

Ringgold External Identify Database Schema

Introduction.....	2
New in this version	2
General	2
Coverage and frequency.....	3
Data supplied.....	3
Institutions and ringgold_ids.....	3
Tables.....	4
Loading Order	4
Packages.....	4
Character set (charset).....	4
Dataset Options	4
Table content description.....	5
Table detail	6
institutions	6
alt_names.....	7
alt_names - note types	7
classifications	8
classifications – classification types	8
consortia	9
deleted_ids.....	9
identifiers	10
identifiers – external identifier types.....	10
ips (optional for an additional fee)	10
multies.....	11
notes	11
sizes	12
sizes – size kinds.....	12
taxonomy	12
tiers	13
tiers – tier types	13
url	14
urls – note types.....	14
XML Database Schema	14

Introduction

This document is designed to provide customers with an overview and detail of organization data files available from Ringgold for their internal use.

New in this version

- ⤴ References to “Parents” replaced with “Institutions” and “p_code” replaced by “ringgold_id”.
- ⤴ Columns in “multies” for “ringgold_id” and “parent_ringgold_id” are different positions.
- ⤴ This version includes two new tables, “classifications” and “taxonomy”. The classifications table contains the information about the new classifications of academic institutions in nine categories. The taxonomy table contains all of the possible values defined by Ringgold and referred to throughout the Schema as indicated in the table description.
- ⤴ “Rg_tiers and types” reference tables are gone as they are redundant and handled by the taxonomy table.
- ⤴ All field names are now lower case as is common practice in many database designs.
- ⤴ The “rec_id”, when exported, now has a consistent name for all tables.
- ⤴ “parents.ext_name” and “parents.ext_city” fields have been removed, the diacritic versions of names formerly appearing in these fields have been moved to the alt_names table coded as DN.
- ⤴ Column “ips.date_created” is now “ips.timestamp” for consistency in all tables.
- ⤴ Added description for table “counts.txt”.
- ⤴ Changed “institution.zip” to “post_code”.
- ⤴ “altnames.ext_name” and “altnames.ext_city” fields have been removed, the diacritic versions of names formerly appearing in these fields have been moved to the alt_names table coded as DN.
- ⤴ Renamed “multies.class” to “relationship”. Updated description for “relationship” and “number” (1-4).
- ⤴ Addenda have been moved to sit with the relevant table and several new tables added.
- ⤴ Open Funder Registry IDs added in the Identifiers table.
- ⤴ 20170131: Multies table definition changes to include all of the hierarchy. No longer just the multi parents of some institutions.
- ⤴ BN Brand Name retired as alt names type. BS Brand Site retired as alt url type.

General

The database consists of 14 tables. Each output format consists of data for each table ready for loading into a target program e.g. a database, other local program or an excel spreadsheet.

The data from Ringgold can be supplied in three different basic formats:

- ⤴ A set of delimited text files - one file per table and a counts.txt file listing the row counts per table for input validation.
- ⤴ A set of SQL insert statements for MySQL
- ⤴ An XML file with associated XML Schema Document for import

Coverage and frequency

The normal routine for data delivery is a complete replacement set of data each week. Data will be delivered by ShareFile. You will have the option of using the ShareFile web interface or an FTP interface to ShareFile with different settings (URL and user account settings that you control). Please note that ShareFile is compatible with most FTP clients.

Data supplied

The data is contained in fourteen data tables within the Identify database. The content of each table is described in the following section. Further technical detail and specification for each table is in the Database Schema section below.

Institutions and ringgold_ids

Historically, organizations were called *parents* because many publisher's subscriber records typically belonged to a single organization (as children). What is now universally referred to as the Ringgold ID in our general documentation was originally termed the *p_code* (parent_code).

With this release, we are changing the table names to better fit the table purpose. **Institutions** will now contain the organizations data and the primary link is the **ringgold_id**

Tables

Loading Order

Depending on your database solution we recommend loading the tables as follows:

1. Load the following reference tables: (so references to these records can be tested during loading, if desired.)
 - a) taxonomy (contains all lists)
2. Load the institutions table (so ringgold_id can be used for foreign keys and validated during load of other tables)
3. Load remaining tables in any order.

Packages

- ⤴ All datasets will be compressed into 'zip' package files compatible with Linux zip and WinZip.

