Usage Manual of the Correspondence Table between the ROR and the NISTEP Dictionary of Names of Universities and Public Organizations

May, 2021

National Institute of Science and Technology Policy (NISTEP)

The Research Organization Registry (ROR) is a registry aiming at identifying research organizations in the world and offering the metadata of those organizations (see https://ror.org/). It started in 2018 as a non-commercial project managed by California Digital Library， Crossref, DataCite， and Digital Science.

The ROR 6.0 (December 2020) registers data for about 98,600 organizations, among which 3,711 are ones in Japan. We checked the correspondence between Japanese organizations in the ROR 6.0 and the NISTEP Dictionary of Names of Universities and Public Organizations ver. 2020.1 (hereafter referred to as "Organization Name Dictionary") and found that 2,575 among 3,711 organizations were corresponded with those in the Organization Name Dictionary.

The resulting correspondence table (file name: CT_ROR6.0_NISTEP_dictionary.xlsx) is explained below.

1. The file structure

The file "CT_ROR6.0_NISTEP_dictionary" consists of two spread sheets, "matched" and "unmatched." The sheet "matched" contains data of 2,575 organizations in the ROR 6.0 connected with those in the Organization Name Dictionary and the sheet "unmatched" contains data of 1,136 other organizations in the ROR 6.0.

Each record (row) of these sheets corresponds to one organization in the ROR 6.0, but there are a few exceptions as explained in the data item "group."

2. Data items in the sheet "match"

Each column corresponds to a data item. Columns A～I are data items on the ROR 6.0, columns J~T are data items on the Organization Name Dictionary, and columns U and V are other data items. Although both the original versions of the ROR 6.0 and the Organization Name Dictionary have many other data items, only main items are shown in this table. In the following explanations, the symbols A, B, C... corresponds to their columns.

【Data items on the ROR 6.0】

(A) seq_ROR: Sequence number given to 3,711 Japanese organizations.　It is based on the order in the original ROR 6.0 file (increasing order of grid_id_preferred (column F)).

(B) id: The organization ID of the ROR. It has a URL format, and the metadata of that organization are mentioned in the URL page.

(C) name: The English name of that organization.

(D) aliases: One or more alias names of that organization are entered. In many cases, it is a Roman

letter expression of the Japanese name.

(E) acronyms: The acronym(s) of that organization.

(F) grid_id_preferred: The preferred organization ID of GRID, a research institution database managed by Digital Science. ROR used GRID as the seed data at its starting time and is being updated simultaneously with GRID.

(G) links: URL of the home page of that organization.

(H) wikipedia_url: URL of Wikipedia on that organization.

(I) types: The type of that organization. One of the following is assigned.

Archive; Company; Education; Facility; Government; Healthcare; Nonprofit; Other

【Data items on the Organization Name Dictionary】

For the data items (K)～(T) below, more detailed explanations are seen in "Usage Manual of the NISTEP Dictionary of Names of Universities and Public Organizations (ver.2020.1_E)" http://doi.org/10.15108/data_rsorg001_2020_1_E

(J) seq_NISTEP: The first 4-digit numeral is the sequence number by the primary organization and the last 2-digit numeral is the sequence number in a primary organization ("01" is the primary organization and "02" and after are subsidiary organizations). The records in the table are arranged by the increasing order of this item.

(K) NID: The organization ID of the Organization Name Dictionary.

(L) formal_name_J: The formal Japanese name of that organization

(M) name_E: The English name of that organization.

(N) name_type: Either "Formal" or "Alias" is given for the type of the name_E in column M. There are a very small number of blanks, for which the English organization name is not known.

(O) sector number: Each organization is classified into one of the following sectors.

1: National universities

2: National junior colleges

3: National technical colleges

4: Public universities

5: Public junior colleges

6: Public technical colleges

7: Inter-university research institutes

8: National organizations

9: National R&D agencies and other agencies

10: Organizations of local governments

12: Private universities

13: Private junior colleges

14: Private technical colleges

15: Companies

16: Non-profit organizations

17: Other organizations

(P) layer: "Primary" if the organization is an independent organization and "subsidiary" if the organization is a subsidiary body of a primary organization.

