

Данные обследований на пробных площадях и мониторинга (Sampling Event Data)

Исходный текст: <https://github.com/gbif/ipt/wiki/samplingEventData#templates>

Вступление

Таблицы, которые имеют данные не только о находке вида в определенном месте и времени, но подробную характеристику состав сообщества и относительное обилие видов, могут быть опубликованы в виде набора данных формата *Sampling Event Data*. Такие наборы данных часто используют для стандартных учетов или результатов других количественных исследований биоразнообразия. Примерами являются описания растительности, стандартные учеты птиц, маршрутные учеты животных и т. д. Они дополняют данные о находках видов, указывая, какой метод работы был использован, и в идеале относительную численность видов (с помощью подходящего числового показателя), зарегистрированных в пробе или на учете/площадке. Эти наборы данных могут быть использованы для сравнения данных о структуре сообществ в разное время и в разных местах и позволяют исследователям сделать вывод об отсутствии определенного вида в определенном месте.

Как преобразовать ваши данные в набор данных формата Sampling Event Data

В конечном итоге ваши данные должны быть преобразованы в две таблицы формата Darwin Core (DwC:): одна таблица событий и другая таблица находок видов, полученных из (связанных с) каждым событием. Например, первая таблица будет содержать шапки геоботанических описаний, а вторая – находки видов сделанных в каждом из описаний с их обилием и другими полями.



Попробуйте преобразовать свои данные в шаблон Excel (<https://github.com/gbif/ipt/wiki/samplingEventData#templates>), который имеет два листа: один для событий (описание площадок, маршрутных учетов или проб), а другой собственно находок видов в ходе каждого события.

В качестве альтернативы, если ваши данные хранятся в базе данных связанной с IPT, вы можете написать два запроса SQL, используя имена столбцов DwC: одну для событий, а другую для находок.

Каждая запись события должна включать все обязательные поля DwC и как можно больше рекомендуемых полей DwC. Вы можете дополнить свою таблицу дополнительными столбцами DwC, но только терминами DwC из этого списка (http://rs.gbif.org/core/dwc_event_2015_05_29.xml).

Точно так же запись о находках видов должна включать все обязательные поля DwC и как можно больше рекомендуемых полей DwC. Вы можете дополнить свою таблицу дополнительными столбцами DwC, но только терминами DwC из списка. Некоторые термины DwC будут избыточными, т.е. они добавляются как в таблицу событий, так и в таблицу находок. Однако, старайтесь не добавлять избыточные термины с одинаковыми значениями. Хорошо, если они имеют разные значения, например, если вы хотите определить местоположение события, а затем определить более конкретные местоположения для отдельных находок. В противном случае, когда местоположение отдельных экземпляров не указывается, его местоположение (координаты) наследуется от события.

За дополнительными рекомендациями вы можете обратиться к руководству Best Practices (<https://github.com/gbif/ipt/wiki/BestPracticesSamplingEventData>) по публикации данных Sampling Event Data и посмотреть на шаблон, заполненный примерами данных или списком наборов данных этого формата (ниже).

Обязательные поля DwC:

- eventID - также требуется для связанной таблицы находок (чтобы связать их вместе)
- eventDate – дата события
- samplingProtocol – метод отбора проб/описания площадок

Рекомендуемые поля DwC:

- sampleSizeValue & sampleSizeUnit – размер образца/пробы и соответствующая размерность
- parentEventID - в ситуациях, когда событие является частью серии событий
- samplingEffort - предоставить оценку интенсивности отбора пробы/описания на площадке
- locationID - в ситуациях, когда отобранный участок / трансект имеет уникальный идентификатор
- decimalLatitude & decimalLongitude & geodeticDatum - чтобы указать местоположение конкретной точки
- footprintWKT & footprintSRS - для предоставления определенного контура на карте
- countryCode
- occurrenceStatus - только для данных где нужно указать факт отсутствия вида

Примеры наборов данных формата *Sampling Event Data*:

- Набор данных мониторинга бабочек в Израиле (BMS-IL):
<https://www.gbif.org/dataset/647ae6f8-8e26-4189-b448-02b45b7ad884>
- База данных голландской растительности (LVD): <https://www.gbif.org/dataset/740df67d-5663-41a2-9d12-33ec33876c47>
- Исследование жизни рифов: Глобальный набор данных по рифам:
<https://www.gbif.org/dataset/38f06820-08c5-42b2-94f6-47cc3e83a54a>
- Обследование Lepidurus arcticus на северо-востоке Гренландии 2013:
<https://www.gbif.org/dataset/78360224-5493-45fd-a9a0-c336557f09c3>
- Насекомые из световых ловушек (1992–2009), Зоологический музей, Копенгаген:
<https://www.gbif.org/dataset/f506be53-9221-4b44-a41d-5aa0905ec216>

ЧАСТО ЗАДАВАЕМЫЕ ВОПРОСЫ:

В. Как указать, что событие выборки было частью временного ряда?

О. Все события выборки в одном и том же местоположении должны иметь один и тот же locationID.

В. Как опубликовать иерархию событий (рекурсивный тип данных), используя parentEventID?

О. Классическим примером является подвыборка большой площадки. Чтобы сгруппировать все (дочерние) события подвыборки в (родительской) выборке, parentEventID всех событий подвыборки должен быть идентичен eventID (родительской) площадки. Чтобы быть действительными, все parentEventID должны ссылаться на идентификаторы eventID в таблице событий (площадок). В противном случае parentEventID должен иметь уникальный идентификатор (например, DOI, HTTP URI и т. д.), который ссылается на запись события (площадки), описанной в другом месте. В идеале все (дочерние) события подвыборки имеют ту же дату и местоположение, что и (родительское) событие, на которое они ссылаются.

В. Как опубликовать данные об отсутствии?

А. Включите в таблицу событий факты отбора проб (описания площадок), даже если отбор не выявил находок видов. Это позволяет предположить отсутствие видов в этих случаях. Этот пример набора данных о событиях выборки из Норвегии показывает, как это выглядит (https://gbif.vm.ntnu.no/ipt/resource?r=lepidurus-arcticus-survey_northeast-greenland_2013).

В качестве альтернативы, вы можете сделать отсутствие видов явным, добавив запись о находке вида для каждого вида, установив следующие поля:

Обязательное поле:

occurrenceStatus = "отсутствует"

Необязательные поля (укажите одно или оба):

- individualCount = "0"
- organismQuantity & organismQuantityType, например "0 individuals"

Предупреждение: в настоящее время GBIF индексирует все виды видов независимо от того, присутствуют они или отсутствуют. До тех пор, пока эта проблема не будет устранена, GBIF рекомендует исключить все пропуски видов, используя следующий фильтр на странице IPT:

Фильтр: afterTranslation -> adventrenceStatus -> NotEquals -> отсутствует

Более подробную информацию о том, как применить фильтр, можно найти в руководстве пользователя IPT.