Usage Manual of the NISTEP Dictionary of Names of Universities and Public Organizations (ver.2020.1_E)

1. Aims

As a part of the "Science of Science, Technology and Innovation Policy" program by the Ministry of Education, Culture, Sports, Science and Technology, NISTEP promotes development of data infrastructure on R&D activities in universities and public organizations. It aims at developing basic data infrastructure used for grasping and analyzing the actual state of R&D in Japan as well as evaluating its performance at the organization level and at the sector level. Development of such data infrastructure essentially requires accurate identification of Japanese research institutions that are seen in various data sources (e.g., scientific publication databases, patent databases, R&D statistics). The NISTEP Dictionary of Names of Universities and Public Organizations (hereinafter referred to as "Organization Name Dictionary") contains basic information of such research institutions.

The Japanese version of the Organization Name Dictionary was publicized in December 2012 at first and thereafter has been revised about once a year. This version (ver.2020.1_E) is the first English version derived from the Japanese version "The NISTEP Dictionary of Names of Universities and Public Organizations ver.2020.1." The Japanese version is available from http://doi.org/10.15108/data_rsorg001_2020_1

2. Organizations included in the Organization Name Dictionary

2.1. Organizations recorded

The target of the Organization Name Dictionary is organizations in Japan that are engaged in R&D. Although the main subjects are higher education institutions (universities, inter-university research institutes, junior colleges, and technical colleges) and public research institutions (e.g., national organizations, national research and development institutes), organizations of local governments, private companies, non-profit organizations and other organizations are also included as much as possible.

2.2. Primary organization and subsidiary organization

Independent organizations are called "primary organizations", and subsidiary bodies etc. of primary organizations are called "subsidiary organizations". For about 50 primary organizations, their subsidiary organizations directly connecting to the primary organizations (secondary subsidiary organizations) are comprehensively recorded in the Organization Name Dictionary (see 2.4(2) for details). For other primary organizations, their subsidiary

organizations are recorded as necessary.

2.3. Defunct organizations

Organizations that became no longer existent after 2008 due to abolition, unification, or renaming are also recorded in the Organization Name Dictionary. Succeeding organizations of such defunct organizations, if any, are linked to their preceding organizations.

2.4. Recording criteria of the Organization Name Dictionary

(1) Primary organizations

 a) All existing (active) universities, junior colleges, and technical colleges in Japan

 b) Existing organizations engaged in R&D in Japan, including national organizations, national R&D agencies, independent administrative institutions, organizations of local governments, private companies, non-profit organizations, and other organizations (i.e., international organizations located in Japan)

 c) Main organizations that have been disclosed since 2008

(2) Subsidiary organizations

 a) Secondary subsidiary organizations of the following 33 universities: Hokkaido University; Tohoku University; University of Tsukuba; Gunma University; Chiba University; The University of Tokyo; Tokyo Medical and Dental University; Tokyo Institute of Technology; Tokyo University of Agriculture and Technology; Niigata University; University of Toyama; Kanazawa University; University of Fukui; Shinshu University; Gifu University; Nagoya University; Kyoto University; Osaka University; Kobe University; Okayama University; Hiroshima University; Tokushima University; Kyushu University; Nagasaki University; Kumamoto University; Osaka Prefecture University; Osaka City University; Keio University; Waseda University; Tokai University; Tokyo University of Science; Nihon University; Kindai University

 b) Secondary subsidiary organizations of the following 4 inter-university research institutes: High Energy Accelerator Research Organization; National Institutes of Natural Sciences; Research Organization of Information and Systems; National Institutes for the Humanities

 c) Secondary subsidiary organizations of the following 11 national R&D agencies: Japan Aerospace Exploration Agency; Japan Science and Technology Agency; RIKEN; National Cancer Center; National Center for Global Health and Medicine; National Cerebral and Cardiovascular Center; National Center for Child Health and Development; National Center of Neurology and Psychiatry; National Center for Geriatrics and Gerontology; Japan Fisheries Research and Education Agency; National Agriculture and Food Research Organization

 d) Secondary subsidiary organizations of the following 2 independent administrative institutions: National Hospital Organization; National Institutes for Cultural Heritage

 e) Following divisions belonging to universities (other than those included in (a))

