

Identify Audit Service

Data Exchange Document

Introduction

This document has two functions, first to provide a standard specification for clients to provide their source data files to Ringgold for auditing, and second to explain the deliverables returned to the client when the audit is complete.

1. Transmission of records to Ringgold

In order for Ringgold to perform audits of client data, we require the consistent submission of the data sets that you wish us to clean and to assign Ringgold IDs and metadata. In order to do this it is important that Ringgold receives files which include as much information as possible regarding the organizations contained therein. Please submit as much information as you hold for the end-user organization as well as any information in “bill to” or “supply to” fields, end-user information is the most important in accurately identifying the institution.

1.1 Data to include

The data included in your files is not only required for our programs, but it can help our researchers to improve matches between your records and the Identify Database records.

1.1.1 Minimum requirements

As a minimum we require the following data elements to be submitted. Please ensure each element is in a separate field; i.e., do not concatenate all address elements into a single field.

Field Name	Description of contents	Example
Customer number	The unique ID from your record(s) that identifies the organization/customer. Can be numeric or alpha numeric – but must be unique.	12346562
Organization name	The name for the institution or organization in your record(s).	Roche Pharmaceuticals or Penn State University
Department (where available)	A subordinate part of the main organization	Library Chemistry department
Address field (where available)	Street address	121, Wood Green PO Box 24
City	The city in which the organization is located. Essential to sort out the many duplicate names.	Portland London Mumbai
State (where available)	Political division within which the city exists. Different countries have different levels and names for these divisions; e.g., county, province.	Florida or FL Hertfordshire or Herts Quebec
Postcode/zip (where available)	The post or zip code for the organization. Provides reasonably unambiguous data for identifying organizations.	60637 (zip) BS1 4DW (UK postcode) DE 65929 (German postcode)
Country	Country where organization is located. Use full name or a standard code. Include for all countries including your “home” country.	United States or US UK or GB Germany or DE

1.1.2 Useful fields

We can also find the following fields extremely useful:

Field Name	Description of contents	Example
Institution URL	Domain name	www.university.ac.uk
Delivery address	Where the main information on a record is for an agent, we can apply a more accurate Ringgold Identifier if we have the end user information	Penn State University, Roche Pharmaceuticals

1.1.3 Additional fields

It is also possible to submit additional subscription-level data (order level) for each customer, which can be returned in the deliverables. However, please note that the three deliverable formats vary in the types and number of additional fields that can be displayed or returned. Information on the deliverables follows in Section 2.

This is just a list of suggestions and you may have other fields which you wish to supply to us:

Field Name	Description of contents	Examples
Product name/product code	Name of the title and/or product code.	Journal name or acronym Package name or acronym
Agent	Information on any intermediary, including subscription agents, library booksellers, etc.	EBSCO, USACO, Dawson, Accucoms
Price	Sales amount in your local currency.	\$1565.00, £1200.00
Invoice date	Date customer invoiced, in numeric format.	25/06/2014
Renewal date	Date of renewal of subscription, in numeric format.	01/29/2015
Package details	Package containing Product	Consortium X, Arts Collection
Customer tier	Your internal customer tier.	A5, 12
Sales territory	Your own geographic division of sales territory.	Far East Asia, EMEA
Sales rep	The representative – either internal or external.	iGroup, John Smith
Reference number	Any internal reference numbers that you find useful.	123DFG456
Format	Any format information.	Electronic, Print only, P+E
Consortium	Name or code for the consortium which subscribed	WALDO, NESLI2014
Notes	Any notes you may have.	Customer is on 2 year renewal.

1.2 File Format

Ringgold can accept the following file formats, please see the table for further instructions:

- Field delimited formats such as comma separated value (csv) or tab or pipe separated formats.
- Excel files
- Text files, .txt, which will require the same formatting components as a .csv file (see below)

File Type	CSV	Tab	Excel	Notes
Format	Several programs can produce CSV files. In Excel there is a “save as” option to output the file as csv.	Several programs can produce Tab delimited files. In Excel there is a “save as” option to output as tab delimited.	A standard Excel .xls or .xlsx extension.	Only one worksheet tab is allowed for this format.
Field delimiter	Comma (each field is separated from the next by a comma)	Tab (each field is separated from the next by a “tab” character)	Not required	
Field surround character	Double quotes are recommended to avoid problems with commas in the field data string	Not required	Not required other than to display unusual information as text, whereby it should be preceded by ‘	
End of record character	Carriage return / Paragraph character	Carriage return / Paragraph character		Required for text-based files rather than spreadsheet format files.
Problem characters	Avoid using high-order ASCII characters such as “smart” quotes, apostrophes and control characters in the data, or please advise Ringgold of the use of such characters within your data.			

