLC). ERGA is also acutely aware of the ongoing and fraught history of harms many IPLC have, and continue to experience, as a result of biodiversity-related research and respects the right of any IPLC to decline or defer participation in the project. However, if IPLC do choose to participate, it was important that ERGA had the appropriate infrastructure to support systemically IPLC inclusion, participation, recognition and attribution across the project.

Through synergistic collaborations and partnerships across the ERGA Samples and Sample Processing Committee, ERGA Ethics, Legal and Social Issues Committee, COPO (<https://copo-project.org>) and Local Contexts (<https://localcontexts.org>), ERGA has built the first streamlined process for recognizing IPLC rights by creating novel validatable metadata fields that support the inclusion of both Traditional Knowledge and Biocultural Labels and Notices (see Figure 1). By informatically supporting the streamlined inclusion of the Labels within the metadata manifest and SOP (<https://github.com/ERGA-consortium/ERGA-sample-manifest>) each ERGA member has both the responsibility and opportunity to place a Notice on any samples, data or associated Traditional Knowledge collected as a standard of practice.

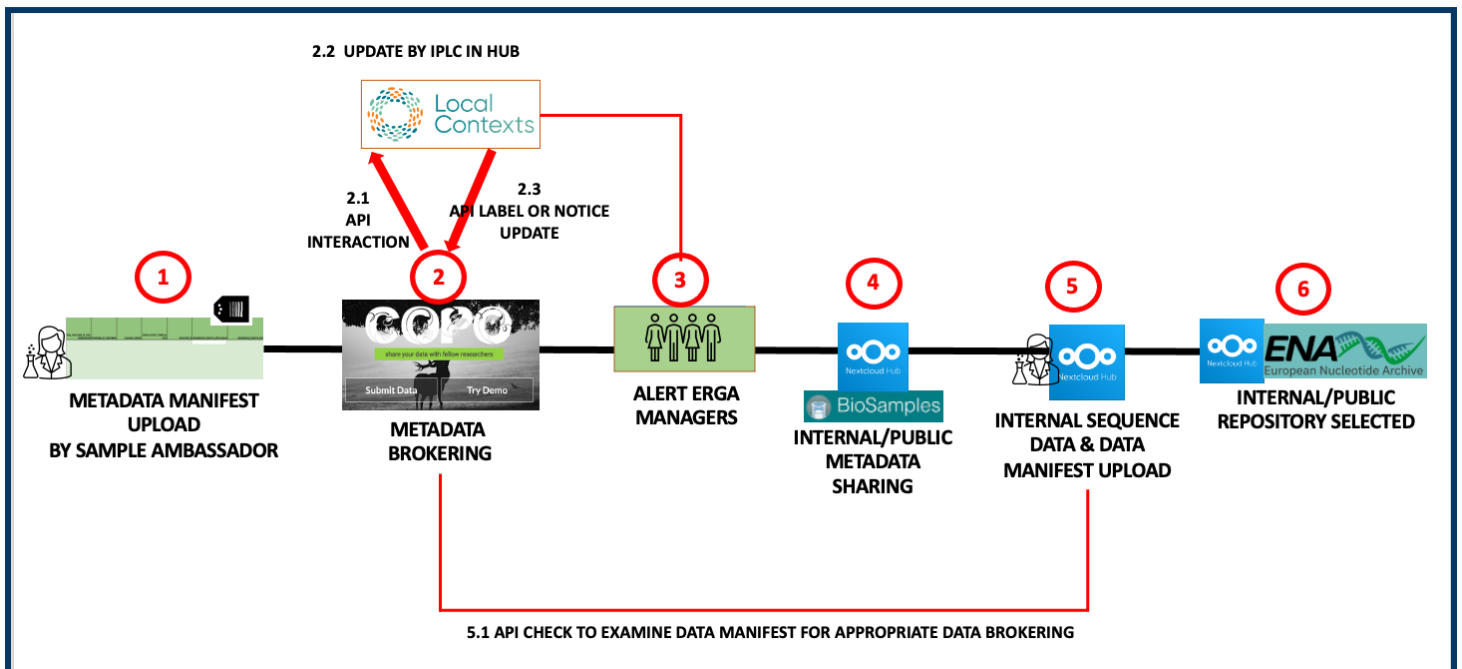


Figure 1: ERGA Label and Notice Implementation Workflow

Section 2: Disclosing IPLC Rights and Interests in Biodiversity Data

Genetic resources are subject to international agreements including the Convention on Biological Diversity (CBD), the FAO ITPGRFA, PCOV, UNCLOS and the WHO PIP Framework. The rights of IPLCs were specifically addressed through the supplementation of the Nagoya Protocol (NP) which supports Fair and Equitable Benefit-Sharing from Genetic Resources (Bavikatte & Robinson, 2011; CBD 1992: Art 8(j)). Similarly, scientific institutions are increasingly promoting equity, diversity and inclusion as key goals to improve the impact of science and research to society.

IPLC lands and resources are regularly a source for biodiversity samples from which to generate reference genomes. As we increasingly recognize the value that metadata brings to scientific endeavors and our understanding of the natural world, we need to consider how IPLC relationships to species can be included within our data frameworks. The inclusion of IPLC provenance metadata in genomic databases is seen to support appropriate attribution, the recognition of IPLC interests in relation to their genetic resources, and provide pathways for local access and benefit-sharing models (Garrison et al., 2019; Hudson et al., 2020; Ambler et al., 2021; Sherkow et al., 2022).