(i) Hospitals

(ii) Research institutes of national universities

(iii) Divisions designated as "Joint Usage / Research Center" or "World Premier International Research Center Initiative (WPI)"

(iv) Some tertiary subsidiary organization under the secondary subsidiary organization in (a), only if the secondary subsidiary organization is very broad

f) Hospitals, colleges, and principal research institutes that are subsidiary divisions of national organizations or national agencies (other than those included in (c) and (d))

2.5. Number of organizations recorded

In this version of the Organization Name Dictionary (ver.2020.1_E), more than 15,000 primary organizations and nearly 4,000 subsidiary organizations are recorded (see Tables 1 for details).

3. Information covered (Explanations of the data items)

3.1. Organization identification information

a) NISTEP organization ID: 18-character alphanumeric organization identification code uniquely used by NISTEP. The code starts with "NID2012", followed by a randomly generated number (11 characters).

b) Record number: This number is given to each record (organization) for arrangement, since the NISTEP Organization ID itself has no meaning. The record numbers have a form "WW-XXXX", "WW-XXXX-YYY", or "WW-XXXX-YYY-ZZ", for the primary organizations, the secondary subsidiary organizations, and the tertiary/quaternary subsidiary organizations, respectively. Here, "WW" is the sector number, "XXXX" is the sequence number of the primary organization in the sector, "YYY" is the sequence number of the secondary subsidiary organization in the primary organization, and "ZZ" is the sequence number identifying the tertiary/quaternary subsidiary organization under the secondary organization.

3.2. Information on organization name

a) Organization name: Within 19,134 organizations in the Organization Name Dictionary, 18,215 (95.2%) are given an English name. English names of the remaining 919 organizations are not well-defined at the moment.

b) Name type: It discriminate whether the organization name in (a) is the formal name (then "1") or the alias or abbreviated name (then "2"). When it cannot be confirmed whether the name is formal or alias, "2" is entered. At present, 10,161 organizations are given "1" and 8,054 given "2".

c) Organization name in Japanese: The formal name of each organization in Japanese. It is

given each of all organizations in this dictionary.

3.3. Information on the sector to which organizations belong

    a)   Sector number and sector class: Each organization is classified into one of 16 sectors (see Table 1).

    b)   Hospital flag: The flag "○" is assigned if the relevant organization is a hospital. This flag is independent of the sector class.

3.4. Information on the layer of organization

    a)   Layer: "1" if a primary organization, "2" if a secondary subsidiary organization,.and "3" or "4" if subsidiary organization belonging to a secondary or tertiary subsidiary organization, respectively.

    b)   NISTEP Organization ID of primary organization (Primary organization ID): ID of the primary organization if the relevant organization is a subsidiary organization. ID of the organization itself if the relevant organization is a primary organization.

    c)   NISTEP Organization ID of direct upper organization (Direct upper organization ID): Given if the relevant organization is a tertiary and quaternary subsidiary organization

    d)   University subsidiary type: One of the following types is assigned to subsidiary organizations of national, public, and private universities: Undergraduate school; Graduate school; Special course; United undergraduate and graduate school; Faculty; Research institute; Hospital; Cross-sectional division.

3.5. Information on transfer of organization

    a)   Status: "no" for defunct (i.e. not-existing) organizations and blank for active (i.e. existing) organizations.

  If a) is "no", b), c), and d) below are given.

    b)   Transfer type: "Unified", "Abolished", or "Renamed".

    c)   Transfer date: Date of transfer in the form "YYYY-MM-DD". "YYYY-MM-00" if only the year and the month is known and "YYYY-00-00" if only the year is known.

    d)   NISTEP Organization ID of succeeding organization (Succeeding organization ID): Given if the Transfer type is "Unified" or "Renamed", or if there is an organization that succeeded the activities of the relevant organization after "Abolished".

3.6. Other information

    a)   Out-of-criteria organization flag: Although the recording criteria for subsidiary organizations are defined as 2.4(2), the Organization Name Dictionary includes exceptionally some subsidiary organizations out of the criteria. A flag "※" is entered in the column of such organizations. Many of such organizations are recorded in order to avoid a failure that may occur when this Dictionary is used to identify the organization names mentioned in various databases.