(Q) primary org: In case that the column P value is "subsidiary", NID of its primary organization is entered in this column.

(R) status: If the organization is defunct (i.e. not-existing), "no" is entered.

(S) transfer year: The year in which the organization became defunct, in case that the column R value is "no."

(T) succeeding org: In case that the column R value is "no", the NID of the succeeding organization is entered if it exists.

【Other data items】

(U) group: Two or more organizations registered in the Organization Name Dictionary are related by a same group number in the following cases.

    (a) Two or more organizations registered in the ROR are connected to a single organization registered in the Organization Name Dictionary. (Group numbers 1～9)

    (b) A single organization registered in the ROR is connected to two or more organizations registered in the Organization Name Dictionary. (Group number 10)

    (c) There are some cases that the organization registered in the ROR is a subsidiary body of the (primary) organization registered in the Organization Name Dictionary but the primary organization is not registered in the ROR, In these cases, the primary organization is added to the table (no ROR data in this row) and a same group number is given to the records of those organizations. (Group numbers 11～47)

(V) remark: Remarks are given in the following cases.

    (a) Link to the URL shown in "links" (column G) has been broken or are linked to a different or a preceding organization. Or, the column G is blank. The correct link is shown if it is identified.

    (b) The organization name entered in "name" (and also "aliases") is that of a different or a preceding organization.

    (c) The organization in the ROR is a subsidiary body of that in the Organization Name Dictionary, but they are connected because the organization in the ROR is not registered in the Organization Name Dictionary.

    (d) There are some other remarks.

3. Data items in the sheet "unmatch"

Data items on the ROR 6.0 (columns A～I) are same as those in the sheet "match."

There are a very few cases that the Japanese name of the organization and a related remark are entered in "formal_name_J" (column J) and in "remark" (column K), respectively.

4. Some statistical data

4.1 Sector distribution of the connected ROR organizations

Sector	Connected ROR organizations			All organizations in Organization Name Dictionary		
	Primary	Subsidiary	Total	Primary	Subsidiary	Total
1. National universities	85	49	134	101	2,042	2,143
2. National junior colleges	1		1	26		26
3. National technical colleges	51		51	59		59
4. Public universities	89	8	97	110	95	205
5. Public junior colleges	9		9	63		63
6. Public technical colleges	2		2	7		7
7. Inter-university research institutes	5	16	21	7	26	33
8. National organizations	26	9	35	124	80	204
9. National R&D agencies and other agencies	60	184	244	167	561	728
10. Organizations of local governments	269	26	295	1,043	409	1,452
12. Private universities	594	52	646	686	619	1,305
13. Private junior colleges	144		144	583		583
14. Private technical colleges	2		2	4		4
15. Companies	408	10	418	4,924	18	4,942
16. Non-profit organizations	438	33	471	7,233	137	7,370
17. Other organizations	5		5	8	2	10
Total	2,188	387	2,578	15,835	3,989	19,824

4.2 The relationship between "types" in the ROR and "sector" in the Organization Name Dictionary

Sector	Types of ROR								
	Archive	Company	Education	Facility	Government	Healthcare	Non-profit	Other	Total
1. National universities	2		87	1		44			134
2. National junior colleges			1						1
3. National technical colleges			51						51
4. Public universities			89			8			97
5. Public junior colleges			9						9
6. Public technical colleges			2						2
7. Inter-university research institutes	3		2	14				2	21
8. National organizations			5	7	16	4		3	35
9. National R&D agencies and other agencies	11	2	11	43	18	146	4	8	243
10. Organizations of local governments	16	1	15	48	37	167	1	10	295
12. Private universities	1		592			52		1	646
13. Private junior colleges			144						144
14. Private technical colleges			2						2
15. Companies	1	396		10		12			419
16. Non-profit organizations	10	3	4	82	2	184	74	112	471
17. Other organizations			1	4					5
Total	44	402	1,015	209	73	617	79	136	2,575