



# Mejorando la FAIRness (“Justicia”) de la literatura gris arqueológica japonesa

*Improving the FAIRness of Japanese Archaeological Grey Literature*

Yuichi Takata<sup>1</sup> y Peter Yanase<sup>2</sup>

<sup>1</sup>Nara National Research Institute for Cultural Properties. Nijō-chō 2-9-1, Nara, Nara pref., Japan. ORCID: <https://orcid.org/0000-0003-3330-3983>

<sup>2</sup>Nara National Research Institute for Cultural Properties. ORCID: <https://orcid.org/0000-0002-7507-7200>

Traducción: Andrés D. Izeta y G. Roxana Cattáneo

## Resumen

*En Japón, la arqueología es predominantemente una actividad administrativa, con solo 15,261 excavaciones impulsadas por la investigación en comparación con 331,574 excavaciones de rescate vinculadas a proyectos de desarrollo entre 1976 y 2022. Este enfoque opera dentro de un marco más amplio centrado en la protección de sitios arqueológicos y como parte de un programa de preservación de propiedades culturales más grande regido por la Ley para la Protección de Bienes Culturales. La ley enfatiza tanto la preservación como la utilización de bienes culturales, siendo la protección una actividad que abarca ambas. Las tareas administrativas, que incluyen excavación, análisis, selección, reparación, divulgación pública, exhibición en museos y preservación, constituyen un proceso cíclico de gestión de propiedades culturales. La estructura administrativa descentralizada implica la Agencia de Asuntos Culturales a nivel nacional y las Juntas Educativas locales a niveles prefecturales y municipales. Cabe destacar que el 95% de las excavaciones en los últimos cincuenta años fueron realizadas por agencias administrativas como parte de la preservación de bienes culturales. Los informes de trabajo de campo se consideran los productos finales de las excavaciones, fundamentales para la difusión de resultados y la reclamación de propiedad común. Sin embargo, debido a la descentralización, la producción y difusión de la literatura gris en la arqueología japonesa carecen de una gestión coherente. Iniciativas como la Base de Datos Integral de Informes de Sitios Arqueológicos en Japón (SORAN) buscan mejorar la Encontrabilidad, Accesibilidad, Interoperabilidad y Reutilización (FAIRness) de la literatura gris arqueológica japonesa.*

**Palabras clave:** Excavaciones de rescate; Literatura gris; Informes de trabajo de campo; Principio FAIR; Japón.

## Abstract

*In Japan, archaeology is predominantly an administrative pursuit, with only 15,261 research-driven excavations compared to 331,574 rescue excavations linked to development projects from 1976 to 2022. This approach operates within a broader framework focused on protecting archaeological sites, part of a larger cultural property preservation program governed by the Law for the Protection of Cultural Properties. The law emphasizes both the preservation and utilization of cultural properties, with protection encompassing these activities. Administrative tasks, including excavation, analysis, selection, repair, public sharing, museum display, and preservation, constitute a cyclical process of cultural property management. The decentralized administrative structure involves the national-level Agency for Cultural Affairs and local boards of education at prefectural and municipal levels. Notably, 95% of excavations over the past fifty years were conducted by administrative agencies as part of cultural property preservation. Fieldwork reports are considered the end-products of excavations, crucial for disseminating results and claiming common ownership. However, due to decentralization, the production and dissemination of grey literature in Japanese archaeology lack coherent management. Initiatives like the Comprehensive Database of Archaeological Site Reports in Japan (SORAN) aim to enhance the Findability, Accessibility, Interoperability, and Reusability (FAIRness) of Japanese archaeological grey literature.*

**Keywords:** Rescue excavations; Grey literature; Fieldwork reports; FAIR principle; Japan.

## Introducción

La arqueología en Japón es principalmente un esfuerzo administrativo. Entre 1976 y 2022, solo hubo 15,261

excavaciones arqueológicas impulsadas por preguntas de investigación en comparación con las 331,574 excavaciones de rescate realizadas antes de proyectos de desarrollo (Maizō bunkazai kankei tōkei shiryō, 2023).

Recibido 27-08-2023. Recibido con correcciones 26-09-2023. Aceptado 24-11-2023

Revista del Museo de Antropología 16 (3): 485-492 /2023 / ISSN 1852-060X (impreso) / ISSN 1852-4826 (electrónico)  
<http://revistas.unc.edu.ar/index.php/antropologia/index>

IDACOR-CONICET / Facultad de Filosofía y Humanidades – Universidad Nacional de Córdoba - Argentina



Las excavaciones de rescate en Japón se llevan a cabo dentro de un marco más amplio preocupado por la protección de sitios arqueológicos, que a su vez es parte de un programa más grande para la preservación de bienes culturales. Las actividades de preservación dentro de este marco se basan en la Ley para la Protección de Bienes Culturales (promulgada en 1950, última enmienda en 2022) (Tanaka, 1984) y se centran en objetos (Okamura, 2011).