ERGA has included the following fields within its metadata manifest to support the inclusion and recognition of critical contextual metadata associated to be collected for IPLC samples;

- ASSOCIATED_TRADITIONAL_KNOWLEDGE_OR_BIOCULTURAL_RIGHTS_APPLICABLE
- INDIGENOUS_RIGHTS_DEF
- ASSOCIATED_TRADITIONAL_KNOWLEDGE_OR_BIOCULTURAL_PROJECT_ID
- TRADITIONAL_KNOWLEDGE_OR_BIOCULTURAL_CONTACT

- ETHICS_PERMITS_REQUIRED
- ETHICS_PERMITS_DEF
- SAMPLING PERMITS_REQUIRED
- SAMPLING PERMITS_DEF
- NAGOYA_PERMITS_REQUIRED
- NAGOYA_PERMITS_DEF

Researchers may not be aware of the specific interests of IPLCs and in some cases may not have been in direct contact with those communities. IPLCs are best placed to provide appropriate provenance information and the protocols and permissions that they associate with the genetic resource. However, as researchers are primarily involved in the generation and submission of reference genome data, it is their responsibility to disclose the presence of IPLC interests (Anderson & Hudson., 2020; Liggins et al., 2021; McCartney et al., 2022). The Local Contexts Hub is a web portal that allows researchers to apply Biocultural or Traditional Knowledge Notices to biodiversity genomic and transcriptomic data in order to provide an explicit disclosure of IPLC interests in the genetic resource. This addresses the ethical expectations of IPLCs and enables access and benefit-sharing mechanisms to be applied, as appropriate, to genetic sequence information or Digital Sequence Information (DSI) as it is known amongst policymakers.

Section 3: The Traditional Knowledge and Biocultural Labels and Notices

Engagement with IPLCs is through customized Traditional Knowledge (TK) and Biocultural (BC) Labels. There are [20 TK Labels](#) and [10 BC Labels](#) that can be grouped into three categories (see Figure 2.1):

- Provenance Labels identify the group or subgroup which is the primary cultural authority for the material, and/or recognizes other interests in the material;
- Protocol Labels outline traditional protocols associated with access to this material and invite viewers to respect community protocols;
- Permission Labels indicate what activities the community has approved as generally acceptable. Other uses require direct engagement with primary cultural authorities.

