

## Data Completeness & Assignment Rules

A key aspect of the ISNI assignment process is the completeness (or otherwise) of the data submitted. In addition, the values or weightings associated with certain data elements is greater than those for others. At an overview level, ISNI distinguishes three broad categories of data completeness: in descending order of effectiveness these are referred to as Rich, Non-sparse and Sparse.

### Rich records

In terms of the types of data element most often submitted, rich records are defined as follows.

#### *For person names:*

The person name is unique if all name variants (surname plus one initial or given name) are unique. Unique names must include at least one full forename or given name.

Person names are considered 'rich' when they meet one of the following sets of conditions:

- Full dates of birth and death
- Year of birth or year of death + (1 title or instrument associated with the person or 1 related name)
- (1 title or instrument) + (1 URL link or 1 related name)
- The request includes an "isNot" statement for an identity with the same name string

*Related name = co-author or affiliated institution*

#### *For organization names:*

Organizations are unique if the requested name does not consist only of initials and the request includes:

- A full UN/LOCODE (<http://www.unece.org/cefact/locode/service/location>)
- and an Organization type
- and Organization URL

Organizations are considered 'rich' when they meet one of the following sets of conditions:

- Both a beginning and an end date
- Beginning year or end date + (1 title or instrument or 1 related name)
- (1 title or instrument) + (1 URL link or 1 related name)
- The request includes an "isNot" statement for an identity with the same name string

*Related name = co-author or affiliated institution*

A rich record may be assigned an ISNI, even if data is supplied by only one source (this is referred to as “single source assignment”).

### **Non-sparse records**

In terms of the types of data element most often submitted, non-sparse records are defined as follows.

Must include a name (for a person or organization)

Must include a local identifier for the name

- Names in the common surnames list must score at least 2 from the following list whereas other names must score at least 1:
  - Birth date (type is “lived”: score 1)
  - Death date (type is “lived”: score 1)
  - Birth date (type is not “lived” AND date is before 1800: score 1)
  - Death date (type is not “lived” AND date is before 1800: score 1)
  - Publisher associated with the name (publisher or publishers: score 1)
  - ISBN of a book associated with the name (ISBN or ISBNs: score 1)
  - Title of a work associated with the name (title or titles: score 2)
  - Co-author or Organization affiliation (one: score 1; two or more: score 2)

Non-sparse records that are admitted to the database but are not rich will result in ISNI assignments if they match with records from other sources.

### **Sparse records**

Sparse records are those that fail to meet the requirements for non-sparse records (see above). Such records cannot be admitted to the ISNI database as without distinguishing metadata, there is no way to relate them to records in external databases.

### **ISNI assignment across many domains**

Please note that the examples given above are largely drawn from ISNI’s involvement in library and publishing workflows. As ISNI is implemented across multiple domains, we already see equivalent but parallel types of data contributed to the ISNI database. Some simple examples include ISRCs (instead of ISBNs) for musical recordings associated with a given name, music labels (instead of publishers) in the music industry and ISANs or EIDR identifiers (instead of ISBNs) to identify films or other audio-visual works.