La ley establece que es importante “preservar” (hozon), así como “utilizar” (katsuyō) “bienes culturales” (bunkazai), y el concepto de “protección” (hogo) incluye ambas actividades. (Para una traducción al inglés relativamente reciente y completa de la Ley, consulte Japón: Ley para la Protección de Bienes Culturales, 2009). El proceso real de protección implica tareas como excavar, recopilar, analizar objetos, seleccionar objetos para designar como bienes culturales, reparar y restaurar objetos, compartir conocimientos adquiridos sobre los objetos con el público, exhibir objetos en museos y preservar objetos. Todas estas actividades forman un ciclo de trabajo administrativo (Nishiyama, 2015).

La maquinaria administrativa que se ocupa de la protección de bienes culturales está descentralizada. A nivel nacional, está coordinada por la Agencia de Asuntos Culturales, adscrita al Ministerio de Educación, Cultura, Deportes, Ciencia y Tecnología; a niveles prefecturales y municipales, la gestionan las Juntas de Educación locales (Tanaka, 1984).

En otras palabras, el 95% de las excavaciones realizadas en los últimos cincuenta años en Japón fueron llevadas a cabo por agencias administrativas como parte de sus actividades de preservación de bienes culturales.

Aunque ninguna legislación lo afirma explícitamente, en este marco, los informes finales de trabajo de campo se consideran los productos finales de las excavaciones arqueológicas. Según la Agencia de Asuntos Culturales, los informes de trabajo de campo son “los más centrales entre los registros dejados para las futuras generaciones como reemplazos de bienes culturales enterrados que no se pudieron preservar en su estado original, y permiten la difusión y utilización de los resultados de las excavaciones” (Gyōsei mokuteki de okonau maizō bunkazai no chōsa ni tsuite no hyōjun, 2004: 5). Además, la Agencia afirma en otro lugar que “debido a que los informes de trabajo de campo son registros de bienes culturales enterrados que resumen con precisión los resultados de las excavaciones, es necesario tomar medidas para preservarlos perpetuamente y hacerlos públicos para que las personas puedan reclamar la propiedad común y utilizarlos” (Bunkachō Bunkazai Bu Kinenbutsu Ka & Nara Bunkazai Kenkyūjo, 2010).

Sin embargo, debido a la naturaleza descentralizada de

la maquinaria administrativa japonesa, la producción y difusión de la literatura gris no se gestiona de manera coherente. Como resultado, ni la extensión de la literatura gris arqueológica ni la medida en la que se ha publicado son fácilmente identificables en Japón.

Se han realizado varios intentos en el pasado para resolver este problema, siendo el más reciente el desarrollo de la Base de Datos Integral de Informes de Sitios Arqueológicos en Japón (SORAN), un índice en línea de informes de trabajo de campo, actualmente gestionado por el Instituto Nacional de Investigación de Bienes Culturales de Nara (NABUNKEN).

En este documento, presentaremos brevemente SORAN y los esfuerzos relacionados para mejorar la “Encontrabilidad”, Accesibilidad, Interoperabilidad y Reutilización (FAIRness) de la literatura gris arqueológica japonesa.

## El nacimiento de SORAN

SORAN es actualmente gestionado por NABUNKEN. NABUNKEN fue establecido en 1952 como una organización auxiliar bajo la Comisión Nacional para la Protección de Bienes Culturales (posteriormente la Agencia de Asuntos Culturales) para el estudio in situ de bienes culturales muebles e inmuebles en la región de Nara. Es miembro de los Institutos Nacionales de Patrimonio Cultural, una entidad paraguas formada en 2007 que comprende nueve entidades legales especializadas en asuntos relacionados con el patrimonio cultural. NABUNKEN se encarga de liderar y educar a la comunidad japonesa de expertos en patrimonio cultural, especialmente a la comunidad arqueológica. Desde su creación el instituto en sí, es un productor activo de información arqueológica, ya que ha participado en innumerables excavaciones en Japón y en el extranjero. También gestiona tres museos que exhiben sus hallazgos y se esfuerza constantemente por mejorar los métodos actuales de preservación, conservación, restauración, presentación y mantenimiento de bienes culturales (Nara Bunkazai Kenkyūjo, 2022).

