

Inventory and Monitoring Division

National Park Service

U.S. Department of the Interior

Natural Resource Stewardship and Science



Minimum Implementation Standards for Network Projects

Version: 3.0

Date Released: 10/13/17

Last Updated: 13 October 2017

Table of Contents

1 Introduction..... 3

2 Database Standards..... 3

3 Water Quality..... 4

4 Protected Records..... 5

5 Data Processing Level/Certification..... 12

6 Taxonomic Concept – Identification 17

7 References 20

Revision History Log:

Previous Version #	Revision Date	Author(s)	Changes Made	Reason for Change	New Version #
NA	8/17/2016	Frakes, Kingston	First version	NA	1.0
1.0		Frakes, Kingston	1. Clarified that minimum standards apply to new databases as well as existing 2. Added protected status guidance	1. Central office does not currently have the resources to implement all standards at present 2. Integration of QAP guidance	1.1
1.1		Frakes, Kingston	Added guidance for: • Data Processing Level • Change Tracking • Flagging • Taxonomic Concept • Protocol Configurations	Not have minimum standards for each is taking too long to reconcile individually with each network	1.2
1.2	6/14/2017	Kingston	Updated Protected Records to include example schema diagrams. Updated draft Certification.	Clarification on Protected Records Completion on Certification	1.3
1.3	6/22/2017	Kingston	Updated Taxonomic Concept to include examples in same format as those in Protected Records. Modified column names for Taxonomic Concept.	Consistency in diagrams and consistency with lookup table conventions	1.4
1.4	7/24/2017	Kingston	Accepted changes for Protected Records and Data Processing Level	Publish version 2.0	2.0
2.0	9/18/2017	Kingston	Moved Taxonomic Concept – Identification from section 8 to section 6 and changed from draft to published Updated db diagrams to show crow's feet instead of arrows. DPL: Clarified that it is appropriate to delete unused values where lookup values are specified for DPL	Publish version 3.0 to include Taxonomic Concept – Identification Minimize confusion Clarification	3.0

1 Introduction

This document describes the minimum policy implementation level for any network assistance project implemented by the central office. Minimum implementation does not imply never going beyond the minimum; however, in cases where time/resources are a constraint, these standards will ensure that projects are manageable.

These standards apply to all network projects, including those developing a new database application as well as working with existing systems.

The target audience for these standards is primarily the central office staff although it may also be beneficial and informative for network data managers.

Although we hope to always attain the minimum implementation level, we anticipate cases where even doing the “minimum” is unattainable/unreasonable given existing resources. Thus, exceptions will be accommodated on a case-by-case basis with the expectation that they be uncommon, and negotiated and documented in the project plan.

These minimum implementation standards will be updated regularly as policy is introduced/revised and as the central office determines reasonable levels of minimal implementation.

2 Database Standards

See <https://irma.nps.gov/DataStore/Reference/Profile/2224205?Inv=true> (Frakes et al, 2015) for explanation of standards

For any project involving an existing database that needs revision/improvement, the central office will ensure that the database meets the following standards:

General Database

- Referential Integrity (1.2)
- Data Stand Alone (1.4)

Table Organization and Naming

- Table Description (2.2)
- Primary Keys (2.4)
- Unique Records (2.8)
- Distinct Lookup Tables (2.9)

Column Organization and Naming

- Column Names Representing Unit of Measure (3.6)
- Column Descriptions (3.7)
- Geospatial Units and Datum (3.9)
- Default Values (3.12)
- Zero-Length Strings (3.13)

3 Water Quality

For all network assistance projects involving discrete water quality data, the central office will ensure that the product has the capacity to provide the data to the Water Resource Division (WRD) instance of EQuIS.

Databases that are developed/augmented by the central office will have an EQuIS EDD export routine for submission to WRD. Specific guidance on the mapping is found in the "Water Resource Division EQuIS EDD Mapping: IMD Implementation" [document](#). At a minimum, the central office will:

- Map data to all required fields defined by EQuIS
- Attempt to map all data collected by the network that is considered optional by the WRD.

In cases where there are water quality data that do not clearly have an appropriate location in the EQuIS EDD, the central office will not attempt to determine work-arounds or negotiate with the WRD about additions/enhancements to the existing EDD.

Finally, the central office will direct networks to request that labs submit data to Networks in WQX structure. If labs are unaware of EQuIS WQX, a template is [here](#).

4 Protected Records

The desired outcome of this standard is to explicitly communicate the protected status of all records to ensure that all unprotected records are shared with the public and that protected content is only made accessible to appropriate audiences.

Because data are intended to be open and public by default, Networks will identify potential types of data that must be classified as protected (See [QAP Guidance](#)) as early as possible in the data management workflow so as to limit the risk of accidental dissemination (ideally immediately upon data entry). Networks will work with appropriate park staff (typically resource management chiefs or above) to define what data are protected, the degree of protection warranted, and if/how data will be disseminated.

Types of information to evaluate include:

- Resource Data and Information – Information about biological taxa
- Location Data and Information – Location-specific information
- Personally Identifiable Information – Information that is used to uniquely define an individual

One of three levels of protected status are to be assigned to each of the main information types to generate a protected status matrix (See Example Table 4.1; also refer to DO 66 RM 66B for policy guidance):

- Legally Protected. Covered under existing statute as being exempt from public dissemination.
- Operationally Protected. Data withheld from release to protect resources or site fidelity but may be subject to release upon request.
- Not Protected. Data is neither legally nor operationally protected and therefore does not have any distribution restrictions conditional on protected status.

Example Table 4.1. Example table summarizing data protection standards for inclusion with quality assurance plans developed for protocols implemented by the NPS Inventory & Monitoring Division. With the exceptions noted, all data collected are to be made publicly available in a timely fashion.

Type of Data		Level of Protection	Rules for Dissemination
Resource Data and Information			
	T&E Species Locations	Legally Protected	Locations of observations reported at park centroid only.
	<i>Galax</i> species locations	Legally Protected	Commercially valuable; location information redacted from all dissemination
	T&E migratory shorebird locations	Legally Protected	Locations of observations reported at park centroid for three years; fuzzed to within 50m thereafter.
Location Data and Information			
	Cave entrances	Legally Protected	Redacted
	Spring locations	Legally Protected	Redacted; features associated with protected archaeological sites.
	Permanent Sampling Locations	Operationally Protected	Coordinates provided upon request only. Relative position shown in reports
Personally Identifiable Information			
	Non-NPS Staff Information	Legally Protected	All but first name and last initial redacted from public release

Managing Resource Data and Information

- The master taxonomy lookup table will contain the current protected status for each species tracked by the protocol. This status should be updated regularly. It may also contain a Protected Status Notes column (entry optional) that can be used to describe the reason for the protected status, when it was updated, etc.
- When a taxon is recorded in an observation table, it will inherit the current protected status from the master taxonomy lookup table, but retain the ability to change the protected status, based on defined user roles, between initial data entry and certification
- Reporting of a taxon (whether by export, report, query, etc.) should use the protected status defined in the taxon observation table (not the lookup table) along with the dissemination rules to determine the content of the output