4. Data sources

The names and transfer information of organizations and the relation between primary and subsidiary organizations were identified mainly through investigations on the websites of the relevant organizations, organization lists on the websites of governmental agencies etc., which were supplemented, if necessary, by information on Wikipedia and other information sources. Systematic investigations have been carried out at least once a year for higher education institutions (universities, inter-university research institutes, junior colleges, and technical colleges) and public research institutions (e.g., national organizations and national R&D and other agencies). For other organizations (organizations of local governments, companies, and non-profit organizations), systematic investigations have been done once a few years.

5. File structure

The original file of the Organization Name Dictionary is in a relational database format and consists of 6 sub-files.

This Dictionary file for public view is integrated into a single Excel table format from the original sub-files, considering usability. Each record (row) corresponds to one organization ("NISTEP Organization ID" becomes a record key).

6. On the usage of the Organization Name Dictionary

The Creative Commons "Attribution + ShareAlike" (BY-SA) license is applied for the usage of the Organization Name Dictionary (ver.2020.1_E). Namely, you can remix, adapt, and build upon this work even for commercial purposes, as long as you credit us and license your new creations under the identical terms. The credit should be expressed as follows:

> Creator: National Institute of Science and Technology Policy
> Title of the work: The NISTEP Dictionary of Names of Universities and Public Organizations (Ver.2020.1_E)
> DOI： http://doi.org/10.15108/data_rsorg001_2020_1_E

7. Related databases

The following data (in Japanese) are available from:
https://www.nistep.go.jp/research/scisip/randd-on-university

(1) NISTEP Dictionary of Names of Universities and Public Organizations (ver.2020.1)
http://doi.org/10.15108/data_rsorg001_2020_1
This is the Japanese version of this Dictionary.

(2) Name Variation Table for Universities and Public Organizations (ver.2019.1)

http://doi.org/10.15108/data_rsorg002_2019_1

This table lists various name variations for the organizations registered in the Organization Name Dictionary. The name variations includes ones from the following data sources.

    a) The formal names, aliases, and abbreviated names recorded in the Organization Name Dictionary,

    b) The organization name variations used in Web of Science Core Collection, (WoSCC) and

    c) The organization name variations used in Scopus.

(3) Concordance Table Connecting Scopus to NISTEP Organization Name Dictionary (ver.2018.1.1)

http://doi.org/10.15108/data_rsorg004_2018_1

This is a concordance table that connects the article ID of Scopus (Eid of the Scopus) and the NISTEP Organization ID of the Organization Name Dictionary. This table is based on the organization-name identification for organizations in Japan that were found in the author-affiliation data in Scopus since 1996.

(4) Concordance Table Connecting WoSCC to NISTEP Organization Name Dictionary (ver.2020.1)

http://doi.org/10.15108/data_rsorg003_2020_1

This is a concordance table that connects the accession number (ut) of Web of Science and the NISTEP Organization ID of the Organization Name Dictionary. This table is based on the organization-name identification for organizations in Japan that were found in the author-affiliation data in WoSCC since 1998.

Table 1 Number of Organizations and Number of Names Recorded in the Current Published Version (ver.2020.1_E)

	Sector	Primary Organizations			Subsidiary Organizations		
		Active	Defunct	Total	Active	Defunct	Total
1	National universities	86	15	101	1,509	533	2,042
2	National junior colleges		26	26			
3	National colleges of technology	51	8	59			
4	Public universities	93	17	110	82	13	95
5	Public junior colleges	16	47	63			
6	Public technical colleges	3	4	7			
7	Inter-university research institutes	4	3	7	25	1	26
8	National organizations	54	70	124	65	15	80
9	National R&D agencies and other agencies	80	87	167	388	173	561
10	Organizations of local governments	791	252	1,043	312	97	409
12	Private universities	614	72	686	522	97	619
13	Private junior colleges	310	273	583			
14	Private colleges of technology	3	1	4			
15	Companies	4,169	755	4,924	14	4	18
16	Non-profit organizations	3,572	3,661	7,233	89	48	137
17	Other organizations	6	2	8	2		2
	Total	9,852	5,293	15,145	3,008	981	3,989