Por otro lado, SORAN surgió como respuesta a la necesidad de mejorar la localización y accesibilidad de la literatura gris arqueológica japonesa y la información contenida en ella. En 2008, cinco bibliotecas universitarias nacionales ubicadas en el extremo occidental de Japón, con la Biblioteca de la Universidad de Shimane a la cabeza, formaron una alianza para publicar los textos completos de informes locales de trabajo de campo arqueológico en Internet. Este esfuerzo se convirtió eventualmente en un proyecto a nivel nacional con la participación de veintiuna bibliotecas universitarias nacionales llamado el Proyecto de Repositorios de Información de Sitios Arqueológicos a Nivel Nacional. Durante la duración de este proyecto, se digitalizaron casi 20,000 informes y se

pusieron a disposición gratuitamente en línea. Aunque el proyecto fue un esfuerzo conjunto, los conjuntos de datos de las prefecturas se gestionaron por separado. No fue hasta que NABUNKEN asumió las responsabilidades de gestión en junio de 2015 que los datos acumulados se fusionaron en la base de datos monolítica SORAN (Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni tsuite 2, 2017).

### El contenido de SORAN y servicios relacionados

En los últimos años, SORAN ha crecido considerablemente con la ayuda de un número creciente de proveedores de datos de todo Japón. Contiene más datos y más tipos de datos cada año, y nuevos servicios se implementan gradualmente en el portal. A partir del 26 de julio de 2023, contiene la siguiente información pública:

- 126,455 piezas de información bibliográfica sobre literatura gris relacionada con bienes culturales elaborada por administraciones locales y museos.
- 143,857 conjuntos de metadatos basados en resúmenes de informes de trabajo de campo.
- 118,105 piezas de información bibliográfica sobre artículos académicos publicados dentro de informes de trabajo de campo y otras revistas no arbitradas y oscuras.
- 665,621 piezas de ubicaciones geográficas relacionadas con bienes culturales.
- 35 modelos tridimensionales de restos estructurales.

### Políticas sobre datos digitales

El principio FAIR no es ampliamente conocido ni aplicado en Japón, especialmente en las ciencias sociales, y no existen repositorios públicos donde los investigadores puedan cargar datos o preimpresiones de manera gratuita. La Sociedad Japonesa para la Promoción de la Ciencia, la única agencia independiente de financiamiento en Japón, gestiona un portal llamado Catálogo de Datos de Japón para las Humanidades y Ciencias Sociales (JDCat), que es un sitio de agregación de metadatos donde se puede analizar y descargar metadatos a granel. Sin embargo, el portal aún está en sus inicios: hasta julio de 2023, solo cuenta con la participación de cinco instituciones. Se cree que esta situación mejorará gradualmente a partir de este año, ya que JDCat acaba de comenzar a buscar nuevos socios. Es importante destacar que el proyecto solo acepta solicitudes de bases de datos con una trayectoria bajo los principios FAIR.

Por otro lado, la Agencia de Asuntos Culturales compartió su política con respecto a los datos arqueológicos digitales en tres volúmenes entre 2017 y 2019 (Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni tsuite 1–3, 2017–2019). El primer volumen trata sobre la digitalización de fotografías y la fotografía digital, el segundo sobre la digitalización de informes de trabajo de campo, y el tercero sobre la digitalización de fuentes

primarias y el manejo adecuado de fuentes primarias digitales como mediciones realizadas con dispositivos digitales. Estas pautas promueven la digitalización de materiales físicos con fines de respaldo y para mejorar la accesibilidad. Al mismo tiempo, también destacan la importancia de la gestión adecuada de archivos digitales nativos. Sin embargo, estas son solo pautas, y depende de las administraciones locales seguirlas o no. ¿Qué tan justo es SORAN?

NABUNKEN ha aprendido mucho sobre el principio FAIR y otras mejores prácticas de gestión de datos a través de varios proyectos internacionales en los que participó, como ARIADNE (Infraestructura de Investigación Arqueológica para la Red de Datos Arqueológicos en Europa) y SEADDA (Salvando la Arqueología Europea de la Era Digital Oscura) y lanzó una traducción al japonés de las Guías de Buenas Prácticas de ADS (<https://archaeologydataservice.ac.uk/help-guidance/guides-to-good-practice/>) en 2022. Sin embargo, aplicar todo ese conocimiento a SORAN llevará tiempo. A continuación, evaluaremos brevemente qué tan FAIR es SORAN hasta 2023 basándonos en los principios rectores disponibles en el sitio web de GO FAIR (<https://www.go-fair.org/>).