Managing Location Data and Information

- The site/location table will have a required protected status column and an optional Protected Status Notes column
- Reporting of any location (whether by export, report, query, etc.) should use the protected status defined at the site/location table along with the dissemination rules to determine the content of the output

Managing Personally Identifiable Information (PII)

- The staff/contact table will have a required protected status column and an optional Protected Status Notes column
- Reporting of any staff/person (whether by export, report, query, etc.) should use the protected status defined at the staff/person table along with the dissemination rules to determine the content of the output
- Efforts should be made to avoid including PII in IMD databases unless absolutely necessary

Protected Status Lookup Table

The following ProtectedStatus lookup table will be used

Code	Label	Summary
LP	Legally Protected	Covered under existing statute as being exempt from public dissemination
OP	Operationally Protected	Data withheld from release to protect resources or site fidelity but may be subject to release upon request.
NP	Not Protected	Data is neither legally nor operationally sensitive and therefore has does not have any distribution restrictions conditional on protection status

4.1 Example Database Schema Elements – Resource Data and Information (Biological Taxa)

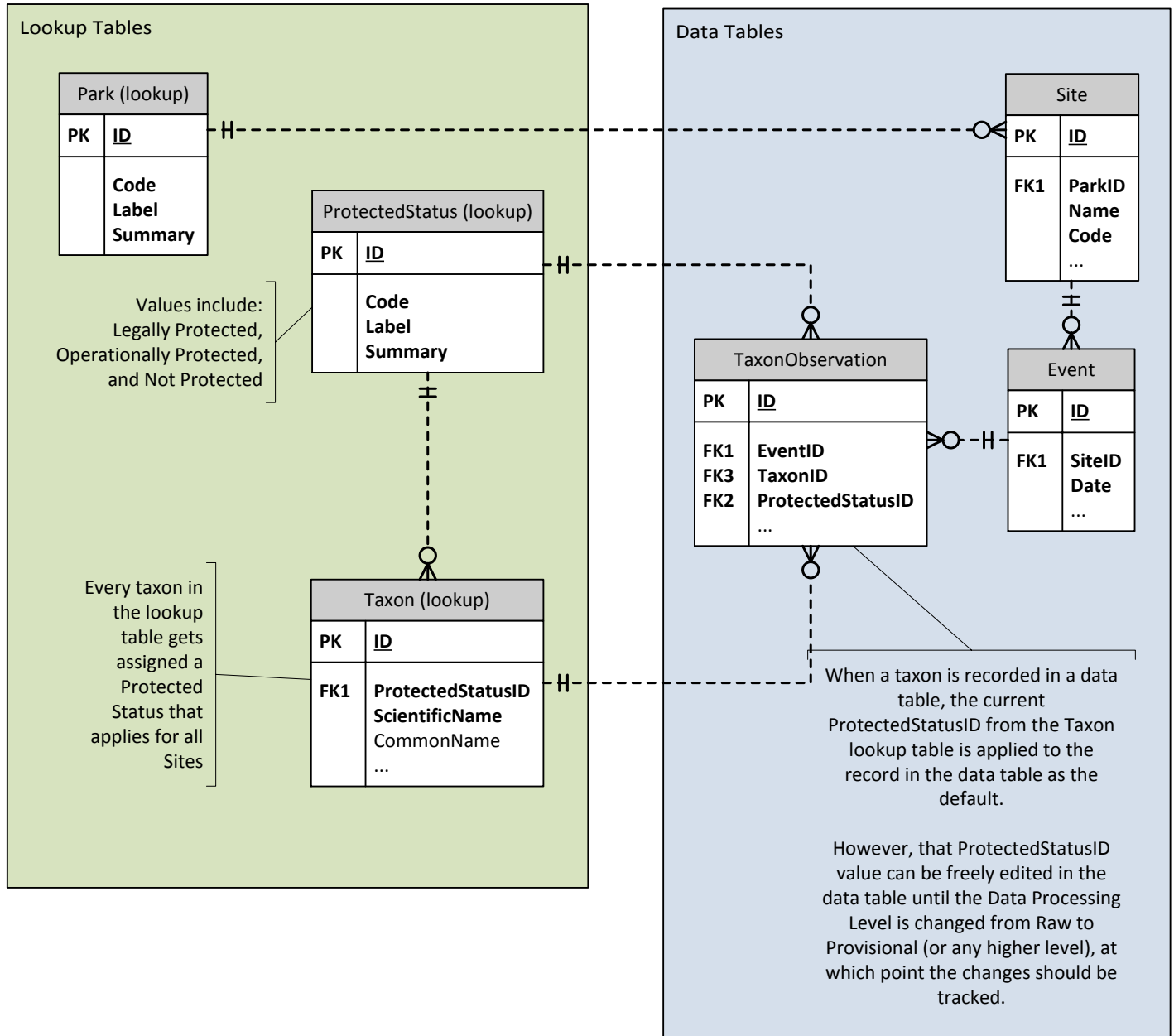
The simplest possible schema model should be selected for implementing Protected Status in the database. The models are as follows (listed in increasing order of complexity):

- Basic model – Protected Status is the same for a taxon for every site in the database
- Park model – Protected Status is the same for a taxon for each park in the database
- Site model – Protected Status for a taxon depends on the site (i.e. sampling/observation location)
- Single Taxon Protocol model – Protocol is only about one taxon (e.g., SWAN Bald Eagles)

Note that in all models it is assumed that the desired default Protected Status for a taxon may change over time. The decision on which model to use is based on whether or not a taxon has one status value for all sites, or not, at a point in time.