IPLCs can embed their local and traditional protocols for how their biodiversity data is used, displayed, managed and accessed going forward into data infrastructures. Although the BC Labels are mainly for biodiversity data, the IPLCs can also apply the TK Labels.

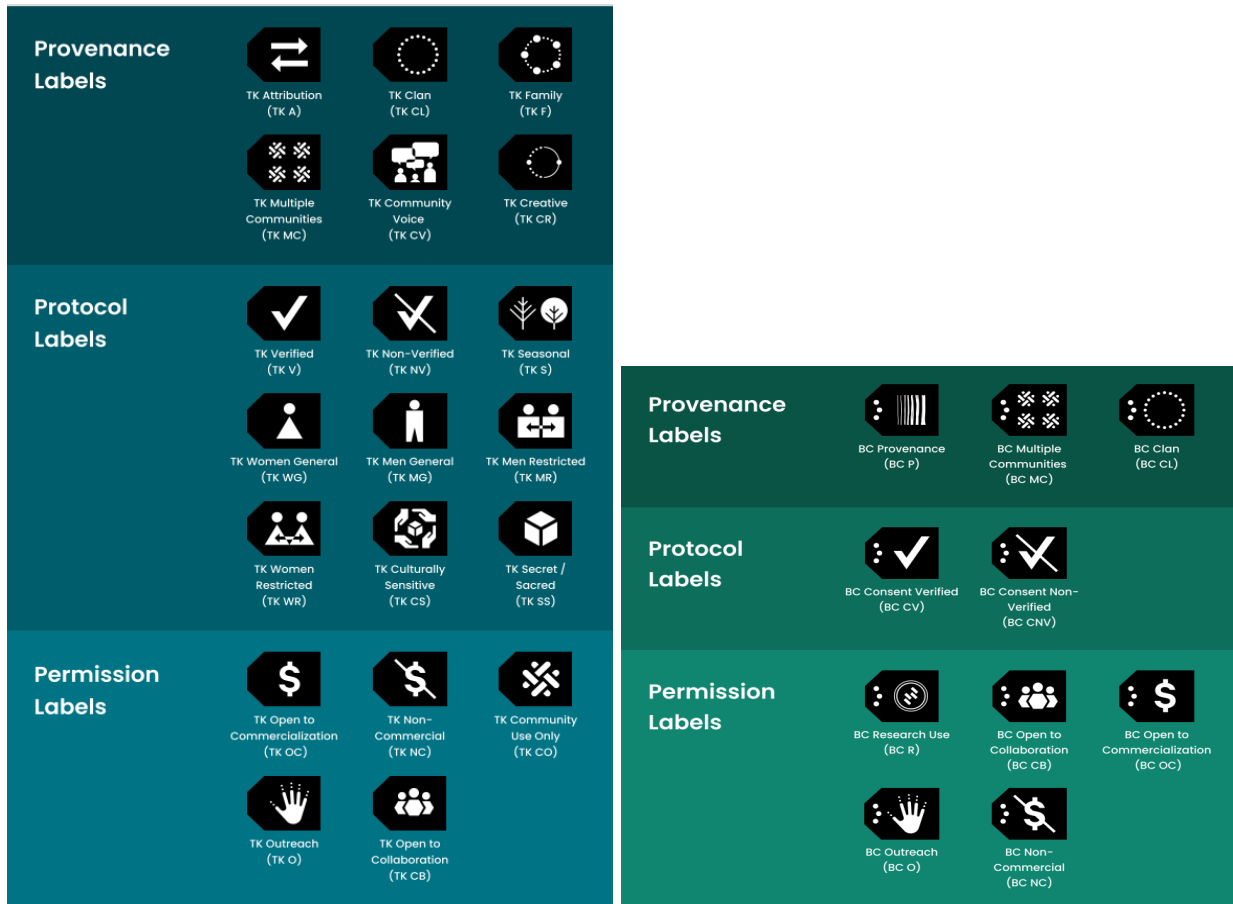


Figure 2: Overview of the available Traditional Knowledge (TK) and Biocultural (BC) Labels

