

# 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 that are returned to the client when the audit is complete.

## 1. Transmission of records to Ringgolds

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. Please submit as much address information as you hold for the end-user organization to help us accurately identify the institution.

### 1.1 Data for inclusion

#### 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
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 postal 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
Department	A subordinate part of the main organization	Department of Chemistry

### 1.1.3. Additional fields

It is also possible to submit additional data for each account, which Identify Database clients may choose to host in the client’s secure instance of the Identify Database online interface.

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 Database name or acronym
Format	Any format information.	Electronic, Print only, P+E
Price	Sales amount in your local currency.	\$1565.00, £1200.00
Invoice date	Date customer invoiced, in numeric format.	25/06/2021
Service Start and/or End Dates	Date that the particular ordered item begins and/or ends	01/01/2023
Package details	Package containing Product	Consortium X, Arts Collection
Tier and/or Rate Code	Internal fields used to price the particular client, order, or product	A5, 12
Sales territory	Your own geographic division of sales territory.	Far East Asia, EMEA
Reference number	Any internal reference numbers that you find useful.	123DFG456
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	Tab	Not required	
Field surround character	Double quotes are recommended	Not required	Only to display as text, where text should be preceded by ‘	
End of record character	Carriage return / Paragraph character	Carriage return / Paragraph character		Required for text-based files
Problem characters	Avoid using high-order ASCII characters such as “smart” quotes, apostrophes and control characters in the data			

## 1.3. File consistency

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

- Use consistent account numbers between data submissions. We use the customer number to map a new data submission to the previous one, and identify new records.
- Use 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.
- Keep the 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 field always contains the name of an organization rather than the first line of a street, a mailing address, or an individual's name.

## 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)
2. Uploads to Identify Online for analysis within the Identify Database online interface (optional for active Identify Database clients).

### 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 is the primary file in the set. It contains the mapping of the original account ID provided to the equivalent Ringgold Identifier. This file also contains the organization name and address information provided to Ringgold. The number of records in this file may be less than was originally supplied, either because of duplication of account IDs, or because we have not been able to determine the equivalent Identify Database record and allocate a Ringgold ID. Any records that fall into the latter category are returned in the Unallocated and Distributor file (see below).

All subsequent files in this set contain only Ringgold Identify Database data, linked via Ringgold ID.

Subs deliverable file overview:

Column Header	Description
Customer Account ID	Unique ID provided in source data
Ringgold ID	Assigned Ringgold ID
Type	Ringgold Type
Account name	Organization name provided in source data
City	Organization city provided in source data
State	Organization state/province/equivalent provided in source data
Zip	Organization zip/postal code provided in source data
Country	Organization country provided in source data

## 2.1.2. Parents File

This file contains much of the descriptive metadata from the Identify Database for each organization assigned to an organization record in the Subs file. This file also contains the relationship to the parent organization which is the next level up in the Identify Database hierarchy for each organization, and so on up to the top the hierarchy. Each hierarchical parent record also includes the appropriate Identify Database metadata. Because parent records are included, there are more records in this file than in the Subs file. It is possible for systems to reconstruct a hierarchy by following the Ringgold IDs and Parent ID columns; see example below:

Parent ID	Ringgold ID	Name	City	State	Notes:
1438	115095	University of California Berkeley College of Engineering	Berkeley	CA	<i>Assigned ID from subs file</i>
1439	1438	University of California Berkeley	Berkeley	CA	
96640	1439	University of California System	Oakland	CA	
0	96640	State of California	Sacramento	CA	<i>Highest level of the hierarchy, as there is no Parent ID referenced.</i>

Here is another representation of this hierarchy, from the Identify Online interface:

Ringgold ID	Name
96640	State of California
1439	University of California System
1438	University of California Berkeley
115095	University of California Berkeley College of Engineering

The upward hierarchy records enable you to reveal related organizations in the audit. There will be some fields which appear to contain no data, such as size or state. This is because that metadata element does not apply to the institution, or because we have not been able to obtain that piece of data from an authoritative source.

Parents deliverable file overview:

Column Header	Description
Parent ID	Unique numerical organizational identifier of direct parent
Ringgold ID	Unique numerical organizational identifier, assigned by Ringgold
Name	Organization Name
City	City of administrative HQ
State	State, province, or equivalent.
Zip	Postal Code or Zip Code, part of the mailing address of the institution in the format used in each country
Country	Country of administrative HQ in ISO3166 2 digit country code format
Size	Definition depends on category, e.g., Number of students for academic organizations (where possible FTE figures are used but may be total enrollment). Size represents the total number of employees for other organization types.

