**GSCID/BRC Project and Sample Application Standard**

**Core Project**

**v1.3**

**Finalized by the GSCID/BRC Metadata Working Group**

How to interpret the document:

**BOLD**: Field name

*ITALICS*: Attributes of the field

1. **Project Title**

*Core Project Field ID*: CP1

*Description*: Descriptive title of the sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: Sequencing of isolates from 2023 outbreak of pseudopneumovirus in Antarctica

*Data Source*: Sample Provider

*Comments*:

*OBO Foundry Synonym*: investigation title

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001622

*OBO Definition*: A textual entity that denotes an investigation.

*Other Synonym*: BioProject:Title\*;MIxS: project name;

1. **Project ID**

*Core Project Field ID*: CP2

*Description*: Unique identifier of the sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: RRI-112233445

*Data Source*: GSCID

*Comments*:

*OBO Foundry Synonym*: investigation identifier

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001628

*OBO Definition*: A CRID symbol used to identify an investigation.

*Other Synonym*:

1. **Project Description**

*Core Project Field ID*: CP3

*Description*: Textual description of the project including hypothesis, rationale, goals, etc.

*Data Categories*: Investigation

*Allowed Values*: textual abstract or hyperlink to project description document; for GSCID projects this could simply be the white paper proposal

*Syntax*: free text

*Example Values*: Nasal swabs were collected from 45 patients experiencing respiratory distress during emergency room visit from May - August of 2023 in Antarctica to determine the genome sequences of this new outbreak virus for comparison with previously circulating strains

*Data Source*: Sample Provider

*Comments*:

*OBO Foundry Synonym*: investigation description

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001615

*OBO Definition*: A textual entity that describes an investigation.

*Other Synonyms*: BioProject:Description\*;

1. **Project Relevance**

*Core Project Field ID*: CP4

*Description*: Indicates how the knowledge derived from the project can be applied, and to what field(s).

*Data Categories*: Investigation

*Allowed Values*: Agricultural, Medical, Industrial, Evolution, Environmental, Model organism, Other

*Syntax*: pick list

*Example Values*: Medical

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: study design

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0500000

*OBO Definition*: A study design is a plan specification comprised of protocols (which may specify how and what kinds of data will be gathered) that are executed as part of an investigation and is realized during a study design execution.

*Other Synonyms*: BioProject:Relevance\*;

1. **Sample Scope**

*Core Project Field ID*: CP5

*Description*: Indicates the scope and purity of the biological sample used for the project.

*Data Categories*: Investigation

*Allowed Values*: Monoisolate, Multiisolate, Multispecies, Environment, Synthetic, Single cell, Other

*Syntax*: pick list

*Example Values*: Monoisolate

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: specimen-based scope of investigation specification

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001884

*OBO Definition*: specimen-based scope of investigation specification

*Other Synonyms*: BioProject:Sample Scope\*;

1. **Target Material**

*Core Project Field ID*: CP6

*Description*: Indicates the type of material that is isolated from the sample for use in the project.

*Data Categories*: Investigation

*Allowed Values*: Genome, Transcriptome, Proteome, Purified chromosome, Reagent, Phenotype, Other

*Syntax*: pick list

*Example Values*: Genome

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: target material in specimen specification

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001882

*OBO Definition*: target material in specimen specification

*Other Synonyms*: BioProject:Material\*;

1. **Target Capture**

*Core Project Field ID*: CP7

*Description*: Indicates the scale, or type, of information that the project is designed to generate from the sample material.

*Data Categories*: Investigation

*Allowed Values*: Whole, Targeted locus/loci, Clone ends, Exome, Random survey, Other

*Syntax*: pick list

*Example Values*: Whole

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: target capture specification

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001899

*OBO Definition*: target capture specification

*Other Synonyms*: BioProject:Capture\*;

1. **Project Method**

*Core Project Field ID*: CP8

*Description*: Indicates the general approach used to obtain data.

*Data Categories*: Investigation

*Allowed Values*: Sequencing, Array, Mass Spectrometry, Other

*Syntax*: pick list

*Example Values*: Sequencing

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: investigation assay specification

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001896

*OBO Definition*: investigation assay specification

*Other Synonyms*: BioProject:Methodology\*;

1. **Project Objectives**

*Core Project Field ID*: CP9

*Description*: Indicates the project goals with respect to the type of data that will be generated and submitted to an INSDC database.