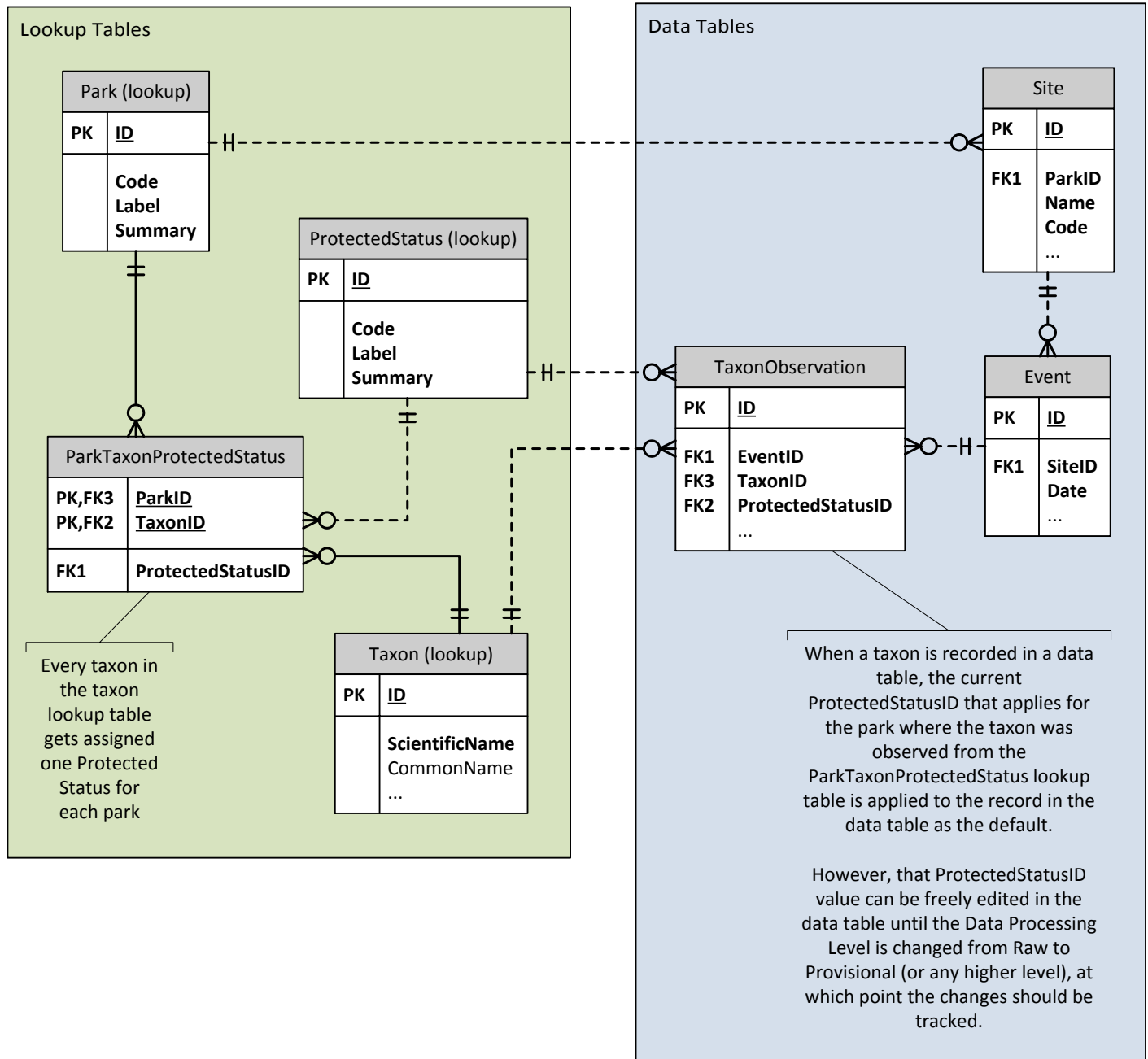
Protected Status Database Example for Resource Data and Information – Basic Model (simple)

Assumptions: This model assumes that each taxon has one Protected Status that applies for every site in the database. If the Protected Status is not the same for a taxon at every site, see either the Park Model or the Site Model.



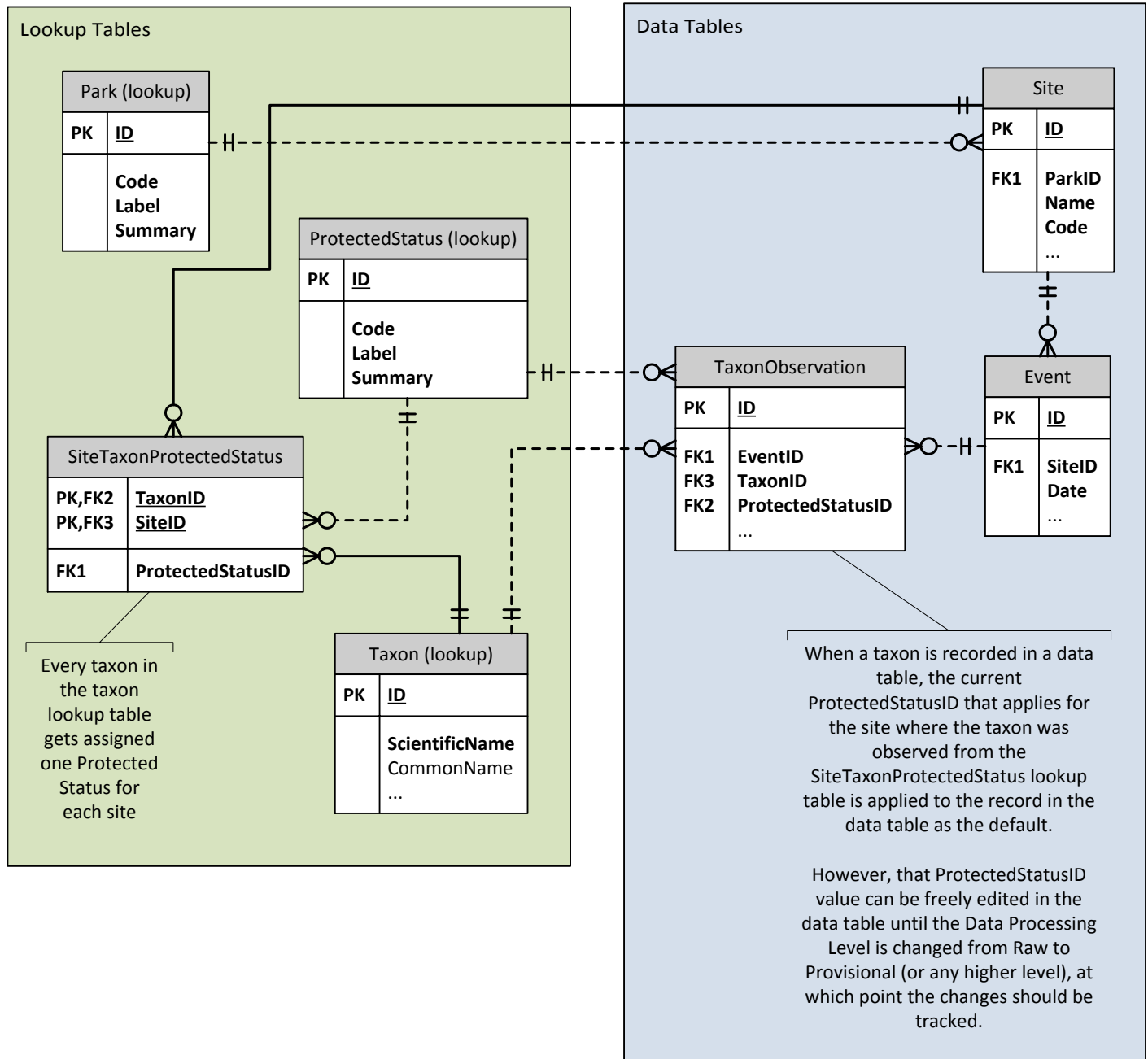
Protected Status Database Example for Resource Data and Information – Park Model (medium complexity)

Assumptions: This model assumes that each taxon has one Protected Status per park in the database. If the Protected Status for a taxon is not the same for all sites at a particular park, see the Site Model.