Academic Staff	Number of faculty or professionals (e.g. lawyers)
Type	Ringgold Type. Specifically developed to categorize organizations worldwide. The classification has two parts: sector/subject or type.
Url	Main domain, sub-domain or specific department URL.
JISC Band	JISC band assignment
IPEDS	Numerical identifiers assigned by the Integrated Postsecondary Education Data System.
Notes	Miscellaneous notes
Carnegie	Carnegie Basic Classification
RG Tier	Ringgold Tier. Specifically developed to indicate industry sector, size and type of the organization worldwide.
ISNI	International Standard Name Identifier

### 2.1.3. Alt Names File

This file contains alternative names for all organizations in the Parents file, one row per alternate name. These include translations, acronyms, former names, names of acquired companies, and other versions of names. The altnames file, if loaded into your internal systems, can support better searching and entry points.

Altnames deliverable file overview:

Column Header	Description
Ringgold ID	Unique numerical organizational identifier, assigned by Ringgold
Alt name	Alternate name
City	City of administrative HQ, affiliated with alternate name
Country	Country of administrative HQ in ISO3166 2-digit country code format, affiliated with alternate name

### 2.1.4. Consortia File

This file contains details of every consortium that has members within the Parents file, with one row for each active consortium relationship from the Identify Database. Consortia and member organizations may appear more than once, particularly if an institution is a member of several consortia. This information can be used to assess sales and partnership possibilities for consortium deals.

Consortia deliverable file overview:

Column Header	Description
Ringgold consortium ID	Unique numerical organizational identifier of the related consortium
Consortium Name	Consortium name
City	City of consortium's administrative HQ
State	State, province, or equivalent of consortium's administrative HQ
Zip	Postal Code or Zip Code of consortium's administrative HQ
Country	Country of consortium's administrative HQ in ISO3166 2 digit country code format
Member Name	Member organization's official Identify Database organization name

Member Ringgold ID	Unique numerical organizational identifier of the member organization
City	City of member organization's administrative HQ
State	State, province, or equivalent of member organization's administrative HQ
Zip	Postal Code or Zip Code of member organization's administrative HQ
Country	Country of member organization's administrative HQ in ISO3166 2-digit country code format

## 2.1.5. Classifications file

This file contains nine additional metadata elements, applied only to academic organizations. Classifications data enables granular analysis of academic organizations using metadata such as subject interest, reach, governance etc. Definitions of the values contained in these fields can be found [in this overview of Identify Database academic classifications](#).

Classifications deliverable file overview:

Column Header	Description
Ringgold ID	Unique numerical organizational identifier, assigned by Ringgold.
Ringgold Name	Organization Name
activity_status	The current operational status of an organization.
economic_model	The legal status of the organization, for-profit or non-profit.
governance	The way the organization is managed, by whom and for whom.
level	Nine numerical designations assigned to a fixed level within an Identify Hierarchy.
mission	What an organization does, is trying to achieve, and what its purpose is.
number_of_sites	Contiguous locations that an organization operates from; either single or multi-site.
self_description	The simple terms used by an organization to describe itself or its role.
reach	The geographical scope of the organization.
subjects	A description of what the specific organization is involved in, pursues and investigates. These are defined in terms of academic or vocational discipline.

## 2.1.6. Unallocated and Distributor file

Occasionally supplied organization records cannot be mapped unambiguously to a single institution within the Identify Database. Ringgold returns these records as a separate Excel file with three worksheets containing the following information:

**Distributors** 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. The data in this file can be used to request enhanced end user information from the agents and distributors. If enhanced information is supplied back to Ringgold as part of an Audit Update, we can re-audit the record, and allocate the end user Ringgold ID instead.

**Personal** The provided organization record refers only to an individual's name and/or home address; it is therefore not possible to unambiguously link the customer record to an organization. Ringgold returns these records, unaudited, in this worksheet. If further information is supplied in a subsequent Audit Update, we will re-audit the record, and map it to the correct Identify Database record.

**Pass** This sheet contains all the records where we have not received enough information to identify a single organization. If further information is supplied in a subsequent Audit Update, we will re-audit the record, and map it to the correct Identify Database record.

For further assistance please contact [support@ringgold.com](mailto:support@ringgold.com)