Launching the Operation of
"Japan Data Catalog for the Humanities and Social Sciences" (JDCat)
–Support for Searching Research Data
for research & education, policy planning –

The Japan Society for the Promotion of Science (JSPS, President: SATOMI Susumu) and the National Institute of Informatics (NII, Director General: KITSUREGAWA Masaru), have announced the public release of "Japan Data Catalog for the Humanities and Social Sciences" (JDCat), a system for searching research data in the humanities and social sciences. JDCat is a product of JSPS's Program for Constructing Data Infrastructure for the Humanities and Social Sciences [*1]. It was created based on the repository software WEKO3 developed by the NII's Research Center for Open Science Infrastructure [*2] (RCOS, Center Director: YAMAJI Kazutsuna, Professor, NII Digital Content and Media Sciences Research Division), and it enables cross-disciplinary searches of data in a variety of humanities and social sciences publications by affiliated research institutions and that provides access to the published data of each institution.

Japan Data Catalog for the Humanities and Social Sciences (JDCat)
https://jdcat.jsps.go.jp

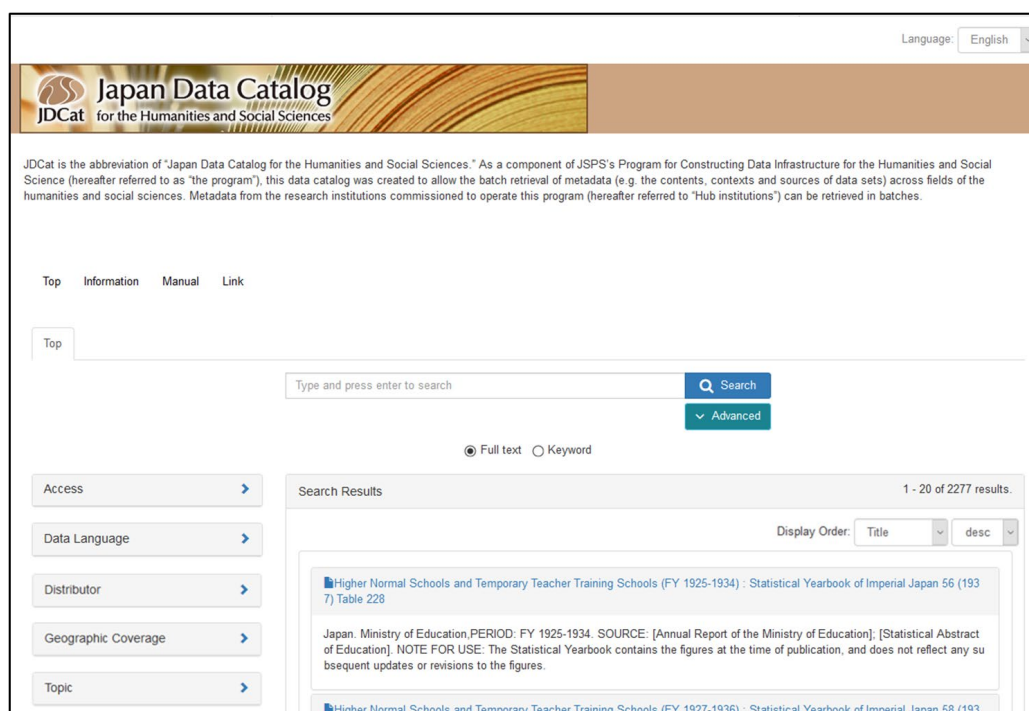Scheduled to be released at 10:00 on Thursday, July 15 2021, JST

Figure 1: JDCat provides cross-disciplinary data-search of research data in the diverse fields of the humanities and social sciences, allowing access to data published by 5 research institutions.

Japan Society for the Promotion of Science
University-Industry Cooperation and Research Program Division
Kojimachi Business Center,
5-3-1 Kojimachi, Chiyoda-ku, Tokyo
102-0083 JAPAN
Direct: +81(0)3-3263-4645   FAX : +81(0)3-3263-1716
E-Mail: di-hs@jsps.go.jp

Research Organization of Information and Systems
National Institute of Informatics
Publicity Team
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo
101-8430 JAPAN
Direct: +81(0)3-4212-2164   FAX : +81(0)3-4212-2150
E-Mail: media@nii.ac.jp

Research data in fields of the humanities and social sciences are a means of capturing human activities and societal phenomena, and are used in various forms of societal decision-making, including policy formulation based on objective evidence. These data include individual data from social surveys, statistical tables from official statistics, texts of historical materials, image data, and many other types of data. In other countries, infrastructures are being actively developed to make data in the humanities and social sciences available to the public and share them with societies. JDCat is a cross-disciplinary search system for research data maintained at present by the five research institutions selected by JSPS via an open call for proposals. These are the Institute of Economic Research, Hitotsubashi University; Center for Social Research and Data Archives, Institute of Social Science, The University of Tokyo; Panel Data Research Center at Keio University; JGSS Research Center at Osaka University of Commerce; and Historiographical Institute, The University of Tokyo. At the time of this release, the JDCat system targets research data in the social sciences, with plans to add research data in the humanities around October this year.