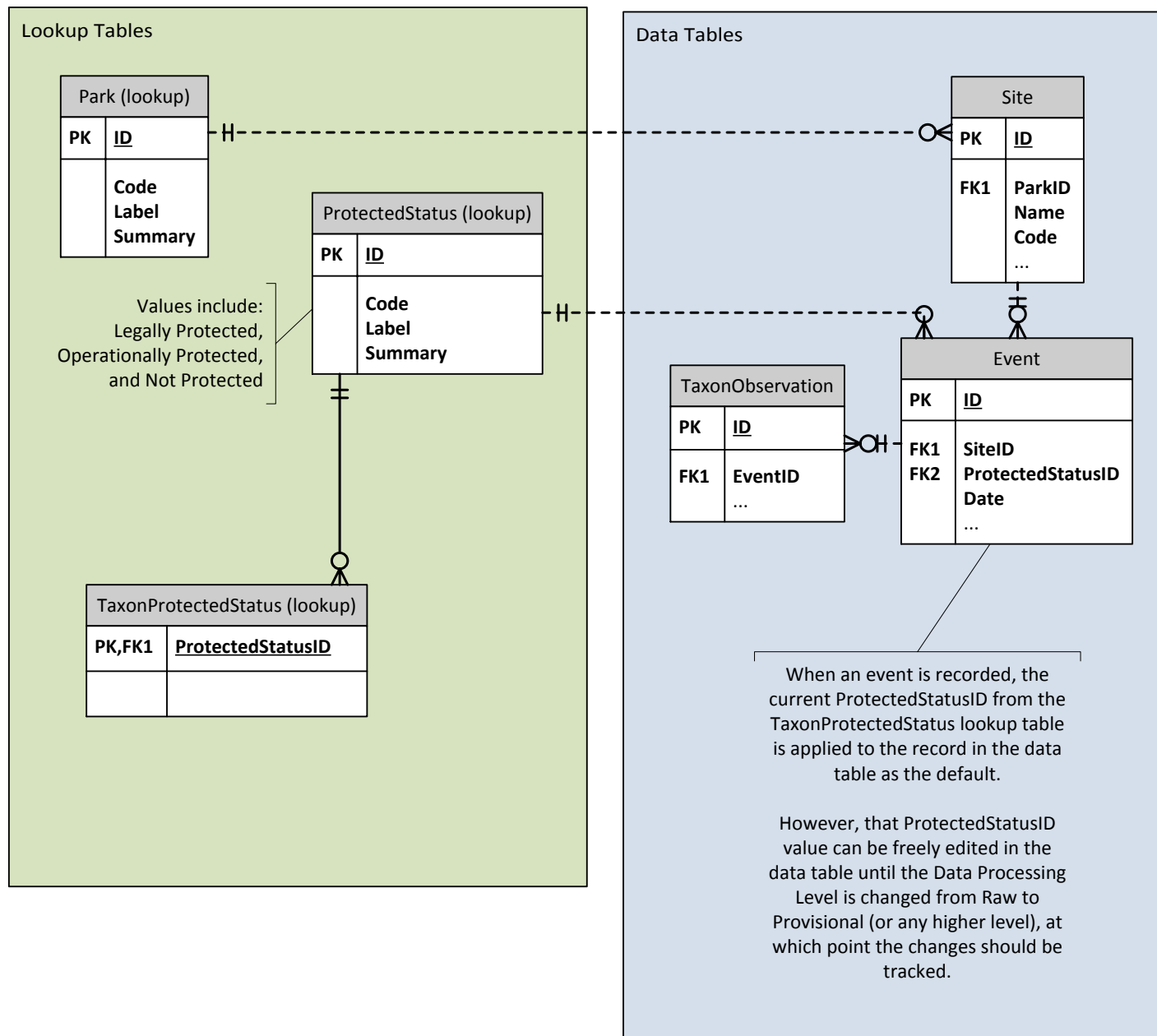
Protected Status Database Example for Resource Data and Information – Site Model (higher complexity)

Assumptions: This model allows for a Protected Status per site in the database. Only use this model if the other models will not suffice.



Protected Status Database Example for Resource Data and Information – Single Taxon Protocol Model (simple)

Assumptions: This model assumes that the protocol is only about a single taxon and that taxon has one Protected Status that applies for every site in the database. If the Protected Status is not the same for a taxon at every site, this model must be extended like the Park Model or the Site Model.

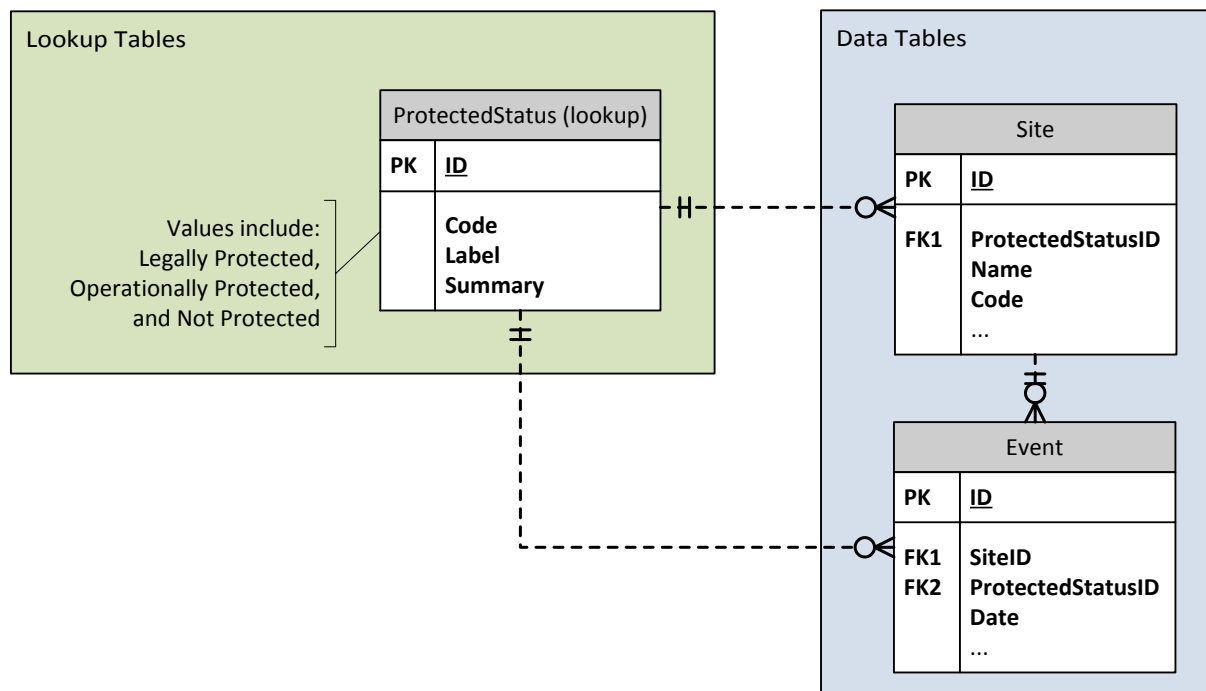


4.2 Example Database Schema Elements – Location Data and Information

ProtectedStatusID is simply added to the table that manages monitoring locations (in this example, Site).