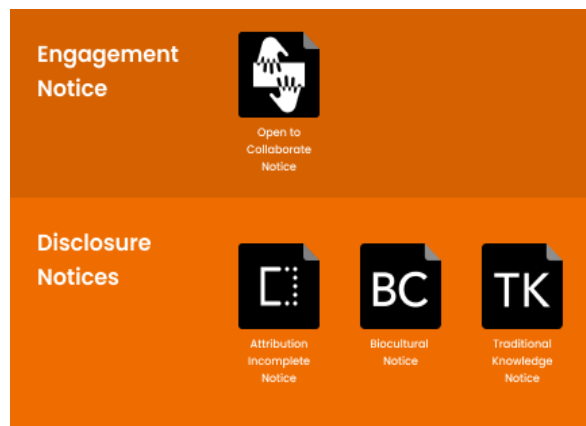


Figure 3: Overview of the available Engagement and Disclosure Notices

The Labels and Notices are interoperable with Darwin Core as they can be mapped to the pre-existing rights field (Table 1, Figure 4).

Table 1: Interoperable Metadata: This table highlights the appropriate Rights metadata fields [Darwin Core] for Labels and Notices.

LABELS			NOTICES		
Field attribute	Value	Example value(s)	Field attribute	Value	Example value(s)
Rights	The custom Label text created by a community for this collection/item using the <u>Local Contexts Hub</u> . For clarity, this should be preceded by the name of the Label. Do not include the generic URL for the Label.	Local Contexts TK Attribution: This Label is being used to correct historical mistakes or exclusions pertaining to this material. This is especially in relation to the names of the people involved in performing or making this work and/or correctly naming the community from which it originally derives. As a user you are being asked to also apply the correct attribution in any future use of this work.	Rights	Notice text from the landing page of the Notice. For clarity, this should be preceded by the name of the Notice and the URL for the Notice. Use the exact text from the landing page; do not adapt or change the text.	Local Contexts Traditional Knowledge (TK) Notice https://localcontexts.org/notice/tk-notice : The TK Notice is a visible notification that there are accompanying cultural rights and responsibilities that need further attention for any future sharing and use of this material. The TK Notice may indicate that TK Labels are in development and their implementation is being negotiated.
rightsURI	The URL of the project in the Local Contexts hub.	https://localcontextshub.org/projects/29fc8acd-d647-48db-a766-b2c1bd9264b9/	rightsURI	The URL of the project in the Local Contexts hub.	https://localcontextshub.org/projects/29fc8acd-d647-48db-a766-b2c1bd9264b9/
rightsIdentifier	The identifier of the specific Label being applied.	TK-A BC-P	rightsIdentifier	The identifier of the specific Notice being applied.	TK-Notice,BC-Notice, Attribution-Incomplete,Open-To-Collaborate
rightsIdentifierScheme	Local Contexts	Local Contexts	rightsIdentifierScheme	Local Contexts	Local Contexts
schemeURI	Local Contexts website URI	https://localcontexts.org			

```
<rightsList>
<rights rightsURI="https://localcontextshub.org/projects/29fc8acd-
d647-48db-a766-b2c1bd9264b9/" rightsIdentifier="TK-Notice"
rightsIdentifierScheme="Local Contexts"
schemeURI="https://localcontexts.org">
  Local Contexts TK Notice https://localcontexts.org/notice/tk-notice: The TK Notice is a visible notification that there are accompanying cultural rights and responsibilities that need further attention for any future sharing and use of this material. The TK Notice may indicate that TK Labels are in development and their implementation is being negotiated.
</rights>
</rightsList>
```

```
<rightsList>
<rights rightsURI="https://localcontextshub.org/projects/29fc8acd-
d647-48db-a766-b2c1bd9264b9" rightsIdentifier="TK-A"
rightsIdentifierScheme="Local Contexts"
schemeURI="https://localcontexts.org">
  Local Contexts TK Attribution: This Label is being used to correct historical mistakes or exclusions pertaining to this material. This is especially in relation to the names of the people involved in performing or making this work and/or correctly naming the community from which it originally derives. As a user you are being asked to also apply the correct attribution in any future use of this work.
</rights>
</rightsList>
```

Figure 4: Rights field Implementation: Examples of implementation of Rights fields for Notice (Left) and Label (Right).