The information used to describe the research data (metadata) handled by JDCat conforms to the Data Documentation Initiative [*3], an international standard in the social sciences. The system is designed to enable searches in both Japanese and English. In addition to general text searches, it supports advanced and faceted searches [*4]. JDCat provides an intuitive and easy-to-use search environment for a wide range of users who are interested in research data in the humanities and social sciences. It is expected that the research data provided by this system will be utilized in various fields such as collaborative research, education, and policy planning.

Comments from HIROMATSU Takeshi, Director of JSPS's Center for Constructing Data Infrastructure for the Humanities and Social Sciences:

"JDCat is a data catalogue created by JSPS as a part of the Program for Constructing Data Infrastructure for the Humanities and Social Sciences. It gives open access to everyone. During past three years, preparatory works were carried out that laid the foundations for this launching of the JDCat system. They started with surveying the state of data archives in other countries and selecting and translating the controlled vocabulary in their metadata schema. Then, we decided what functions should be installed in JDCat, test-operated the system, and made refinements to it.

JDCat has four main operational objectives based on FAIR Date Principles [*5]. These are (1) making it easy to identify data, (2) optimizing interoperability both in and outside Japan, (3) making possible access to Japan data in English, and (4) creating a data infrastructure that facilitates the automatic collection metadata from foreign countries. Technologically, these objectives are realized by (1) assigning digital object identifiers to data, (2) mapping the schema in the Data Documentation Initiative (DDI) and Japan Consortium for Open Access Repository (JPCOAR), (3) maintaining metadata in both Japanese and English, and (4) providing open access to metadata through "Creative Commons licenses zero" (CC0) [*6].

At the present stage, the information provided by JDCat covers only the above-described five research institutions, and its scope and types are limited. Both of data scope and types are expected to expand as the usership of the JDCat will increases. As one of staff members involved in this Program, I hope that "what is starting out as something small will be nurtured into something big", and it will have big impacts."

(*1) Program for Constructing Data Infrastructure for the Humanities and Social Sciences:

This program was established by JSPS in April 2018 with an aim of promoting the humanities and social sciences. By building a comprehensive infrastructure for sharing and utilizing data related to humanities and social science research across disciplines and countries, the project aims to encourage researchers to share data together and promote joint research, both domestically and internationally.　See https://www.jsps.go.jp/english/e-di/index.html for details

(*2) Research Center for Open Science and Data Platform:

In response to the global momentum toward open science, this research center was established within NII in April 2017 to develop and operate an academic platform to serve as an infrastructure for open science. It is expected that, through the wide sharing of academic papers and research data in academia and society, and a wide range of research activities being carried out openly, research activities will be accelerated, problem solving based on close cooperation with society will be promoted, and that this will push academic activities to a new dimension (open science). See https://rcos.nii.ac.jp/en/ for details

(*3) Data Documentation Initiative:

The Data Documentation Initiative is an international metadata standard for describing data obtained through observation such as surveys and experiments in the fields of social, behavioral, economic, and health sciences. It has been adopted by data archives in other countries, such as ICPSR and UKDS, and is being used in the operation of search systems. See https://ddialliance.org/ for details

(*4) faceted searches

A navigation system in which the search system prepares search conditions in advance and the user can narrow down the search target by selecting the search conditions.

JDCat adopts faceted search for some items such as "distributor", aiming to create an environment where first-time researchers who do not know specialized keywords can easily search for data.

(*5) FAIR Date Principles (FAIR)

These are principles that create an international standard for the data creation and sharing. FAIR is an abbreviation of Findable, Accessible, Interoperable, and Reusable, which gives expression to an appropriate method of providing public access to data. International momentum is gaining for creating standardized data based on these principles. For more information on FAIR as a standard for sharing data, see

https://DOI.org/10.18908/a.2018041901

(*6) Creative Commons licenses zero (CC0)

CC0 is a type of copyright licenses that have been released by Creative Commons in order to allow for the free distribution and reuse of copyrighted material under certain conditions, while retaining copyright. CC0 indicates 'no rights reserved' and that anyone can use the work in any way. For more information, see the Creative Commons Japan website: https://creativecommons.jp/licenses/

National Institute of Informatics