Protected Status Database Example for Location Data and Information

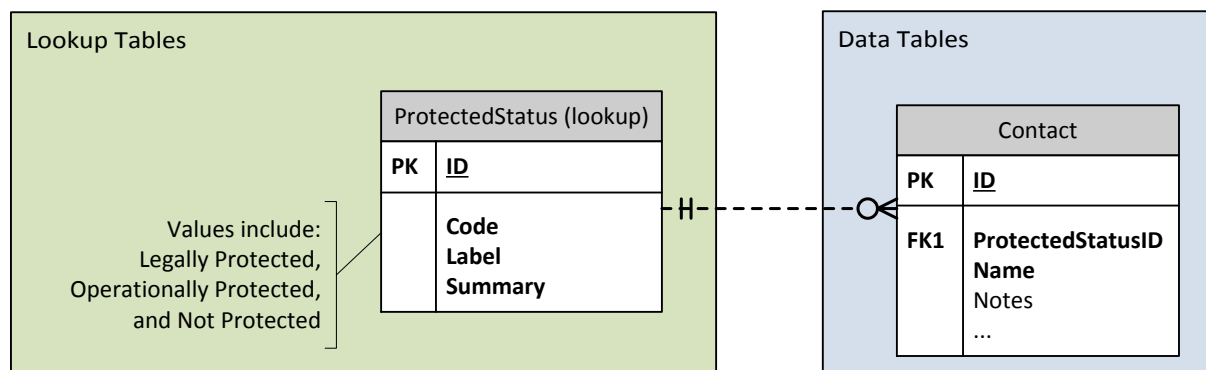
Assumptions: Current Protected Status for sites is maintained in Site table. When an Event is recorded at a site, the ProtectedStatusID from the related Site table is copied to the ProtectedStatusID in the Event table.



4.3 Example Database Schema Elements – Personally Identifiable Information (PII)

ProtectedStatusID is just added to the table that manages information about people (in this example, Contact) if such a table exists in the database.

Protected Status Database Example for Personally Identifiable Information (PII)

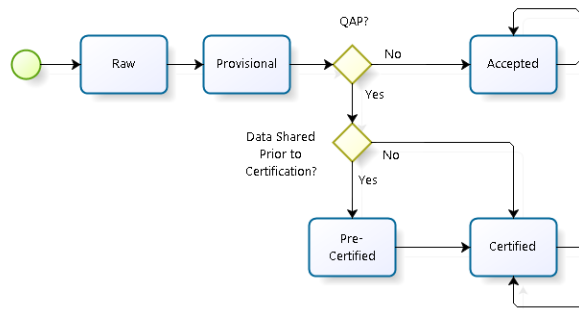


5 Data Processing Level/Certification

The desired outcome of this standard is to communicate the data processing level (DPL; heretofore known as Certification) for every observation¹ yet do it in an efficient way that minimizes database bloat that would occur if this were tracked separately for every column.

This standard will be applied to all new databases created and/or reviewed in relation to protocols pending approval in 2017. For existing approved protocols, this is entirely optional although encouraged.

DPL was addressed in the “[Certification Guidelines \(NPS, 2016\)](#)” and further refined in the “[Draft QAP guidance \(DeVivo, 2016\)](#)”. The documentation addresses five possible data processing levels, which include: Raw, Provisional, Accepted, Pre-Certified, and Certified. Allowable levels are dependent on whether or not a Quality Assurance Plan (QAP) is present.



The functional analyst (FA) should ask the following questions of the network client:

1. Do you have a QAP? (This determines the IsActive column values in our lookup table below)
 - a. If no, we will use the following DPL values: Raw, Provisional, Accepted
 - b. If yes, we will use the following DPL values: Raw, Provisional, Pre-Certified (optional), and Certified
2. What is the level of granularity at which you apply your QA or QC processes? For each activity, at what level do you need to certify? (This determines where we put the DataProcessingLevelID columns in the tables).
 - a. At the event level
 - b. At the activity level (e.g., a site condition assessment, a bird count)
 - c. At the individual measurement level (e.g., pH measurements for a water quality activity)

The FA **should not** be involved in determining when a data processing level should be changed or otherwise interpreting the draft [QAP guidance](#) (DeVivo, 2016). If the network has questions about this, they should defer to their protocol lead.

¹ An observation is any value collected/measured in the field

Certification/Data Processing Lookup Table

The following DataProcessingLevel lookup table structure will be used:

Name	Data Type	Required	Is Unique	Description
ID	tinyint	Yes	Yes	Unique primary key identifier
Code	Text, 2	Yes	Yes	Unique alpha code for each lookup value
Label	Text, 20	Yes	Yes	Unique name of each lookup value
Summary	Text, 255	Yes	Yes	Definition
IsActive	Boolean/bit	Yes	No	Indicates whether the data processing level is appropriate for the given application/protocol.
ProcessOrder	Tinyint	Yes	Yes	Used to enforce the forward progress of data processing levels and prevent moving backwards.

The following DataProcessingLevel lookup table values²³ will be used

Code	Label	Summary	IsActive	Process Order
R	Raw	Data in their state of original collection, either in the form of paper or electronic field forms, electronic sensor files, or digital media (photo, video, or audio). These data generally have been collected following some QA procedures, but have not undergone any QC procedures.	Yes	10
P	Provisional	Data that have undergone all planned QA procedures, have been entered into the system of record, and have passed verification procedures to ensure that data in the system of record match source data (field forms, sensor data files, etc.). Provisional data have not undergone all QC procedures.	Yes	20
A	Accepted	Data that have undergone all QA and QC procedures as defined in a published protocol and associated SOPs. Used in cases where a Quality Assurance Plan has not been developed and approved.	Yes if no QAP	30
PC	Pre-Certified	Data within a system of record that have undergone defined minimum levels of QC as defined in the QAP, rendering the data suitable for use for a limited set of designated or defined purposes only. The QC procedures that have been performed are communicated clearly to the end user. In general, this status should only be used in cases where data are routinely used to support park management decisions before certification can be completed.	Yes if QAP	40
C	Certified	Data have been through all QC procedures as outlined in the published protocol and the associated published QAP: Data either meet the documented quality standards, or data that do not meet the standards are appropriately identified at the observation level; procedures for data collection, processing, and transformation have been documented in appropriate metadata formats; data are suitable for use.	Yes if QAP	50