Section 4: Researcher Workflow within the Local Contexts Hub

Registering on the Local Contexts Hub can be done by going to the official [Local Contexts Hub's Website](#). To register, first make a profile by choosing a username and password, and confirming your email. All registrants must read and agree to both the Local Contexts Hub's Terms and Conditions and Indigenous Data Sovereignty Agreement.

After this, registrants create or join an account. There are three account types: Researcher, institution, and community. A researcher account is for an individual researcher who is either affiliated with an institution or working independently. Researcher (and institution) accounts enable the development of Projects with [Notices](#) on certain biodiversity data.

Box: Which Account type is Right For Me?

Accounts in the Local Contexts Hub are workspaces created for different entities. An account enables the development, collaboration and management of Local Contexts Labels, Notices and Projects. Any user can be a member of multiple accounts and multiple account types.

A "Researcher Account" is for an individual while an "Institution Account" is for a collective institution such as universities, research institutions, museums and libraries. Both researcher and institution accounts can only use Notices and do not have the ability to apply Labels. A "Community Account" is the only account type that can customize and apply Labels to biodiversity data.

Create a researcher account if the institution has yet to make a commitment to applying Notices. In the future we expect to be able to link Researcher Account to Institution Accounts.

After completing the researcher registration process, researchers can create a Project with Notices on biodiversity data to notify an IPLC. Local Contexts Notices allows researchers to acknowledge and make visible IPLC interests in collections, information and data. Notified IPLCs may then choose if, how and when to

apply their Labels. Notices and Labels provide a point of connection for people and place as they engage with biodiversity samples/data. While the primary function is to make transparent Indigenous interests in samples/data, they can also guide appropriate use/re-use.

There are four Notices that can be used by researchers (see Figure 3). There is one Engagement Notice, the Open to Collaborate Notice. There are three Disclosure Notices: The Attribution Incomplete Notice, the Traditional Knowledge (TK) Notice and the Biocultural (BC) Notice. Each Notice has a different purpose and can be used individually or together. The BC Notice is the most commonly used for biodiversity data. Details on the purpose of each Notice can be found on the Local Contexts website. The Local Contexts team is available to help with registration and Notice use queries. They can be contacted at support@localcontexts.org.

Table 1: Researcher and Institution Notice and Label Implementation Workflows

Activity	Internal Hub Actions	External Hub Actions
Registration, Profile, and Account Creation Workflow		
Create a profile	Sign up to Local Context (LC) Hub, add personal details to establish a profile in the Hub. Choose profile settings and personalize them as needed.	Support from LC staff (incl formal onboarding and informal engagement).
Create or join an institution account	View all account options, and decide whether to establish or join an existing institution account (can choose multiple).	Support from LC staff (incl formal onboarding and informal engagement). Creating an institution account requires a letter of support or contact person to validate. Joining an institution account may require contacting the administrator responsible for the account.
Create a researcher account	Establish an account on the LC Hub for an individual researcher.	Letter of support not needed, done autonomously. You are encouraged to connect an ORCID ID.
Project Creation and Disclosure Workflow		
Create a Project	Select project type and enter details as well as contacts, identifiers, and privacy levels. A Unique Project ID* is generated upon completion.	Can be done by an individual, although support is available from LC staff.
Apply Notice(s) to a Project	Notices are applied to a project at creation. Can be referenced via the API using the Project ID to copy Notices into the metadata.	Ensure Notices are applied appropriately for the context.
Notify community (or communities) of a Project	Can notify community users, or have invites sent to known IPLC contacts. LC Hub has a searchable community register.	Identifying a contact person or organization for the relevant IPLC may be required before notifying an IPLC through in the Hub.

