

Editorial Policy for Ringgold Identify Database

Table of Contents

Table of Contents.....	1
1. Introduction	2
2. Scope of the Identify Database.....	2
2.1 Hierarchical relationships	3
3. Identify Database	3
3.1 Mandatory Identify Metadata	4
3.2 Non-mandatory Identify Metadata.....	5
3.3 Material Changes	8

1. Introduction

This document outlines the policy for inclusion of organizations and metadata in Ringgold's Identify Database.

Ringgold's *Identify* has been developed to serve a wide range of needs throughout the scholarly communications ecosystem. *Identify* contains uniquely identified institutions, each institution or part thereof is assigned a Ringgold Identifier. The organizational records provide extensive additional metadata and are mapped into their respective hierarchies and provide relationships with other organizations.

The database includes any identifiable organization, and its component parts that have an interest in scholarly communications, including as consumers, creators, funders, or intermediaries in the scholarly supply chain.

2. Scope of the Identify Database

Identify covers all sectors of the scholarly communications environment, including academic, corporate, government (both national and local), healthcare and non-profit organisations. It is global in reach and identifies both main organizations and their subsidiary parts (children) in structured hierarchies.

Identify has been developed since 2003 and new records have been added primarily via our provision of the Identify Audit Service for a range of clients, particularly publishers. New records are created when Ringgold becomes aware of them, in auditing client files, and through market research. New records are also created when an organizational entity requires a parent organization in order to fit into a hierarchy, for example a department must have an institution of which it is a component part.

Institutions are not added under the following circumstances:

Libraries

Libraries are not included unless they are standalone or serve more than one legally separated organization. Examples of libraries which are included in *Identify* are, but not limited to, the following:

- Stand-alone national and public libraries
- University libraries where they provide services to the university itself and also function as a regional or national library
- Health libraries, where they serve university health faculties and departments, and also a local hospital or local health system
- Independent archives

Temporary programs

Temporary programs, are not included.

Individual retail outlets

Individual retail outlets are not included, with the exception of book stores, including campus book stores.

Admissions and Administrative Offices of Universities

Admissions or administrative offices of universities or colleges are not included unless they appear in client files received as part of the Identify Audit Services.

Publications and Products

Publications and products are not added as separate records unless they are a legally recognized organization in their own right, for example a company may use the name of its main product, or a newspaper may be a legally recognized company. The Ringgold ID for the owning organization would be assigned. Publications, e.g. journals, databases, or book titles are not added as a separate record, but imprints of a publisher are provided with a separate *Identify* record and Ringgold ID.

Fictional Organizations

The Identify Database does not include fictional organizations.

2.1 Hierarchical relationships

Organizational entities are mapped into Identify Hierarchies based on demonstrable management of the subsidiary part. Other relationships appear as linked relationships, for example, consortium members are not owned or managed by the consortium, but are members. Imprints are considered to be organizational entities, even if they are not institutions in their own right and are mapped into the hierarchy of a publisher or owning institution. Where an institution is related to more than one organization, multiple IDs are listed.

3. Identify Database

Identify is a database of over 440,000 organizational entities and consortia. Each entity or part thereof is assigned a unique Ringgold Identifier. These entities consume, produce, supply, fund or have another demonstrable interest in scholarly and online content.

Each record contains basic metadata: Ringgold ID, the name, location and URL (where a website is available).

Records contain additional metadata where applicable and reliably sourced:

- Alternative names
- Metric data: Size, Staff, Number of Beds, Number of Hospitals, Number of Doctors, etc.
- Classification data: Ringgold Type, Ringgold Tier, Carnegie Classification, JISC Band, Ringgold's new classifications.
- Additional Identifiers: ISNI, NCES, IPEDS, OFR
- Access and Authentication Data: Athens Codes, IP Address Ranges

The database is organized hierarchically. Each individual entity's record is, where appropriate, mapped to the organization's larger structural hierarchy, a 'family tree'.

The database has become more granular over time and Ringgold routinely adds the first subdivision in the hierarchies of academic and research institutions and within hospitals. The institutions contained in the database can be broadly grouped into six sectors:

- Academic
- Hospital/Health systems

- Government, central and local
- Corporate
- Non-Profit organizations
- Consortia

3.1 Mandatory Identify Metadata

For each uniquely identified organization the following metadata is provided. The first five metadata categories are compulsory for any institution. All other data will be provided where the data is reliably available or the data category applies to the organizational record.