La información bibliográfica sobre informes de trabajo de campo es fundamental para el conjunto de datos de SORAN. Cada pieza de información bibliográfica tiene una URL única que puede ser utilizada para referirse a ella. Los datos bibliográficos en sí no son muy útiles cuando se trata de informes de trabajo de campo arqueológicos porque no transmiten mucho más allá de los nombres y tipo de sitio. Para resolver este problema, la Agencia de Asuntos Culturales comenzó a solicitar a los gobiernos locales que adjuntaran información a través de formularios semiestructurados a los informes desde 1994 (Morimoto, 2017). Estos formularios deben contener información sobre cada intervención arqueológica sobre los informes de trabajos de campo y registrar los nombres, direcciones, coordenadas, tamaños, tipos y edades de los sitios excavados, las fechas y los motivos de las excavaciones, así como listas de los tipos de restos estructurales y artefactos encontrados. NABUNKEN comenzó a agregar la información de estas hojas de datos en una base de datos electrónica pública en 2003. Esta base de datos se fusionó con SORAN en 2019 y ahora proporciona metadatos detallados sobre los informes de trabajo de campo y está junto a la información bibliográfica relevante. A partir de 2019, la nueva información se carga directamente en SORAN a través de una interfaz web. Para generar metadatos sistemáticamente para informes anteriores a 1994, la Agencia de Asuntos Culturales envía un aviso anual pidiendo a todas las partes relevantes que preparen datos para un año específico. Por ejemplo, en 2023, la Agencia solicitó la creación de metadatos para informes publicados en 1973. La información recién preparada para informes anteriores a 1994 no se adjunta a las copias en papel, sino que solo existe dentro de

SORAN.

Además, cuando se adjuntan archivos PDF a la información bibliográfica, la página con los metadatos se le asigna un DOI para facilitar aún más la referencia a la publicación.

Toda esta información, incluido el texto completo de los archivos PDF, se indexa y es buscable en SORAN.

SORAN no requiere registro de usuarios. Los archivos PDF se pueden descargar de forma gratuita, y la información bibliográfica está disponible en formatos BibTeX, TSV y JSON. De los tres, el formato JSON es el más informativo e incluye referencias relacionadas con metadatos en bases de datos externas, como la Biblioteca Nacional sobre Dietas.

En SORAN, se indican claramente las licencias de uso tanto de los informes de trabajo de campo como de los metadatos, y el proceso de creación de metadatos está bien documentado y disponible en el portal. Además, los metadatos en SORAN son muy completos, ya que se basan en los resúmenes adjuntos a los informes de trabajo de campo, y los usuarios pueden verificar fácilmente la validez de los metadatos contra las páginas respectivas de los informes.

En resumen, los informes de trabajo de campo y los metadatos relacionados en SORAN, en su mayoría, pueden considerarse FAIR, aunque no fueron diseñados con el principio FAIR en mente. Los dos mayores problemas que deben resolverse antes de que se pueda considerar verdaderamente FAIR son: 1) los metadatos basados en las hojas de datos no siguen estándares internacionales como Dublin Core, y 2) dichos metadatos no están disponibles para su descarga. Este último es el problema más complicado, ya que requiere negociaciones con los proveedores de datos de SORAN. En otras palabras, es necesario convencer a las administraciones locales de que es de interés común que los metadatos estén disponibles para su descarga gratuita.

## Conclusión

La administración de la información arqueológica en Japón es una carrera contra el tiempo. Las personas involucradas en excavaciones de rescate intentan extraer la mayor cantidad de información posible de los sitios arqueológicos antes de que comiencen los proyectos de desarrollo y luego continúan procesando los artefactos e información hasta que, eventualmente, terminan compilando todo lo que han interpretado en los informes finales de trabajo de campo. Luego, reinician el proceso una y otra vez. En este marco, los informes de trabajo de campo se han convertido en los productos finales *de facto* de las excavaciones y se consideran el registro más completo y fácilmente comprensible que queda de los sitios. Por lo tanto, hoy en día hay un

enfoque más orientado a la preparación y preservación de estos informes que a la difusión de la información que contienen. SORAN de NABUNKEN surgió como respuesta a la necesidad de mejorar la localización y accesibilidad de esta literatura gris arqueológica. Lo hace proporcionando un portal donde los interesados pueden cargar el texto completo de sus informes en formato PDF junto con metadatos detallados sobre las excavaciones. SORAN ha mejorado considerablemente la FAIRness de la literatura gris arqueológica japonesa, pero aún hay margen para mejorar. Hasta julio de 2023, solo alrededor del 25% de los informes de trabajo de campo publicados en Japón están disponibles en la web en línea. Además, no todos los metadatos están disponibles para su descarga, ya que requiere más discusiones con los proveedores de datos. Finalmente, en el futuro, para lograr una arqueología verdaderamente basada en datos, SORAN debe ser capaz de extraer y proporcionar datos de mayor detalle, basados en los archivos PDF. Con este enfoque de la administración de los datos de la arqueología japonesa centrados en los conceptos de velocidad y eficiencia, y el enfoque del gobierno japonés en “utilizar” activos culturales, se cree que la calidad de los datos digitales disponibles en línea mejorará y se volverá más FAIR día a día.