Notice to Label Workflow		
Provide receipt of Notice(s) and Project	After a community has been notified of a project, the community may provide a status update or send a message through the Hub interface.	Community will receive an email notification that a project has been shared with them.
Apply a Label to replace Notice(s)	Community/ies can choose to add Labels or take no action. If Labels are added, then they automatically appear on the Project page, replacing the Notices.	Communities will decide if to apply Labels and determine which one(s) should be applied if so. The Indigenous Community Working Group may be a resource for communities to identify which Label(s).
Front facing finalization	API calls Project ID and provides relevant Notices and/or Labels. Relevant files can be downloaded to display the correct Labels.	Institutions must determine which field the Notice(s) and/or Label(s) will go in.
Visualizing Notices and Labels within Systems		
Visualizing Notices and Labels on digital record	Access Notices and Labels from the API	Discussions with relevant database administrators to ensure Notices and Labels can be visualized.

*When you create a Project in the Local Contexts Hub, it will automatically be assigned a Project ID, a randomly generated 36 alphanumeric UUID in the following format: 8-4-4-4-12, where each of the numbers represent the number of randomized alphanumerics (hyphens are constant) with all letters always being lowercase. The Project ID is found on the "View Project" page under the "Project Identifiers" section.

Section 5: What is the Right Scope or Scale For My Project?

When developing a Local Contexts Project, a judgment must be made relating to the scope and scale of the project. Projects could be generated for each sample but that potentially creates an administrative burden for both researcher and the community. Projects can be created for whole or partial collections where a single Project ID can be used across multiple samples. If there is an established relationship with a community, then the project might be tailored to incorporate a series of samples or projects within their geographic areas of interest. This will often be a sensible starting point for researchers.

A Project can cover a wide range of topics and that can include many different types of IPLC biodiversity data. These could include but are not limited to:

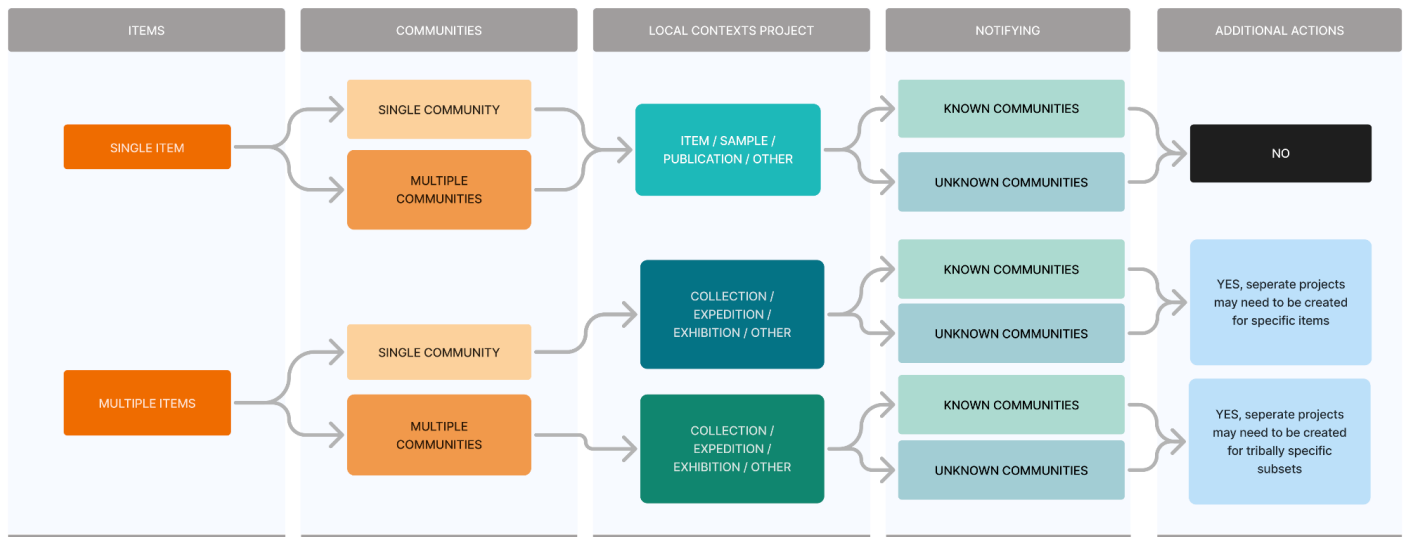
- Item (a single record, photograph (digitized vouchers), sound recording, manuscript, archival document, digital heritage item, artwork);
- Collection (papers, language materials, DNA sequences, voucher specimens, researcher collection);