Rules and Logic

- DPLs will never be assigned to lookup tables
- DPL columns will be required with a default value of “raw”
- The DPL column used in each of the data tables should be named DataProcessingLevelID
- The last date of DPL assignment/status change is required. The Date will default to the current date. If a column is inserted into the table to accomplish this outcome, it should be named DataProcessingLevelDate.
- The individual who made the last DPL assignment/status change is required. If a column is inserted into the table to accomplish this outcome, it should be named DataProcessingLevel_UserID.
- Ideally, applications will have business logic that prevents values from violating the state change diagram above by moving backwards in state (e.g., from Provisional to Raw).
- DPLs may be re-“Certified” or re-“Accepted”, causing the DataProcessingLevelDate to change
- The application (i.e., user interface) will clearly communicate what is being certified when DPL values are applied

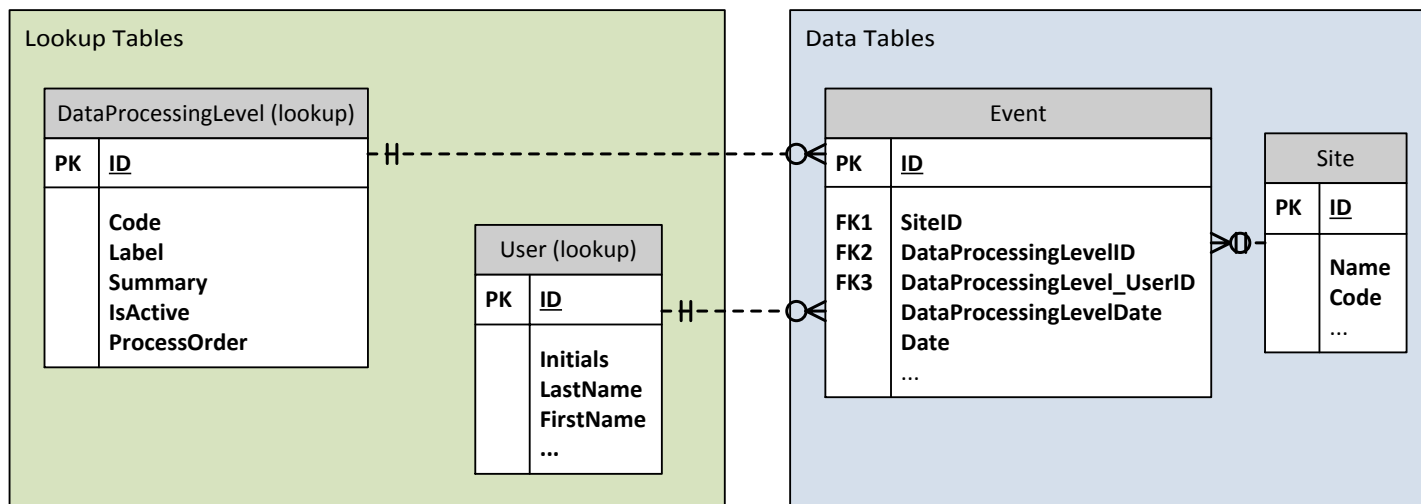
² values and definitions from draft QAP guidance, Phase IIIc section (DeVivo, 2016)

³ it is appropriate to delete values that will not be used

5.1 Example Database Schema Elements

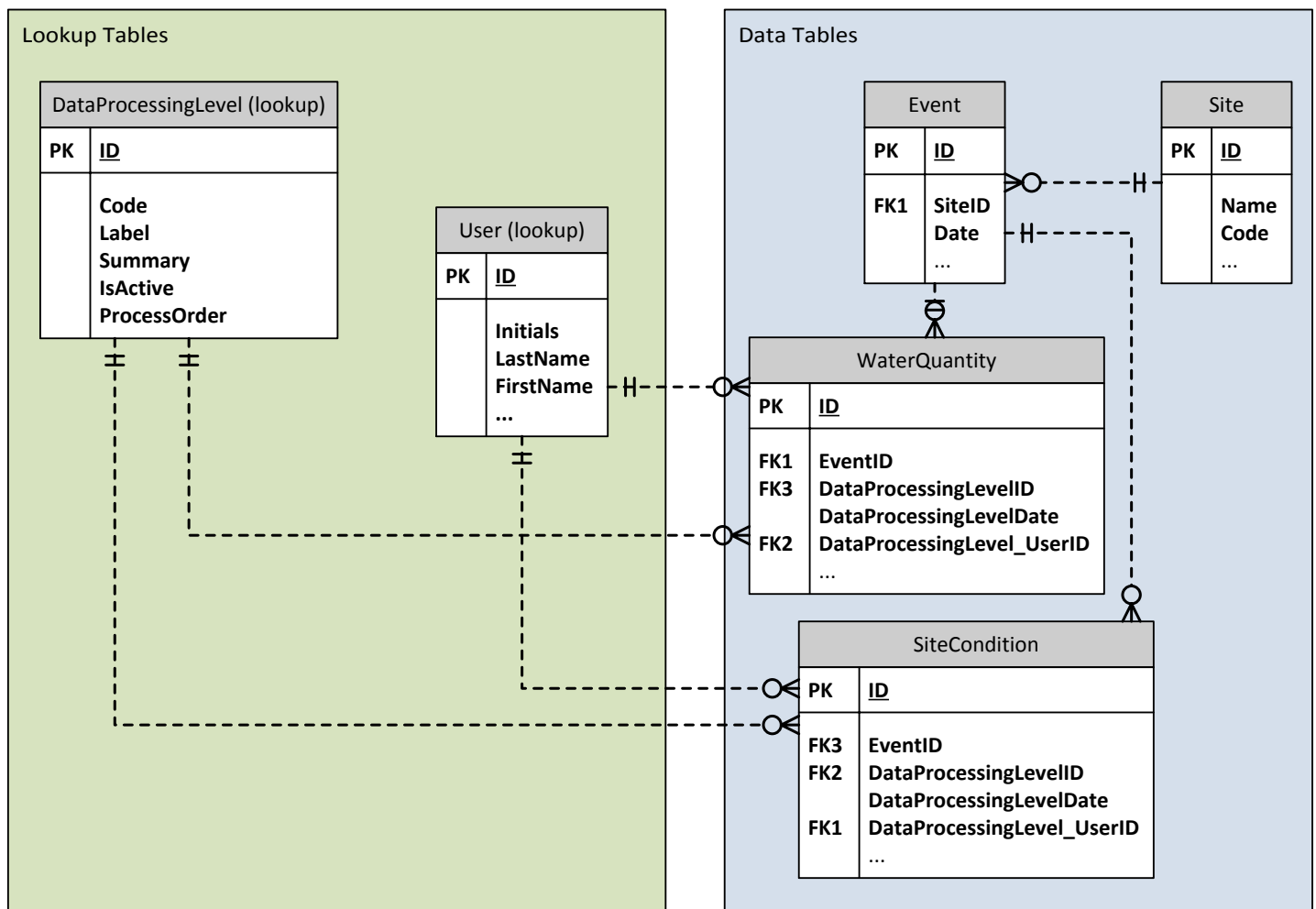
Certification/Data Processing Level Database Example – Event Level Model (simple)

Assumptions: This model assumes that Certification/Data Processing Level flags can be applied at the Event level and that these flags apply to all data for a particular Event. DataProcessingLevelID defaults to “raw”.