3.1.1 Ringgold Identifier

Ringgold IDs are numeric, unique and sequentially assigned. They contain between four and six digits, but there is no limit to the number of digits, and the individual digits do not signify anything. Ringgold IDs are not recycled, so the number for a record which is deleted is never used again.

3.1.2 Institution Name

The legal name is given in the native language or transliteration to Latin characters.

- For institutions that provide their own name only in a non-native language, the name used by the institution is provided in the name field.
- If the institution provides its name in multiple languages, the native language name is used in the name field and the name in other languages are provided in the alternative name field.

For the transliteration of languages that use non-Latin characters, ISO or international accepted standards are used where available.

- We transliterate: Russian, Belarusian, Ukrainian, Serbian, Bulgarian, Greek, and Japanese.
- For others we provide the name in English.

The format for subdivisions is the name of the main institution followed by the name of the division.

- There are no commas in main names
- Campuses are preceded by a hyphen
- 'The' is never used
- Main names do not contain quote marks
- All abbreviations are spelled out, except Inc, GmbH, Ltd, LLC etc.
- No diacritics are used
- The ampersand (&) is only used when a legal part of the main name

3.1.3 City

Ringgold has identified the problem in the variances in the naming of world-wide cities and is developing an authority file which will include a number of standard names and sources, e.g. UNLOCODE, Google Places.

3.1.4 Country

Provided as both the full country name and the ISO3166 2 digit country code.

3.1.5 URL

Main domain name, but not a long string. These are checked and amended every time a record is reviewed. <http://> is not included, <https://> is included to enable the link.

3.2 Non-mandatory Identify Metadata

3.2.1 Region/State

Ringgold has identified the problem in the variances in the naming of world-wide states and regions, and is developing an authority file which will include a number of standard names and sources, e.g. UNLOCODE, Google Places.

3.2.2 Post Code

Postal Code or Zip code part of the mailing address of the institution in the format used in each country.

3.2.3 Alternative Names

There are several forms of alternative name. These include:

TR = English translation of the institution's name, which is mandatory for organizations in countries where the first language is not English. Corporate designations and company names are not translated

AB = Abbreviation

AB TL = Abbreviation of the transliterated name

AB TR = Abbreviation of translated name into English

AC = Acronym, added if it appears on the institution's website, or appears in customer records.

AC FN = Acronym of former name

AC TL = Acronym of the transliterated name

AC TR = Acronym of translated name into English

AK = a/k/a (also known as)*

AK FN = also know as, former name

AK TL = A/K/A (also known as) of transliterated name

AK TR = also known as, translation

AQ = Acquisition

CS = Ceased (please select year if known)

DN = Diacritic name

FN = Former name

FN until = Former name with date

FN LB = Former name of Library

LB = Library name

LB TR = translation of library name

MG = Merged (please select year if known)

TL = Transliterated Name

TL FN = Transliterated former name

TR FN = Former name translated to English

*For countries with more than one official language, alternative names in the various languages are listed, if provided on the institution's website.

3.2.4 Alternative URL

Other domains; e.g. former domains, e-mail domains. For ceased institutions, the current URL is moved to the alternative URL field as a former site (FS).

AS = alternative site

ES = Email domain

FS= former site (included current URLs for ceased organisations)

GS = Graduate site (Japan only)

LS = Library site

3.2.5 Ringgold Type

The Ringgold Type classification has been specifically developed to categorize institutions worldwide. The Ringgold Type classification has two parts: sector/subject or type. The first part before the / will give an indication of the sector of the organization:

- academic
- corporate
- government (central government)
- hospital
- other (non-profit institutions)
- public (local government)
- consortium

The part after the / gives an indication of the subject or type of the institution. A full list is available as a separate document and in the taxonomy table in *Identify* exports.

3.2.6 Ringgold Tier

Ringgold Tiers classify by sector, size and type; these are applied globally. A full list is available as a separate document and in the taxonomy table in *Identify* exports.

3.2.7 Ringgold Classification Metadata

Ringgold has expanded the descriptive metadata associated with the records in the Identify Database by developing a new classification system in consultation with customers. The new data will not replace the existing classifications, but will serve to enhance the depth and utility of the *Identify* data and increase value for customers. At present, the new classifications are only applied to academic sector institutions (but not to academic schools, meaning organizations teaching K-12).

The nine new areas of classification are:

- 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
- Ringgold 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

3.2.8 Carnegie classifications

The classification of colleges and universities by the Carnegie Classifications of Institutions of Higher Education at Indiana University to recognize and describe institutional diversity in US. The classification consists of a number of components, Ringgold records the Basic Classification (the traditional Carnegie Classification Framework), and Size.