*Data Categories*: Investigation

*Allowed Values*: Raw sequence reads, Sequence, Analysis, Assembly, Annotation, Variation, Epigenetic markers, Expression, Maps, Phenotype, Other

*Syntax*: pick list

*Example Values*: Assembly; Annotation

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: specification of data to be generated in an investigation

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001892

*OBO Definition:* specification of data to be generated in an investigation

*Other Synonyms*: BioProject:Objective\*;

1. **Grant Agency**

*Core Project Field ID*: CP10

*Description*: The name of the agency providing funding support for the specimen collection and sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: U.S. National Institutes of Health; Bill and Melinda Gates Foundation

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: grant agency

*OBO Foundry Purl*: submitted to OBI

*OBO Definition*: An organization that provides funding support for projects such as investigations.

*Other Synonyms*:

1. **Supporting Grants/Contract ID**

*Core Project Field ID*: CP11

*Description*: Unique identifier of the grant(s) and/or contract(s) that supports the specimen collection and sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text; not applicable allowed

*Syntax*: free text

*Example Values*: N01AI40041; HHSN266200400041C

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: grant ID

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001629

*OBO Definition*: A CRID symbol used to identify a grant.

*Other Synonyms*: BioProject:Grant ID;

1. **Publication Citation**

*Core Project Field ID*: CP12

*Description*: Citation for scientific publication(s) related to the specimens included in the project and/or the sequence derived from the project and/or its interpretation

*Data Categories*: Investigation

*Allowed Values*: PubMed ID; Digital Object Identifier (DOI); not applicable allowed

*Syntax*: free text

*Example Values*: PMID:22260278; DOI:10.1126/science.323.5915.713a

*Data Source*: Sample Provider

*Comments*: Separate multiple entries with a semicolon (;).

*OBO Foundry Synonym*: PubMed ID

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001617

*OBO Definition*: A CRID symbol that is sufficient to look up a citation from the PubMed, a literature database of life sciences and biomedical information.

*Other Synonyms*: BioProject:PubMed ID, DOI;MixS: ref\_biomaterial;

1. **Sample Provider Principal Investigator (PI) Name**

*Core Project Field ID*: CP13

*Description*: Name of the responsible investigator providing the sample set and metadata for the sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: Jane Z. Doe

*Data Source*: Sample Provider

*Comments*: BioProject:will be included as custom attribute and concatenated together with "Sample Provider PI's Institution" and "Sample Provider PI's email" fields.

*OBO Foundry Synonym*: specimen provider principal investigator

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001889

*OBO Definition*: A person who is a principal investigator and provides the specimen

*Other Synonyms*:

1. **Sample Provider PI's Institution**

*Core Project Field ID*: CP14

*Description*: Institutional affiliation of the responsible investigator providing the sample set and metadata for the sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: Random Research Institute (RRI)

*Data Source*: Sample Provider

*Comments*: BioProject: will be included as custom attribute and concatenated together with "Sample Provider PI's Institution" and "Sample Provider PI's email" fields.

*OBO Foundry Synonym*: organization of specimen provider principal investigator

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001880

*OBO Definition*: An organization that is the affiliation of the principal investigator providing the specimens for the investigation

*Other Synonyms*:

1. **Sample Provider PI's email**

*Core Project Field ID*: CP15

*Description*: Preferred email address of the responsible investigator providing the sample set and metadata for the sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: jane\_doe@aol.com

*Data Source*: Sample Provider

*Comments*: BioProject: will be included as custom attribute and concatenated together with "Sample Provider Principal Investigator (PI) Name" and "Sample Provider PI's Institution" fields.