## Bibliografía

- Maizō bunkazai kankei tōkei shiryō: Reiwa 4 nendo. (2023). *Bunkachō Bunkazai Dai Ni Ka*. [https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/93861801\\_01.pdf](https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/93861801_01.pdf) [Last accessed: 26 July 2023].
- Tanaka, N. (1984). Japan. In H. Cleere (Ed.), *Approaches to the Archaeological Heritage: A Comparative Study of World Cultural Resource Management Systems* (pp. 82–88). Cambridge University Press.
- Nishiyama, N. (2015). Heritage Management in Present-day Japan. In J. Babb (Ed.), *The SAGE Handbook of Modern Japanese Studies* (pp. 80–98). SAGE Publications.
- Japan: Law for the Protection of Cultural Property. (2009). *Japan Center for International Cooperation in Conservation*, National Research Institute for Cultural Properties: Tokyo.
- Okamura, K. (2011). From Object-Centered to People-Focused: Exploring the Gap Between Archaeologists and the Public in Contemporary Japan. In K. Okamura & A. Matsuda (Eds.), *New Perspectives in Global Public Archaeology* (pp. 77–86). Springer.
- Gyōsei mokuteki de okonau maizō bunkazai no chōsa ni tsuite no hyōjun. (2004). *Maizō Bunkazai Hakkutsu Chōsa Taisei Tō No Seibi Jūjitsu Ni*

*Kansuru Chōsa Kenkyū linkai*. [https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/hokoku\\_06.pdf](https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/hokoku_06.pdf) [Last accessed: 26 July 2023].

Bunkachō Bunkazai Bu Kinenbutsu Ka & Nara Bunkazai Kenkyūjo. (2010). *Hakkutsu chōsa no tebiki: Seiri, hōkokusho*. Monuments and Sites Division, Agency for Cultural Affairs.

Nara Bunkazai Kenkyūjō. (2022) *Nara Bunkazai Kenkyūjō gaiyō 2022. Nara Bunkazai Kenkyūjō*. <http://doi.org/10.24484/sitereports.130391> [Last accessed: 26 July 2023].

Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni suite 1. (2017) *Maizō Bunkazai Hakkutsu Chōsa Taisei Tō No Seibi Jūjitsu Ni Kansuru Chōsa Kenkyū linkai*. <http://doi.org/10.24484/sitereports.71612> [Last accessed: 26 July 2023].

Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni suite 2. (2017) *Maizō Bunkazai Hakkutsu Chōsa Taisei Tō No Seibi Jūjitsu Ni Kansuru Chōsa Kenkyū linkai*. <http://doi.org/10.24484/sitereports.71613> [Last accessed: 26 July 2023].

Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni suite 3. (2019) *Maizō Bunkazai Hakkutsu Chōsa Taisei Tō No Seibi Jūjitsu Ni Kansuru Chōsa Kenkyū linkai*. <http://doi.org/10.24484/sitereports.71614> [Last accessed: 26 July 2023].

Takata, Y. (2023) 2022 nendo sūji de miru zenkoku iseki hōkoku sōran. In Y. Takata (Ed.), *Dejitaru gijutsu ni yoru bunkazai jōhō no kiroku to rikatsuyō 5: LiDAR, 3D dēta, dejitaru ākaibu, SNS, GIS, chitekizaisan ken*. <http://doi.org/10.24484/sitereports.130529> [Last accessed: 26 July 2023].

---

## Improving the FAIRness of Japanese Archaeological Grey Literature

Yuichi Takata<sup>1</sup> y Peter Yanase<sup>2</sup>

<sup>1</sup>Nara National Research Institute for Cultural Properties. Nijō-chō 2-9-1, Nara, Nara pref., Japan. ORCID: <https://orcid.org/0000-0003-3330-3983>

<sup>2</sup>Nara National Research Institute for Cultural Properties. ORCID: <https://orcid.org/0000-0002-7507-7200>

Trabajo en idioma original

### Introduction

Archaeology in Japan is primarily an administrative effort. Between 1976 and 2022, there were only 15,261 archaeological excavations driven by research questions compared to the 331,574 rescue excavations undertaken prior to development projects (*Maizō bunkazai kankei tōkei shiryō*, 2023).

Rescue excavations in Japan are conducted inside a larger framework concerned with the protection of archaeological sites, which in turn is one aspect of a larger program for the preservation of cultural properties. Preservation activities inside this framework are based on the Law for the Protection of Cultural Properties (enacted in 1950, last amended in 2022) (Tanaka, 1984) and center around objects (Okamura, 2011).