Certification/Data Processing Level Database Example – Activity Level Model (slightly more complex)

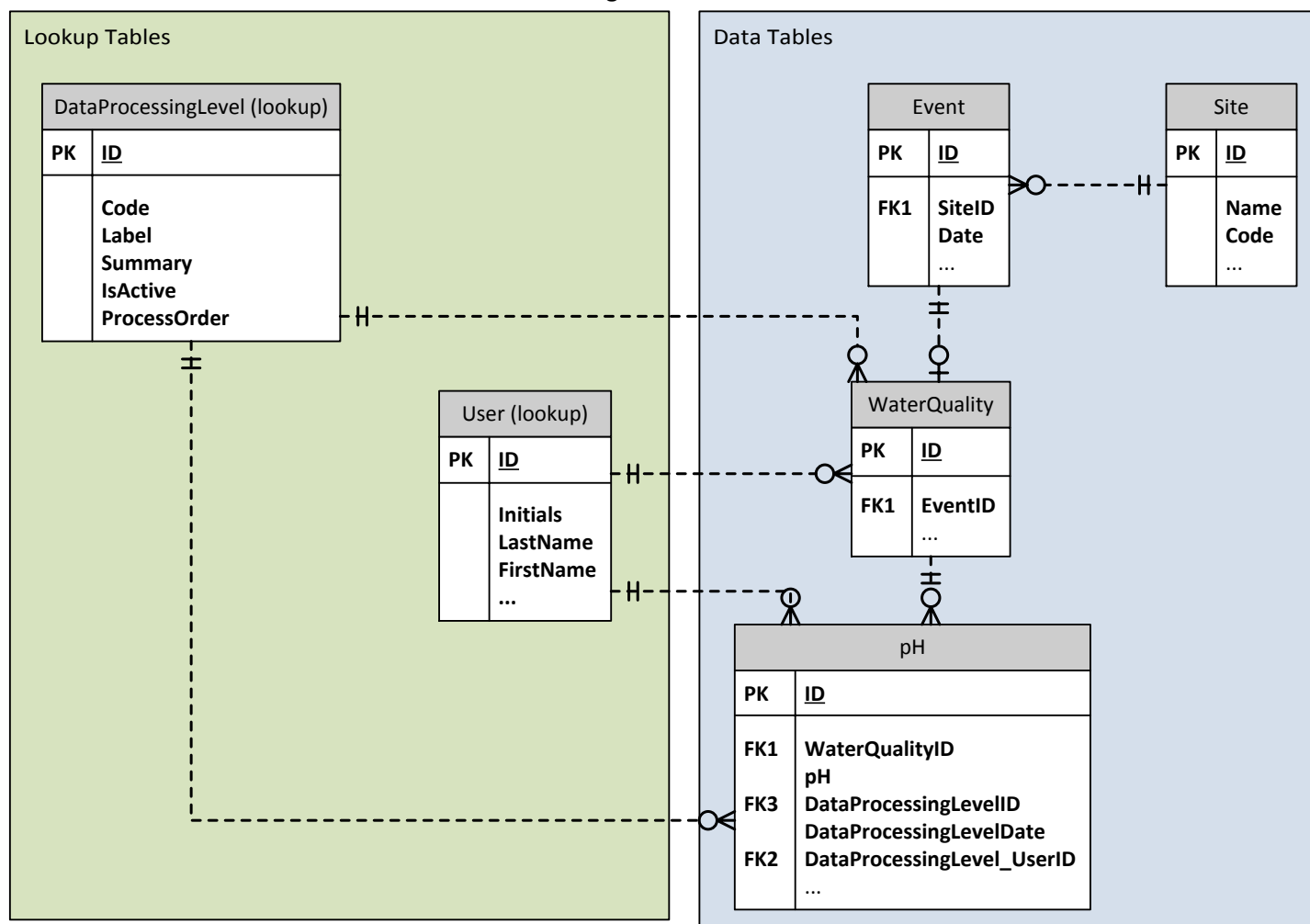
Assumptions: This model assumes that Certification/Data Processing Level flags can be applied at the Activity level and that these flags apply to all data for a particular Activity. DataProcessingLevelID defaults to “raw”.



Certification/Data Processing Level Database Example – Measurement Level Model (most complex)

Assumptions: This model assumes that Certification/Data Processing Level flags must be applied at the Measurement level and separates the Measurement out into its own table in order to do so.

DataProcessingLevelID defaults to “raw”.



6 Taxonomic Concept – Identification

The desired outcome of this standard is to ensure that:

- All taxa recorded refer to the standard used to identify the taxon (taxon + standard = taxonomic concept)

This standard only applies to protocols that record taxa. This standard will be applied to all new databases created and/or reviewed in relation to protocols pending approval in 2017. For existing approved protocols, Networks have the following two options: (1) Adopt new taxonomic concept standard or (2) defer to their own process if it is clearly defined in the protocol and/or related SOPs.

This standard requires networks to:

- Track current taxonomic references/authorities used for identification in the taxonomy table
- Apply the current reference/authority for identification from the taxonomy table to the table that uses a taxon ID at the time of record creation (like Protected Status in the Protected Records section above)

The following tables will be used:

TaxonomicStandard lookup table structure:

Name	Data Type	Required	Is Unique	Description
ID	tinyint	Yes	Yes	Unique primary key identifier
Code	Text, 1	Yes	Yes	Unique alpha code for each lookup value
Label	Text, 40	Yes	Yes	Unique name of each lookup value
Summary	Text, 255	Yes	Yes	Definition

TaxonomicStandard lookup table values^{4 5}:

Code	Label	Summary
F	Field Subject Matter Expert	Use if/when data are collected in the field or verified in the office to one or more undetermined standards. Record the position(s) responsible for performing the identifications as the reference/authority.
T	Taxonomic Subject Matter Expert	Use if/when data are collected and verified by an individual with documented taxonomic expertise. Record the name and organization of taxonomic subject matter expert as the reference/authority.
S	Scientific Publication	Use if there is a Scientific Publication that serves as the reference for identification or reporting. Record the publication reference and year of publication (to distinguish among versions) as the reference/authority.