1.3 File consistency

The data within the file can be organized as you desire, bearing in mind the following:

- File layout being consistent within the file – if you add files together from different sources – make sure that the data in each file is aligned. For example, that columns are consistently applied. If you cannot guarantee this, then please provide separate files.
- Consistent layout between dataset submissions – if you are providing data on a regular basis please keep the layout the same for each submission. If the format changes please inform Ringgold of the changes.
- Data consistent within the fields – we are aware that this is not always possible – however, it is desirable for the contents of fields within a column to be consistent. For example, that the organization name contains the names of organizations rather than the first line of a street or mailing address or an individual’s name.
- Consistent customer numbers between data submissions – we use the customer number to map a new data submission to the previous one, and identify new records.

2. Deliverables

Ringgold returns files in two delivery formats following the audit:

1. Completed returned audit files from Ringgold in the standard format for uploading to internal systems. (System Files)

- Completed audit files which are uploaded to Identify Online for analysis within the *Identify* interface. (IDO Data Upload), optional for Identify Database licence customers.

The Ringgold audit service returns the client's customer identification number with the addition of the Ringgold Identifier – a unique number to identify the organization – and associated metadata. There are two delivery options. Clients may select either or both of these options:

2.1 System Files

Completed returned audit files from Ringgold in the standard format for uploading to internal systems. This consists of five .csv files, as follows:

2.1.1 Subs File

This file contains the original Customer Number, and original location data from the customer, mapped to the Ringgold ID and Ringgold Type for the organization. The original location data may be that which was attached to the shipping/subscription agent field if no location information is available for the end-user. If data was not consistent within the file, there may be blanks in a field, which indicate where data was inconsistent.

Column heading	Definition
Customer Account ID	The customer number from the original data
Ringgold ID	The organization's Ringgold ID
Type	Ringgold classification of sector and subsector
Account name	The customer name as supplied in the source file
City	The customer city as supplied in the source file
State	The customer state/province/county as supplied in the source file
Zip	The customer zip/post code as supplied in the source file
Country	Country as supplied in the original subscriber file

2.1.2 Parents File

This file contains the descriptive metadata (except for alternative names) from the Identify Database for each organization assigned to a customer record in the Subs File. It contains the links to the parent organization which is the next level up in the hierarchy for each organization.

Column heading	Definition
Parent ID	The Ringgold ID for the next level up in the institutional hierarchy of the organization. The parent or owning organization.
Ringgold ID	The organization's Ringgold ID to which the metadata applies.
Name	The official name of the organization
City	The city where the organization is based
State	The state or provincial governmental authority
Zip	The zip or post code
Country	The country of the organization
Size	For academic organizations, the FTE, for others, the number of employees
Academic Staff	The number of faculty, or doctors, lawyers, depending on academic type
Type	Ringgold classification for institutional sector and subsector
Url	. Uniform Resource Locator (URL) or website.
JISC Band	JISC Band applied to the organization (UK records only)
IPEDS	Integrated Postsecondary Education Data System number (US only)
Notes	Contains additional information depending on institutional characteristics
Carnegie	Carnegie2015 Classification applied to an academic institution (US records only)

RG Tier	Ringgold's own tier assignment
ISNI number	ISNI number
OFR ID	Open Funder Registry ID

2.1.3 Alt Names File

This file contains alternative names (abbreviations, acronyms, former names, translations etc.) for all organizations in the Parents File.

Column heading	Definition
Ringgold ID	The organization's Ringgold ID
Alt name	The alternative names of the organization, one name per row
City	City where the organization is based
Country	Country for the alternative name – may be different from the main organization record in some circumstances (mergers, closures, acquisitions)

2.1.4 Consortia File

This file shows details of the consortia which the organizations in the Subs File are members; it includes metadata on each consortium itself.