The Law states that it is important to “preserve” (*hōzon*), as well as “utilize” (*katsuyō*) “cultural properties” (*bunkazai*), and the concept of “protection” (*hogo*) includes both of these activities. (For a relatively recent complete English translation of the Law, see *Japan: Law*

*for the Protection of Cultural Property*, 2009) The actual process of protection involves such tasks as excavating, collecting, analyzing objects, selecting objects to be designated as cultural properties, repairing and restoring objects, sharing insights gained about the objects with the public, displaying objects in museums, and preserving objects. All of these activities form a cycle of administrative work (Nishiyama, 2015).

The administrative machinery dealing with the protection of cultural properties is decentralized. At the national level, it is coordinated by the Agency for Cultural Affairs, attached to the Ministry of Education, Culture, Sports, Science and Technology; at the prefectural and municipal levels, it is handled by the local boards of education (Tanaka, 1984).

In other words, 95% of the excavations done in the past fifty years in Japan were conducted by administrative agencies as part of their cultural property preservation activities.

Although no legislation explicitly states so, in this

framework, the final fieldwork reports are considered to be the end-products of archaeological excavations. According to the Agency for Cultural Affairs, fieldwork reports are “the most central among the records left for future generations as replacements for buried cultural properties that could not be preserved in their original state, and they make it possible to disseminate and utilize the results of the excavations” (*Gyōsei mokuteki de okonau maizō bunkazai no chōsa ni tsuite no hyōjun*, 2004, p. 5). Furthermore, the Agency states elsewhere that “because fieldwork reports are such records of buried cultural properties that accurately summarise the results of excavations, it is necessary to take measures to preserve them perpetually and make them public so that people could claim common ownership of and utilize them” (Bunkachō Bunkazai Bu Kinenbutsu Ka & Nara Bunkazai Kenkyūjo, 2010).

However, because of the decentralized nature of the Japanese administrative machinery, the production and dissemination of grey literature are not managed coherently. As a result, neither the extent of archaeological grey literature nor the extent to which it has been published is easily identifiable in Japan.

Various attempts have been made in the past to solve this problem, the most recent being the development of the Comprehensive Database of Archaeological Site Reports in Japan (SORAN), an online index of fieldwork reports, currently managed by the Nara National Research Institute for Cultural Properties (NABUNKEN).

In this paper, we will briefly introduce SORAN and related efforts to improve the FAIRness of Japanese archaeological grey literature.

### The Birth of SORAN

SORAN is currently managed by NABUNKEN. NABUNKEN was established in 1952 as an auxiliary organization under the National Commission for Protection of Cultural Properties (the later Agency for Cultural Affairs) for the *in situ* study of movable and immovable cultural properties in the Nara region. It is a member of the National Institutes for Cultural Heritage, an umbrella entity formed in 2007 comprising nine legal entities specializing in cultural heritage-related matters. NABUNKEN is in charge of leading and educating the Japanese community of cultural heritage experts, especially the archaeological community. The institute itself is an active producer of archaeological information as it has been involved in countless excavations in and outside of Japan since its inception. It also manages three museums showcasing its finds and constantly strives to improve current methods of preservation, conservation, restoration, presentation, and maintenance of cultural properties (Nara Bunkazai Kenkyūjo, 2022).

SORAN emerged as a response to the need to improve the findability and accessibility of Japanese archaeological grey literature and the information contained therein.

In 2008, five national university libraries located on the western edge of Japan, with Shimane University Library taking the lead, formed an alliance to publish the full texts of local archaeological fieldwork reports on the Internet. This endeavor eventually morphed into a nationwide project with twenty-one national university libraries participating called the Zenkoku Iseki Shiryō Ripojitori Purojekuto [Nationwide Archaeological Site Information Repository Project]. During the lifetime of this project, nearly 20,000 reports were digitized and made freely accessible online. Although the project was a joint effort, the prefectural datasets were managed separately. It was not until NABUNKEN took over management duties in June 2015 that the accumulated data was merged into SORAN's monolithic database (*Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni tsuite 2*, 2017).