Character set (charset)

- ⤴ UTF-8 is our standard character set format
- ⤴ Latin 1 can be delivered on request
- ⤴ UCS-2 can be delivered on request

Dataset Options

- ⤴ Delimited
- ⤴ MySql Sqldump
- ⤴ XML

Table content description

Table name	Content	Notes
institutions	The core table for the database. Contains the Ringgold ID, name, city, country, type etc.	All other tables are linked to this table via the ringgold_id.
alt_names	The alternate names for an organization e.g. former names, other language variants and acronyms etc.	Linked to institutions via the ringgold_id field.
classifications	A linked table containing different classifications fields for organizations. Links to institutions using ringgold_id.	Available with two options, single aggregated row for each parent+classification where value contains all terms in comma separated list, or multiple rows with one row per parent+classification+term.
consortia	A linked table holding data about consortia and their membership.	Linked to the institutions table at the consortium and member level via the Ringgold ID.
counts.txt	Row counts for primary tables.	Only used for package verification and not normally stored in a database.
identifiers	Contains other systems' identifiers e.g. ISNI (when available), IPED, Athens, OFR IDs.	Linked to institutions table via the Ringgold ID.
deleted_ids	Provides onward links for now disused Ringgold Ids. Contains columns for the old and the new Ringgold IDs	Ringgold Ids become obsolete for many reasons – mergers, duplicates etc.
ips	IP address ranges for organizations.	Optional for an additional fee.
multies	Provides hierarchy links via the Ringgold ID to other organizations that related to an organization as parent organizations.	There can be up to five multiple owners of an organization. Each is linked to the owned organization by the Ringgold ID.
notes	Multiple notes can be linked to any organization via the Ringgold ID. Notes cover diverse subjects.	Linked via the Ringgold ID to the institutions table.
sizes	Demographics to measure the size of an organization. Measurements such as, staff, students (universities), beds and doctors (hospitals).	Linked via the Ringgold ID to the institutions table.
taxonomy	A reference table (or list) of standardized tiers and terms used in exports.	Provides extended descriptions for terms occurring in classifications, rg_tiers, types.
tiers	Contains both the tier type e.g. Carnegie, JISC, RG – for Ringgold Tiers	Linked via the Ringgold ID.
urls	Contains multiple links to institutions table.	Linked via the Ringgold ID to the institutions table.

Table detail

institutions

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold in the production system. NOT the Ringgold ID.
parent_ringgold_id	11	ringgold_id	Provides a link (if present) to the Ringgold ID (ringgold_id) of the organization immediately above in the hierarchy – the owner of this organization. Other values: set to 0 if there is no upward link. Set to 888 if there are multiple parent institutions.
ringgold_id	11	All other related tables	This is the Ringgold ID and is used in other tables to link that data to the core organizational record. This relationship is known as a foreign key relationship.
name	150		Name of the organization in the language of the country where the organization is based. Other names are in the alt_names table. Diacritical characters are normalized to non-diacritical equivalents.
city	60		City name – no diacritics.
post_code	20		Zip or postcode.
country	4		ISO-3166 2 digit code. Hong Kong is CN with the state of Hong Kong, Macau is CN with the state of Macau and Puerto Rico is USA with the state of Puerto Rico.
state	25		State/region/county/province or other admin area.
type	50		One of a set of codes used to characterize the activity of any organization e.g. academic/medsch for a teaching hospital. See current types table for possible values.
timestamp			SQL timestamp field for the time/date added to the database.

alt_names

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold in the production system.
ringgold_id	11	Institutions table	The link to the organization record in the institutions table.
name	150		Alternate name e.g. library name, other language name, former name, acronym, diacritics, etc.
city	60		City including diacritics
country	4		ISO 3166-2 2-digit code for the country.
language	35		Future use.
notes	100		See table below for current possible values and descriptions.
timestamp			SQL timestamp field for the time/date added to the database.

alt_names - note types

Note	Description
AB	Abbreviation
AB TL	Abbreviation of the transliterated name
AB TR	Abbreviation of translated name into English
AC	Acronym
AC FN	Acronym of former name
AC TL	Acronym of the transliterated name
AC TR	Acronym of translated name in English
AK	A/K/A (also known as)
AK FN	A/K/A (also known as) of former name
AK TL	A/K/A (also known as) of transliterated name
AK TR	A/K/A (also known as) translated into English
AQ	Acquired (with year if known)
CS	Ceased (with year if known)
DN	Diacritic name
FN	Former name
FN LB	Former name of library
FN until	Former name(until YYYY if known)
IM	Imprint (deprecated)
LB	Library name
LB TR	Library name translated to English
MG	Merged (with year if known)
TL	Transliterated name
TL FN	Transliterated former name
TR	Translated name in English
TR FN	Former name translated into English

classifications

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold in the production system.
ringgold_id	11	Institutions (ringgold_id)	The link to the member Ringgold ID in the institutions table.
classification	255		See table below for current possible values and descriptions.
value	65k text		Classification terms (see taxonomy table for acceptable terms).
timestamp			SQL timestamp field for the time/date added to the database.

classifications – classification types

Classification	Description
activity_status	Status of the organization's operation
economic_model	The economic aim of the organization
governance	The way the organization is managed, by whom and for whom
level	Place of organization in the fixed facet hierarchy
mission	What the organization does, is trying to achieve, what are its reasons for existing and purposes
number_of_sites	Whether the organization has one single site or multiple sites
self_description	A label used by the organization to describe itself
reach	The reach of the organization in terms of its locations and services to its audience and customers
subjects	Areas the organization is involved in

consortia

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold in the production system.
consortia_ringgold_id	11	Institutions (ringgold_id)	A link to the organizational record for the consortium.
member_ringgold_id	11	Institutions (ringgold_id)	The link to the member Ringgold ID in the institutions table. When one record has the same member_ringgold_id as consortia_ringgold_id, then the name in the next column is the consortium name not the member name.
name	230		The name of the member of the consortium. Or the name of the consortium if the member_ringgold_id = consortia_ringgold_id.
timestamp			SQL timestamp field for the time/date added to the database.