Column heading	Definition
Ringgold Consortium ID	The consortium's Ringgold ID
Consortium Name	The name of the consortium
City	City where the consortium is based
State	The state or provincial governmental authority
Zip	Postal code for the consortium
Country	The country where the consortium is based
Member Ringgold ID	Ringgold number for the consortium member that appears in the Subs File.
Member Name	Name of the consortium member that appears in the Subs File.
City	Ringgold city name for the consortium member
State	Ringgold state name for the consortium member
Zip	Ringgold zip/postal code for the consortium member
Country	Ringgold country for the consortium member

2.1.5 Unallocated and Distributor File

Occasionally customer records cannot be identified properly. Ringgold returns these records as a separate Excel file with three worksheets containing the following information:

Distributors - records sometimes cannot be audited if there is insufficient end-user information. Where end-user information is lacking Ringgold allocates the Ringgold ID for the distributor (e.g. subscription agent, bookseller, sales agent), which is returned in this worksheet. This information is also included in the Subs file.

Personal – the customer record is for an individual's home address and it is not possible to unambiguously link the customer to an institution. Ringgold returns these records, unallocated in this worksheet.

Pass – it is not possible to unambiguously link what is probably an institutional address to a particular institution. These are returned unallocated in this worksheet. There are usually very few instances of this in returned audit files.

2.1.6 Academic Classifications file

This file is returned for Academic records only. The nine new areas of classification include:

- Economic Model
- Governance- the way the organization is managed
- Reach- geographical scope of the organization
- Mission - what the organization is trying to achieve and what its purpose is
- Granular Subjects - five hierarchical subject levels and over 800 subjects
- Self-Description- how an organization describes itself
- Fixed Hierarchy Level - nine Ringgold levels
- Activity Status- whether an organization is active or not
- Number of Sites- single or multiple

2.2 Identify Data Upload (optional)

For clients of auditing who also license the Identify Database, both subscriber (customer level) and subscription-level (order level) data can be held in a secure custom view of Identify.

The Identify Online database contains the following fields:

Identify Field Name	Contents
Ringgold ID	Ringgold assigned institutional identifier
Institution Name	Legal name in native language (where appropriate)
Alternative Names	Commonly used names, English language name, name of library, acquisitions, former names, acronyms
Location	City, State, ZIP/Postcode, Country
URL	Domain, sub-domain or specific department URL
Ringgold Type	Ringgold classification, provided as sector, e.g. Academic, followed by "/" and the subsector, usually subject related e.g. medical
Ringgold Tier (RGT)	Each organization is assigned a Ringgold Tier.
SFJ Tier	Shop For Journals tier
Carnegie 2010	For US academic institutions, the basic Carnegie tier
JISC Band	For UK academic and some research institutions, the current JISC Band
Size	Number of students, employees (depending on sector)
Staff	Number of professional staff (as appropriate, e.g. academic staff, doctors)
Hierarchical Relationship	Links to parent or owning institution(s) and to subordinate institutions
Consortia	The consortia of which the institution is a member
Notes	Free text notes including estimated size, e.g. number of members of a society
ISNI number	ISNI number

In addition we are able to provide you with an upload of some of your own data in a secure view of Identify Online to enable you to run analyses and exports. The data may be held at both the subscriber and subscription level, based upon what you supply us. *Identify* holds both subscriber (customer) and subscription (order level) information for analysis within the system, including counts.

The fields are as follows:

Customer field	Description
Customer ID	The unique ID for the client's customer can be displayed in <i>Identify</i>
Customer Name	
Product Name	Name of the title or product code may be analyzed, displayed and exported from <i>Identify</i> .
Product details	Format or package code
Dates	
Quantity	Number of subscriptions (if no subscription count available Ringgold will assume 1). Can be analyzed, displayed and exported from <i>Identify</i>
Tier or customer class	
Agent	Information on any intermediary, including subscription agents, library booksellers, etc displayed in <i>Identify</i> .
Consortia or multisite deal name	
Address fields	As supplied by customer
Client data 1	A field in your originating data may be selected for inclusion in your version of <i>Identify</i> for display.
Client data 2	A field in your originating data may be selected for inclusion in your version of <i>Identify</i> for display.
Client data 3	A field in your originating data may be selected for inclusion in your version of <i>Identify</i> for display.
Additional Client data fields	We can incorporate a large number of client fields

For further information on how to use Identify Online please see the [Identify Online User Guide](#).