### The Content of SORAN and Related Services

In recent years, SORAN grew considerably with the help of an increasing number of data providers from across Japan. It contains more data and more kinds of data every year, and new services are gradually implemented in the portal. As of July 26, 2023, it contains the following public data:

- 126,455 pieces of bibliographic information on cultural-property-related grey literature made by local administrations and museums
- 143,857 sets of metadata based on the summaries of fieldwork reports
- 118,105 pieces of bibliographic information on scholarly articles published inside fieldwork reports and other obscure non-refereed journals
- 665,621 pieces of geographic locations related to cultural properties:
- 35 pieces of 3D models of structural remains

### Policies Regarding Digital Data

The FAIR principle is not widely known or applied in Japan, especially in the social sciences, and there are no public repositories where researchers can freely upload data or pre-prints. The Japan Society for the Promotion of Science, Japan's sole independent funding agency, manages a portal called Japan Data Catalog for the Humanities and Social Sciences (JDCat), a metadata aggregation site where one can analyze and download metadata in bulk. However, the portal is still in its infancy: as of July 2023, it only has five participating institutions. This situation is believed to gradually improve, though, starting this year, as JDCat has just started calling for new partners. It is important to note that the project only accepts applications from databases with a FAIR pedigree.

On the other hand, the Agency of Cultural Affairs shared its policy regarding archaeological digital data in three volumes between 2017 and 2019 (*Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni tsuite 1–3*, 2017–2019). Volume one is about digitizing photographs and digital photography, volume two is

about the digitization of fieldwork reports, and volume three is about the digitization of primary sources and the proper handling of born-digital primary sources like measurements made with digital devices. These guidelines promote digitizing physical materials for backup purposes and to improve accessibility. At the same time, they also stress the importance of proper management of born-digital files. However, these are but guidelines, and whether local administrations follow them or not is up to them.

### How FAIR is SORAN?

NABUNKEN has learned much about the FAIR principle and other best practices of data stewardship through the various international projects it took part in, such as ARIADNE (Archaeological Research Infrastructure for Archaeological Data Networking in Europe) and SEADDA (Saving European Archaeology from the Digital Dark Ages) and released a Japanese translation the *ADS Guides to Good Practice* (<https://archaeologydataservice.ac.uk/help-guidance/guides-to-good-practice/>) in 2022. However, applying all that knowledge to SORAN takes time. Below we will briefly assess how FAIR SORAN is as of 2023 based on the guiding principles available on the GO FAIR website (<https://www.go-fair.org/>).

Bibliographic information on fieldwork reports is at the core of SORAN's dataset. Every piece of bibliographic information has a unique URL that can be used to refer to it. Bibliographical data in itself is not very useful when it comes to archaeological fieldwork reports because it does not convey much beyond the names and types of the sites. To solve this problem, the Agency of Cultural Affairs started requesting the local governments to attach semi-structured datasheets to the reports in 1994 (Morimoto, 2017). These sheets should contain information on every archaeological intervention covered in a given fieldwork report and record the names, addresses, coordinates, sizes, types, and ages of the sites excavated, the dates and reasons for the excavations, and lists of the types of structural remains and artifacts found. NABUNKEN started aggregating the information from these datasheets into a public electronic database in 2003. This database was merged into SORAN in 2019 and now provides rich metadata on the fieldwork reports and sits next to the relevant bibliographic information. From 2019 onward, new information is directly uploaded to SORAN by its users via a WEB interface. To generate metadata systematically for pre-1994 reports, the Agency of Cultural Affairs sends an annual notice asking all relevant parties to prepare data for a specific year. For example, in 2023, the Agency called for creating metadata for reports published in 1973. Newly prepared information for pre-1994 reports is not attached to the hard copies but exists only inside SORAN.

Furthermore, when PDF files are attached to the bibliographic information, the page with the metadata is assigned a DOI in addition to further ease referencing to the publication.

All of this information, including the full text of the PDF files, is indexed and searchable in SORAN.

SORAN does not require user registration. The PDF files are freely downloadable, and bibliographic information is available in BibTeX, TSV, and JSON formats. Out of the three, the JSON output is the most expressive and includes qualified references to metadata in external databases, such as the National Diet Library.

In SORAN, the usage licenses of both the fieldwork reports and the metadata are clearly stated, and the metadata-creation process is well-documented and available on the portal. Furthermore, as the rich metadata in SORAN is based on the summaries attached to the fieldwork reports, users can easily check the validity of the metadata against the respective pages of the reports.

To sum up, the fieldwork reports and related metadata in SORAN, by large, can be considered FAIR, even though they were not designed with the FAIR principle in mind. The two biggest problems that need to be solved before it could be considered truly FAIR are: 1) the metadata based on the datasheets does not follow international standards like the Dublin Core, and 2) said metadata is not available for download. The latter is the more complicated matter as it requires negotiations with the data providers of SORAN. In other words, local administrations need to be convinced that it is in everyone's best interest that the metadata is made freely downloadable.