- Physical Samples (samples used for taxonomic identification, genomic/transcriptomic/proteomic data generation, vouchering, cryopreserved samples;)
- Expedition (collecting event, research trip);
- Publication (digital publication, website, research publication, archival publication);
- Exhibition (online or on-site); or
- Something else (a university syllabus, metadata).

For issuers of the Labels and Notices it is highly encouraged that customized Labels are created to be as broad as possible. This increases the probability that the Label can be used across many projects, and reduces the burden associated with creating a unique Label/Notice for each Project. This will enable the scale of the Projects to be unlimited.

Decision tree

for creating Local Contexts Hub Projects



ITEMS

- Projects for **single items** create specificity for application of Labels but may require more effort to establish.
- Projects for **multiple items** may require additional actions to create sub and/or linked Projects.

COMMUNITIES

- Projects requiring input from a **single community** create direct bilateral relationships that support the application of Labels.
- Projects requiring input from **multiple communities** need to manage overlapping interests.

LOCAL CONTEXTS PROJECT

- A **Local Contexts Project** can be an item, a sample, publication, collection, expedition, exhibition or manually defined using the 'Other' option.
- There is **no difference in the information required** to establish Projects for single or multiple items.

NOTIFYING

- **Known communities** may be on LC Hub or an invite to join can be generated.
- **Unknown communities** may be found via Native Lands module
- LC Notices will remain on the Hub until a relevant community joins.

ADDITIONAL ACTIONS

Separate Projects may need to be created to...

- Specify tribally focused subsets within larger collections.
- **Or** around specific items within a larger collection.

Figure 4: Scalability considerations when generating Projects

Section 6: Which Communities Should I Notify?

It can be challenging to understand the often-complex relationships that IPLCs have with their territories and each other. Interests can overlap or the distribution of a particular species might cross the territories of many different communities. It is better to surface than silence this complexity and allow IPLCs to determine how their respective interests should be recognized. Existing resources such as Native Lands (www.native-land.ca) are useful resources that can facilitate researchers identifying communities with associations to the territory or land where the samples were/are to be collected.

Section 7: What Else Should I Do After Applying a Notice?

Once a project has been created with Notices applied, a single community or multiple communities can be notified of the project. Community members will be able to view the project information and have the option to respond with a status:

- Acknowledge Notice: The community has seen the project, but have not made a decision about applying the Labels.
- Labels pending: Labels will be applied
- Labels not pending: Labels will not be applied

Once a Label is added by a community, any Notices placed by an institution or researcher will be replaced by the Labels (excluding the Open to Collaborate Notice).

An integral component of the Local Contexts approach is the ability and autonomy of IPLCs to create their own unique set of Labels. There are three parts of a Label: Label icon, Label title, and Label description. Across every use of a Label, the Label icon remains the same. Communities can change the Label title and Label description to create their version of each Label. Some communities choose to write the Label description or Label title in multiple languages. Through the Hub, communities can apply their unique Labels to projects they create, and to projects from multiple institutions and researchers.

Section 8: How should I implement a Notice/Label in My ERGA Metadata Manifest?