TaxonomicReferenceAuthority_Identification lookup table

Name	Data Type	Required	Is Unique	Description
ID	int	Yes	Yes	Unique primary key identifier
TaxonomicStandardID	tinyint	Yes	No	Foreign key to TaxonomicStandard lookup table
Label	Text, 40	Yes	Yes	For Scientific Publications, use unique short alphanumeric citation that includes abbreviated authors and year of publication (e.g., Welsh et al. 2003). For Taxonomic Subject Matter Experts, use shortened name and organization (e.g., J.

⁴ values and definitions from draft QAP guidance, Taxonomic Standards section (DeVivo, 2016)

⁵ note that these are fixed values; additional values should not be added

				Doe, UGA) For Field Subject Matter Experts, use shortened position responsible for performing identifications (e.g., crew lead w/ID skill).
Summary	Text, 500	Yes	Yes	For Scientific Publications, use complete and unique citation (e.g., Welsh, S. L., N. D. Atwood, S. Goodrich, and L. C. Higgins. 2003. A Utah flora. Third edition, revised. Provo, UT.: Brigham Young University. 912 pp.) For Taxonomic Subject Matter Experts, use full name and organization (e.g., John Doe, University of Georgia) For Field Subject Matter Experts, use full position responsible for performing identifications (e.g., field crew lead with documented identification proficiency).

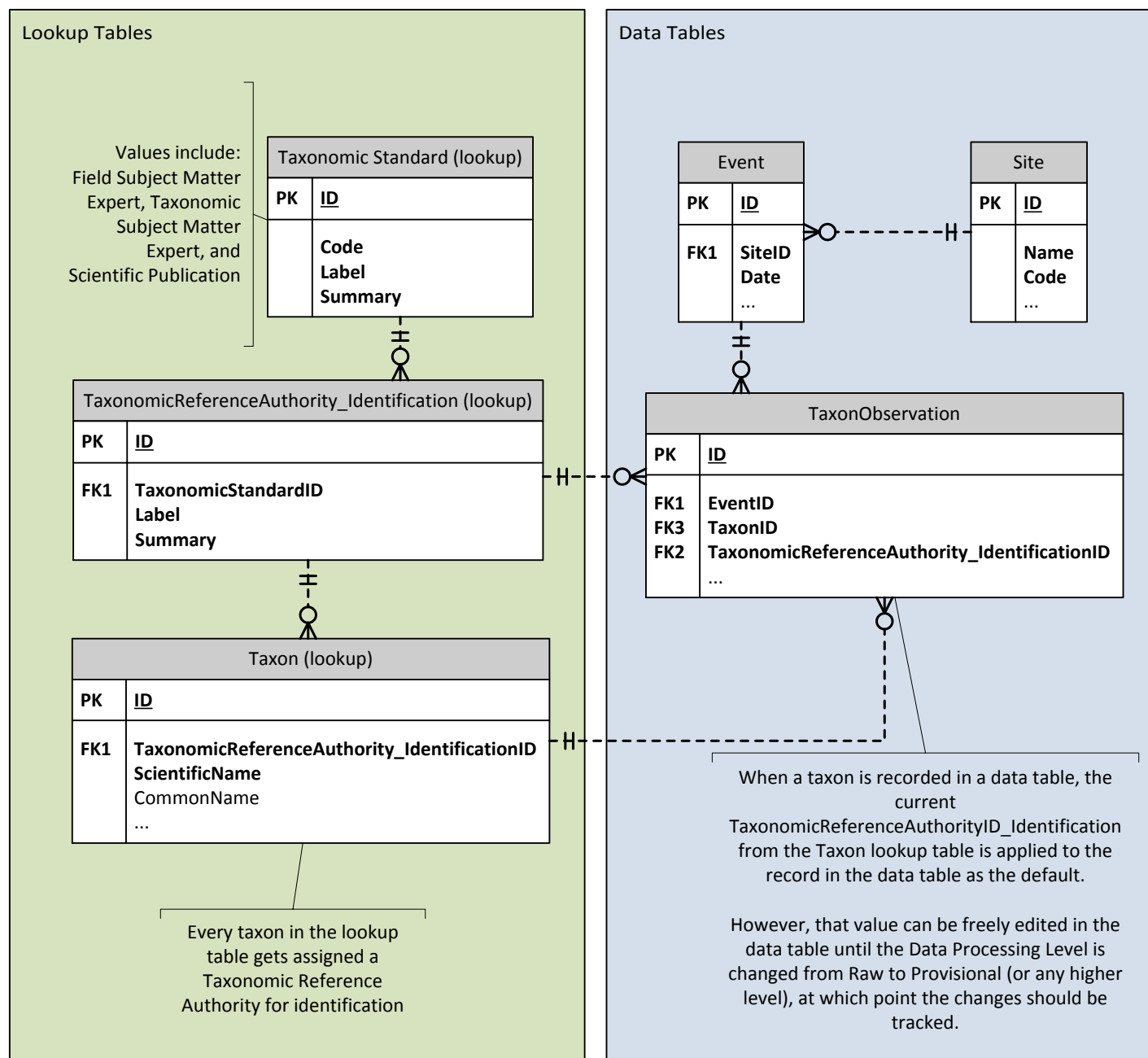
We expect that the taxonomy lookup table will have a `TaxonomicReferenceAuthority_IdentificationID` foreign key that refers to the `TaxonomicReferenceAuthority_Identification` lookup table for the reference/authority used for identification. At the time a taxon ID is used in a data table, the current `TaxonomicReferenceAuthority_IdentificationID` value for the taxon should be applied to a `TaxonomicReferenceAuthority_IdentificationID` column in the data table as a default. This value should be editable in the data table as necessary (see Protected Records section for this model).

Note that for protocols that are about a single taxon (e.g., SWAN Mussels or SWAN Bald Eagles) and do not specifically record the taxon during event measurements, the reference for the taxonomic concept may be recorded in the documentation for the protocol instead of in the database, though if the network requests, we will record the reference in the database.

6.1 Example Database Schema Elements

Taxonomic Concept Database Example – Basic Model (simple)

Assumptions: This model assumes that each taxon has one Identification Reference that applies any time a taxon is identified in the database. If the Identification Reference is not the same for a taxon every time it is identified, see the more complex models from the Protected Records section of this document.



7 References

- [DeVivo, J. 2016. DRAFT Quality Management Plan for the National Park Service Inventory & Monitoring Program: Guidance for Vital Signs Monitoring Protocols \(distributed by K. Gallo email 8/19/2016\)](#)
- [Frakes B and Others. 2015. Inventory and Monitoring Division database standards: September 11, 2015. Natural Resource Report. NPS/NRSS/NRR—2015/1035. National Park Service. Fort Collins, Colorado](#)
- [National Park Service. 2016. Certification guidelines for Inventory and Monitoring data products. National Park Service, Inventory and Monitoring Division. Fort Collins, Colo.](#)