### Conclusion

Japanese administrative archaeology is a race against the clock. The people involved in rescue excavations try to extract as much information as possible from the archaeological sites before the development projects start and then continue to process the artifacts and information until, eventually, finish compiling everything they have learned in the final fieldwork report. Then cleanse and repeat *ad infinitum*. In this framework, fieldwork reports have become the *de facto* end-products of excavations and are considered the most complete and easily comprehensible record that remain after the sites. As such, there is currently more focus on the preparation and preservation of these reports than on disseminating the information found inside them. NABUNKEN's SORAN emerged as a response to the need to improve the findability and accessibility of such archaeological grey literature. It does this by providing a portal where stakeholders can upload the full text of their reports in PDF format along with rich metadata about the excavations. SORAN has considerably improved the FAIRness of Japanese archaeological grey literature, but there is still room for improvement. As of July 2023, only about 25% of the fieldwork reports published in Japan is available online. Furthermore, not all of the metadata is yet downloadable as it requires further discussions with the data providers. Finally, for a truly data-driven archaeology, SORAN needs to be able to extract and serve more granular data based on the PDF files in the

future. With Japanese administrative archaeology's focus on speed and efficiency, and the Japanese government's focus on "utilizing" cultural assets, the quality of digital data available online is believed to become better and more FAIR by the day.

## Bibliography

Maizō bunkazai kankei tōkei shiryō: Reiwa 4 nendo. (2023). Bunkachō Bunkazai Dai Ni Ka. [https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/93861801\\_01.pdf](https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/93861801_01.pdf) [Last accessed: 26 July 2023].

Tanaka, N. (1984). Japan. In H. Cleere (Ed.), *Approaches to the Archaeological Heritage: A Comparative Study of World Cultural Resource Management Systems* (pp. 82–88). Cambridge University Press.

Nishiyama, N. (2015). Heritage Management in Present-day Japan. In J. Babb (Ed.), *The SAGE Handbook of Modern Japanese Studies* (pp. 80–98). SAGE Publications.

*Japan: Law for the Protection of Cultural Property*. (2009). Japan Center for International Cooperation in Conservation, National Research Institute for Cultural Properties, Tokyo.

Okamura, K. (2011). From Object-Centered to People-Focused: Exploring the Gap Between Archaeologists and the Public in Contemporary Japan. In K. Okamura & A. Matsuda (Eds.), *New Perspectives in Global Public Archaeology* (pp. 77–86). Springer.

*Gyōsei mokuteki de okonau maizō bunkazai no chōsa ni tsuite no hyōjun*. (2004). Maizō Bunkazai Hakkutsu Chōsa Taisei Tō No Seibi Jūjitsu Ni Kansuru Chōsa Kenkyū linkai. [https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/hokoku\\_06.pdf](https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/hokoku_06.pdf) [Last accessed: 26 July 2023].

[bunka.go.jp/seisaku/bunkazai/shokai/pdf/hokoku\\_06.pdf](https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/hokoku_06.pdf) [Last accessed: 26 July 2023].

Bunkachō Bunkazai Bu Kinenbutsu Ka & Nara Bunkazai Kenkyūjo. (2010). *Hakkutsu chōsa no tebiki: Seiri, hōkokusho*. Monuments and Sites Division, Agency for Cultural Affairs.

Nara Bunkazai Kenkyūjō. (2022) *Nara Bunkazai Kenkyūjō gaiyō 2022*. Nara Bunkazai Kenkyūjō. <http://doi.org/10.24484/sitereports.130391> [Last accessed: 26 July 2023].

*Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni tsuite 1*. (2017) Maizō Bunkazai Hakkutsu Chōsa Taisei Tō No Seibi Jūjitsu Ni Kansuru Chōsa Kenkyū linkai. <http://doi.org/10.24484/sitereports.71612> [Last accessed: 26 July 2023].

*Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni tsuite 2*. (2017) Maizō Bunkazai Hakkutsu Chōsa Taisei Tō No Seibi Jūjitsu Ni Kansuru Chōsa Kenkyū linkai. <http://doi.org/10.24484/sitereports.71613> [Last accessed: 26 July 2023].

*Maizō bunkazai hogo gyōsei ni okeru dejitaru gijutsu no dōnyū ni tsuite 3*. (2019) Maizō Bunkazai Hakkutsu Chōsa Taisei Tō No Seibi Jūjitsu Ni Kansuru Chōsa Kenkyū linkai. <http://doi.org/10.24484/sitereports.71614> [Last accessed: 26 July 2023].

Takata, Y. (2023) 2022 nendo sūji de miru zenkoku iseki hōkoku sōran. In Y. Takata (Ed.), *Dejitaru gijutsu ni yoru bunkazai jōhō no kiroku to rikatsuyō 5: LiDAR, 3D dēta, dejitaru ākaibu, SNS, GIS, chitekizaisan ken*. <http://doi.org/10.24484/sitereports.130529> [Last accessed: 26 July 2023].