It should be noted the classification is selective and not all US colleges and universities are included in the scheme. The data provides a point in time snapshot. The data for three updates is available; 2005, 2010, and 2015. (<http://carnegieclassifications.iu.edu/>).

3.2.9 JISC Bands

The JISC Banding Model is the basis for tiered pricing models for content available from publishers and vendors via JISC Collections to many UK institutions. Higher and Further Education institutions, Specialist Colleges, Research Councils and a number of other institutions in the UK, are ranked into JISC Bands. The bands are currently based on the amount of government funding each institution receives, rather than any estimation of research levels. (<https://www.jisc-collections.ac.uk/Help-and-information/JISC-Banding/>)

- Higher Education institutions are listed in the following bands: 1, 2, 3, 4, 5A, 5B, 6, 7, 8, 9, 10
- Further Education, specialist colleges, affiliates and Research Councils are listed in the following bands: A, B, C, D, E, F, G, H, I, J

3.2.10 Size

For academic institutions, this is the number of students. Where possible FTE figures are used.

For all other kinds of institution, size represents the total number of employees.

3.2.11 Staff

For academic institutions, this is number of faculty members or teachers. For all other kinds of institution, it is the number of professionals of predominant interest (e.g. lawyers, faculty, researchers, and teachers). This figure is a subset of size, with the exception of student and faculty figures.

Where possible FTE figures are used.

3.2.12 Beds

The number of beds in a hospital.

3.2.13 Hospitals

The number of hospitals in a hospital group or in a health system.

3.2.14 Doctors

The number of doctors in a hospital organization.

3.2.15 Libraries

The number of libraries within a district or organization.

3.2.16 Schools

The number of schools within a system, district, network, or organization.

3.2.17 Notes

Notes provide additional information about the organization size in the Staff field; it indicates which type of staff the size figure refers to. The Notes field is also used to record the date of closure for a ceased institution in the format: ceased year. In due course these two purposes will be separated into unique fields.

3.2.18 NCES

The National Center for Education Statistics identifier for US education institutions.

(<http://nces.ed.gov/>)

3.2.19 IPEDS

NCES's Integrated Postsecondary Education Data System provides identifiers for US post-secondary education institutions, these are linked to the data it collects in seven areas: institutional characteristics, institutional prices, enrolment, student financial aid, degrees and certificates conferred, student persistence and success, and institutional human and fiscal resources.

(<http://nces.ed.gov/ipeds/>)

3.2.20 ISNI

The International Standard Name Identifier number. ISNI is a bridge identifier which is designed to map between proprietary datasets, making them interoperable. It is an ISO standard (ISO 27729) used for identifying contributors to creative works which are public parties, it includes organizations in the Identify Database as Ringgold is a Registration Agency for ISNI and provides services to organizations interested in using ISNI numbers as bridge identifiers. (<http://www.ringgold.com/isni>)

3.2.21 OFR

The Open Funder Registry provides a common taxonomy of over 11,000 funding body names that funding data initiative participants should use to normalize Funder Names and IDs for deposit.

(<http://www.crossref.org/fundingdata/registry.html>)

3.2.22 Athens

The OpenAthens Identifier. Reviewed monthly.

3.2.23 IP addresses

Ringgold has licensed IP address ranges from Publishing Technology's IngentaConnect service. Ringgold receives updates from the IngentaConnect database on a monthly basis.

3.3 Material Changes

The Ringgold ID is a unique and persistent identification number applied to an organization or part thereof. It is a numeric identifier with no set number of digits and has no meaning other than the identification of the organization to which it is assigned. It is used to disambiguate organizational entities to ensure the correct supply of goods and services and the correct recognition of organizational output. It is utilized by academic publishers and intermediaries and by organizations which provide author affiliation information, including ORCID.

Changes to Ringgold IDs occur in a number of circumstances.

3.3.1 Change in the organization itself

If an institution has merged or is acquired, it frequently no longer exists as a separate entity. Ringgold deletes the institution record, and moves the identifier to the new organization so it can still be used and located.

3.3.2 Duplication/Not valid institution

Occasionally duplicate institution records are discovered. In these cases one of the records is deleted. The Ringgold ID for the deleted record is connected to the retained ID and record.

Sometimes it is discovered that an institution that has been given a Ringgold ID is not a valid institution. The record is deleted and the Ringgold ID is connected to an appropriate alternative organization, for example the entity one step up in the hierarchy.

Ringgold IDs that are connected to deleted records are not in fact deleted, but always redirected to another record. Ringgold IDs are never reused for an entirely separate organization.