deleted_ids

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold in the production system.
old_ringgold_id	11	Institutions table	The old link to the organization record in the institutions table – now deleted.
new_ringgold_id	11	Institutions table	The new link to the organization record in the institutions table.
details	100		Contains the reason for deletion.
timestamp			SQL timestamp field for the time/date added to the database.

identifiers

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold in the production system.
ringgold_id	11	Institutions table	The link to the organization record in the institutions table.
identifier_type	30		See table below for current possible values and descriptions.
value	255		The value of the identifier – usually an integer.
timestamp			SQL timestamp field for the time/date added to the database.

identifiers – external identifier types

Identifier	Description
ATHENS	Athens Identifier (see: http://www.openathens.net)
ISNI	International Standard Name Identifier (see: http://www.isni.org) (when available)
IPED	Integrated Post-Secondary Education Data System (see: http://nces.ed.gov/ipeds)
NCES	National Center for Educational Statistics (see: http://nces.ed.gov)
OFR	Open Funder ID see (http://www.crossref.org/fundingdata/registry.html)

ips (optional for an additional fee)

[IP addresses]

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold in the production system.
ringgold_id	20	Institutions table	The link to the organization record in the institutions table.
ip	255		
ip_low	15		
ip_high	15		
ip_int_low	11		
ip_int_high	11		
timestamp			SQL timestamp field for the time/date added to the database.

multies

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold.
ringgold_id	11	Institutions table	The link to the organization record in the institutions table to which the parent_ringgold_id is the owner – above in the hierarchy.
parent_ringgold_id	11	Institutions table	Provides a link to the Ringgold ID (ringgold_id) of the organization(s) immediately above in the hierarchy – the owner of this organization.
relationship	30		The type of ownership relationship that the parent_ringgold_id organization has with the ringgold_id organization. Currently defaults to “admin”. [future use]
number	1		Integer 0 or 1. Zero indicates only single parent. One indicates 2 or more parent institutions.
timestamp			SQL timestamp field for the time/date added to the database.

notes

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold.
ringgold_id	20	Institutions table	The link to the organization record in the institutions table.
notes	255		Text of note
timestamp			SQL timestamp field for the time/date added to the database.

sizes

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold.
ringgold_id	11	Institutions table	The link to the organization record in the institutions table.
size_type	30		See table below current possible values and descriptions.
value	255		Value associated with the “kind” of size.
timestamp			SQL timestamp field for the time/date added to the database.

sizes – size kinds

Kind	Description
beds	Number of beds in a hospital
doctors	Number of doctors in a hospital
hospitals	Number of hospitals owned by the Group
libraries	Number of libraries in an organization
schools	Number of schools in an administrative grouping
size	Number of employees or for academic organizations, FTEs
staff	Number of specialist staff (lawyers, chartered accountants, faculty etc)

taxonomy

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold.
vocabulary	255		Name of the vocabulary list.
name	255		Term in the vocabulary.
description	1024		Description of the term.
timestamp			SQL timestamp field for the time/date added to the database.

tiers

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold.
ringgold_id	11	Institutions table	The link to the organization record in the institutions table.
tier_type	25		See table below for current possible values and descriptions.
value	40		Value for the tier type
timestamp			SQL timestamp field for the time/date added to the database.

tiers – tier types

Tier Type	Description
Carnegie-BASIC2005	2005 Basic Classification – Carnegie classifications (see http://carnegieclassifications.iu.edu/)
Carnegie-BASIC2010	2010 Basic Classification
Carnegie-SizeSet2005	2005 Size and Setting Classification
Carnegie-SizeSet2010	2010 Size and Setting Classification
JISC	Joint Information Systems Committee band (UK) (see http://www.jisc-collections.ac.uk/Help-and-information/JISC-Banding/)
RGT	Ringgold tier
SFJ	Shop for Journals pricing tier (Highwire press) (see http://highwire.stanford.edu/tfocis/)

url

Column	Length	Related to	Content
rec_id	11	None	An internal id used by Ringgold .
ringgold_id	11	Institutions table	The link to the organization record in the institutions table.
url	200		The URL and any alternative URLs for the organization .
notes	100		Reason for url. E.g. "AS for Alternative site". See table below for current possible values and descriptions.
timestamp			SQL timestamp field for the time/date added to the database.

urls – note types

Note	Description
AS	Alternative site
ES	Email site
FS	Former site
GS	Graduate site (Japan only)
LS	Library site
PD	Primary domain

XML Database Schema

Is supplied with the XML database export in ShareFile.