For ERGA, the intention for promoting the Notices/Labels is to promote the rights of IPLCs within the project and provide an inclusive metadata infrastructure that facilitates disclosed rights and interests to persistently travel alongside all IPLC samples collected, and preserved as well as its associated sequencing data. While not all institutional systems participating in the project (collections and biobanks) will be configured to accept and display this information it is important to identify key secondary data infrastructures where this metadata should be enabled.

The generation of a reference genome under the ERGA umbrella is usually conducted by a Genome Team, composed of individual biodiversity experts that lead each aspect of the project lifecycle. A key component of

each Genome Team is the “Sample Ambassador” who coordinates with the Genome Team to organize the samples, permits, barcoding, storage up to shipment to laboratory, storage of vouchers and is generally responsible for metadata collection and manifest completion. For more information on the roles within Genome Teams please see [ERGA Pilot Project Guidelines](#).

A species sample ambassador will generate a Project ID in the Local Context hub, and when submitting samples will enter this ID in the manifest field “ASSOCIATED_TRADITIONAL_KNOWLEDGE_OR_BIOCULTURAL_PROJECT_ID”. Once samples are submitted and validated, COPO will link these two entities together and will pull in the Notice and Label information once available.

8.1. Guidelines for ERGA Managers and Sample Ambassadors

ERGA managers are selected metadata experts with administrative privileges that bestow upon them the ability to accept or reject metadata manifests through ERGAs metadata brokering service, [COPO](#). Here, managers determine whether the manifest is appropriately completed by sample ambassadors prior to BioSample generation. The EMBL-EBI BioSamples database ([Courtot et al. 2022](#)) stores and supplies centralized descriptions and unique metadata about biological samples across the INSDC (including ENA) and facilitates linkages across repositories e.g., the BioSample ID for the European Mink Pilot Project is [SAMEA12922170](#). If the manifest is not of sufficient quality, the manager will reject the manifest and alert the sample ambassador accordingly. Alternatively, if a manifest is sufficiently completed and accepted by the manager all metadata is associated with a publicly available BioSample ID. ERGA managers ensure that metadata manifests uploaded through [COPO](#) containing a Notice are NOT used for BioSample generation until feedback through the LC Hub’s API has been obtained concerning the status of the project (public/private).

After BioSample creation, raw sequencing data and reference genome assemblies are uploaded to the internal ERGA data repository by the sample ambassador. After this, sample ambassadors also upload an ERGA Data Manifest. The ERGA data manifest provides important information about how the sequencing data was generated. This manifest is utilized by COPO to associate the already generated BioSample ID to a newly formed BioProject ID to facilitate data entry into the ENA under the ERGA Pilot Project umbrella.

For manifests containing a Label or a Notice, sample ambassadors have additional obligations as partnering IPLCs may request an alternative mode of data sharing after upload into the ERGA data repository e.g., controlled access. If the IPLC has indicated that open data sharing is to be accommodated, sample ambassadors should follow the typical procedure and complete an ERGA Data Manifest. On upload of this manifest, it indicates to COPO to share this data to the ENA. If the data is to be entered into the public archive, COPO will automatically transfer the Label or Notice ID as a “SAMPLE ATTRIBUTE” to the ENA (note: this is expected to change in future updates, where specific fields will be included in the ENA checklist for the Label and Notice IDs which will aid findability). If an alternative mode of data sharing is requested by the IPLC, the sample ambassador should not upload a Data Manifest, but ensure to alert the ERGA data repository administrative staff to discuss and ensure the appropriate permissions are in place. Importantly, species sample ambassadors are expected to inform the partnering IPLC about how to access all metadata and



genomics data, and to encourage the IPLC to download all of the data as a backup to safeguard permanent access. From the outset, it is important for ERGA sample ambassadors to be clear and transparent when partnering with IPLCs ensuring that communities understand that if it is agreed that data is entered into the ENA, it can be challenging to remove, and it may have already been downloaded and accessed by other users of the repository.