*OBO Foundry Synonym*: email address of specimen provider principal investigator

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001903

*OBO Definition*: An email address of the principal investigator providing the specimens for the investigation

*Other Synonyms*:

1. **Sequencing Facility**

*Core Project Field ID*: CP16

*Description*: Name of the facility resposible for sequence determination; in many cases this will be one of the DMID-supported Genome Sequence Centers for Infectious Diseases (GSCIDs)

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: Awesome Sequencing Center (ASC)

*Data Source*: GSCID

*Comments*:

*OBO Foundry Synonym*: sequencing facility organization

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001891

*OBO Definition*: An organization that provides sequence determination service

*Other Synonyms*:

1. **Sequencing Facility Contact Name**

*Core Project Field ID*: CP17

*Description*: Name of the responsible investigator leading the sequencing project at the sequencing facility

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: John Q. Smith

*Data Source*: GSCID

*Comments*:

*OBO Foundry Synonym*: sequencing facility contact person

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001888

*OBO Definition*: A person who is the contact representative at the sequencing facility

*Other Synonyms*:

1. **Sequencing Facility Contact's Institution**

*Core Project Field ID*: CP18

*Description*: Institutional affiliation of the responsible investigator leading the sequencing project at the sequencing facility

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: University of Excellence (UOE)

*Data Source*: GSCID

*Comments*:

*OBO Foundry Synonym*: organization of sequencing facility contact person

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001897

*OBO Definition*: An organization that is the affiliation of the contact representative at the sequencing facility

*Other Synonyms*:

1. **Sequencing Facility Contact's email**

*Core Project Field ID*: CP19

*Description*: Preferred email address of the responsible investigator leading the sequencing project at the sequencing facility

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: john\_smith@gmail.com

*Data Source*: GSCID

*Comments*:

*OBO Foundry Synonym*: email address of sequencing facility contact person

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001894

*OBO Definition*: An email address of the contact representative at the sequencing facility

*Other Synonyms*:

1. **Bioinformatics Resource Center**

*Core Project Field ID*: CP20

*Description*: Name of the DMID-sponsored Bioinformatics Resource Center (BRC) that contains the metadata linked to the organism sequence

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: New Pathogen Resource (NewPR)

*Data Source*: BRC

*Comments*:

*OBO Foundry Synonym*: Bioinformatics Resource Center

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001626

*OBO Definition*: An organization that is one of the Internet-based research centers established and funded by NIAID (the National Institute of Allergy and Infectious Diseases). The Bioinformatics Resource Centers (BRCs) were formed in response to the threats posed by emerging and re-emerging pathogens, particularly CDC Category A, B, and C pathogens, and their potential use in bioterrorism. The intention of NIAID in funding these bioinformatics centers is to assist researchers involved in the experimental characterization of such pathogens and the formation of drugs, vaccines, or diagnostic tools to combat them.

*Other Synonyms*:

1. **Bioinformatics Resource Center Contact Name**

*Core Project Field ID*: CP21

*Description*: Name of the investigator at the BRC responsible for managing the data derived from the sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: Joe Bioinfo

*Data Source*: BRC

*Comments*:

*OBO Foundry Synonym*: Bioinformatics Resource Center contact person

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001883

*OBO Definition*: A person who is the contact representative of a Bioinformatics Resource Center

*Other Synonyms*:

1. **Bioinformatics Resource Center Contact's Institution**

*Core Project Field ID*: CP22

*Description*: Institutional affiliation of the investigator at the BRC responsible for managing the data derived from the sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: Advanced Bioinformatics Consortium (ABC)

*Data Source*: BRC

*Comments*:

*OBO Foundry Synonym*: organization of Bioinformatics Resource Center contact person

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001881

*OBO Definition*: An organization that is the affiliation of the person who is contact representative of a Bioinformatics Resource Center

*Other Synonyms*:

1. **Bioinformatics Resource Center Contact's email**

*Core Project Field ID*: CP23

*Description*: Preferred email address of the investigator at the BRC responsible for managing the data derived from the sequencing project

*Data Categories*: Investigation

*Allowed Values*: free text

*Syntax*: free text

*Example Values*: joe\_bioinfo@abc.org

*Data Source*: BRC

*Comments*:

*OBO Foundry Synonym*: email address of Bioinformatics Resource Center contact person

*OBO Foundry Purl*: http://purl.obolibrary.org/obo/OBI\_0001887

*OBO Definition*: An email address of the person who is contact representative of a Bioinformatics Resource Center

*Other Synonyms*: