

Snapshot of Japan 2016-2018

AN INTRODUCTION TO STANFORD'S
WEB ARCHIVE ON CONTEMPORARY JAPAN

Regan Murphy Kao
Stanford University Libraries

March 21, 2019

1891

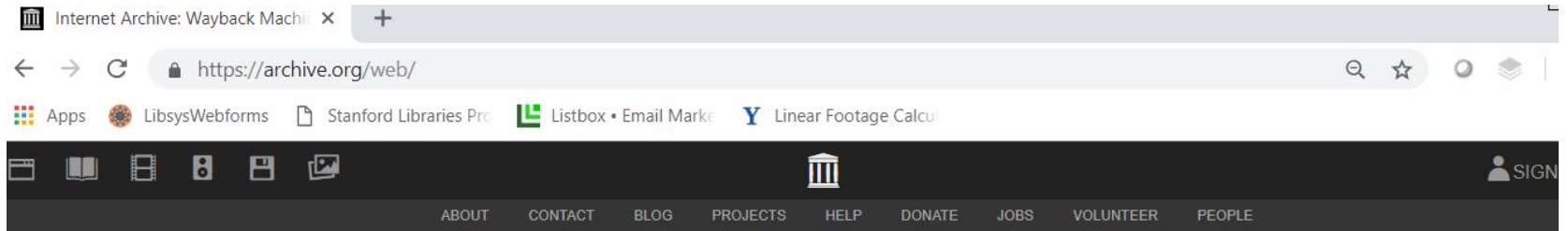
Overview

- Why archive?
- What was archived?
- How was it archived?
- Conclusion: successes and challenges

Why archive?

- To preserve critical writings and artifacts
- To preserve broad spectrum of conversations

Wayback Machine



INTERNET ARCHIVE Explore more than 341 billion web pages saved over time

WayBackMachine

BROWSE HISTORY

DONATE



 **Tools**

- [Wayback Machine Availability API](#)
Build your own tools.
- [WordPress Broken Link Checker](#)
Banish broken links from your blog.
- [404 Handler for Webmasters](#)

 **Subscription Service**

Archive-It enables you to capture, manage and search collections of digital content without any technical expertise or hosting facilities. [Visit Archive-It to build and browse the collections.](#)

 **Save Page Now** 

SAVE PAGE

Capture a web page as it appears now for use as a trusted citation in the future.
Only available for sites that allow crawlers.

Archiving by subject specialists

- Single theme
- Event
- Type (blogs, news media, etc.)

Project's aims?

- Provide a glimpse at a wide range of topics perceived as important at the moment of archiving
- Preserve a diversity of voices
- Compliment not duplicate other efforts

NDL's WARP



国立国会図書館
インターネット資料収集保存事業

Language: **English** | [FAQ](#) | [Help](#) | [Site Map](#)

Quick Search

Search

[Advanced Search](#)

Collection Search



[National Institutions](#)



[Local Governments](#)



[Corporations](#)



[Universities/Colleges](#)



[Political Parties](#)



[Events](#)



[Online Periodicals](#)



[Other](#)

Monthly Special

New Archived Websites (released on WARP in Jul-Sep 2018) (in Japanese)



神戸女子大学



日本ブラインドサッカー協会



野球日本代表

News

2 Nov 2018

An English interface of the WARP is now available!

[Read More](#)

Recommended



[Mechanism of Web Archiving](#)



[Web Archives of the World](#)



[Featured Collections](#)



[Use Cases \(in Japanese\)](#)

Statistics



[Statistics](#)

Monthly Access Ranking

1位: 経済産業省 (2017年6月1日)

51,000

Snapshot of Japan 2016-2018 Web Archive

The screenshot shows a web browser window displaying the Stanford SearchWorks catalog. The browser's address bar shows the URL <https://searchworks.stanford.edu/view/sr233xh9483>. The page header includes the Stanford Libraries logo and navigation links for 'Login', 'My Account', and 'Feedback'. The main search area features the 'SearchWorks catalog' title, a search input field containing 'books & media', and a search button. Below the search bar, there are links for 'Library services', 'Advanced search', 'Course reserves', and 'Selections (0)'. A dark navigation bar contains a 'Back to results' button, a 'Send to' dropdown, and a 'Select' button. The main content area displays the title 'Snapshot of Japan 2016-2018 Web Archive' and a table of metadata:

TYPE OF RESOURCE	Text
DATE CAPTURED	2016-
LANGUAGE	Japanese, English, Chinese, Korean, French

Below the metadata, there is a link to 'View all 156 items' and a horizontal carousel of thumbnail images representing various web pages from the archive.

Topics archived

- **Politics & economics (54):** political parties, financial institutions, political leaders, TPP, security
- **Employment & labor issues (46):** labor unions, industrial cooperatives, irregular employment for youth, at-home work
- **Social issues (156):** aging society, Okinawa, support for women, separate surnames for married couples, food destruction, baby drop box, social withdrawal, international help, LGBTQ, influential social figures, revitalization of rural areas, preservation of old towns, 2020 Tokyo Olympics, blog rankings (politics)
- **Child care & pregnancy (29):** wait lists for child care, authorized nursery schools, egg freezing services
- **Academic & humanities (41):** higher education, religious organizations, university departments (law, economics, sociology)
- **Environment & nature (75):** resources, natural energy, anti-nuclear power generation, global warming, nuclear energy-related institutes
- **Science & technology (41):** robots, advanced materials, research institutes
- **Internet & computers (30):** net harassment, cyber security

Example of archived site

Snap x | Snap x | 遺任 x | Snap x | PAR x | 調査 x | JR通 x | J F x +

← → ↻ https://swap.stanford.edu/20170215215345/http://www.jr-rengo.jp/ 🔍 ☆ 🌐 ⌵ R ⬆

Apps LibsysWebforms Stanford Libraries Proc Listbox • Email Mark Linear Footage Calcu

STANFORD UNIVERSITY LIBRARIES Stanford Web Archive Portal Home | Help

Showing <http://www.jr-rengo.jp/> captured on Feb 15, 2017 < Previous capture | Show overlay | Next capture >

お問い合わせ | リンク | プライバシーポリシー

 約8万人の仲間が加入するJR最大の産業別組合

facebook twitter

JR連合

JR連合とは 安全の確立 交通政策 労働政策 民主化闘争 グループ労組 青年女性活動 機関紙・見解 政治活動

Topics & News

- 2017.02.15 [グループ労組春闘FAXニュースを更新しました](#)
- 2017.02.14 [春闘FAXニュースを更新しました](#)
- 2017.02.13 [民主化闘争情報を更新しました](#)
- 2017.02.07 [機関紙を更新しました](#)
- 2017.02.01 [地域公共交通の課題と展望を更新しました](#)
- 2017.01.24 [JR九州「豊肥本線」の被災現場を視察](#)
- 2017.01.17 [青年・女性委員会機関紙を更新しました](#)
- 2016.12.09 [与党「2017税制改正大綱」にJR連合からの要望事項が反映！！](#)
- 2016.11.22 [石井国土交通大臣へ要請行動！](#)
- 2016.11.22 [2017税制改正に向け、高市総務大臣へ要請行動！（11月21日）](#)



JR九州「豊肥本線」被災現場視察 (1月17日)



2017春季生活闘争

JR北海道問題

結成20周年

東日本大震災

ボランティア活動

JR北労組

JR東海ユニオン

Example: archived site

STANFORD UNIVERSITY LIBRARIES
 Showing <https://www.shinfujin.gsjz/> captured on Feb 14, 2017

New Japan Women's Association
新日本婦人の会 * 新婦人の紹介

女性の願いで行動する会です
平和とジェンダー平等へ
 ようこそ、新日本婦人の会（新婦人）のホームページへ。
 1962年に創立し、国連NGOに認定されている女性団体です。

New Japan Women's Association
新日本婦人の会 * 新婦人の紹介 * 全国の新婦人 * ENGLISH

女性の願いで行動する会です
平和とジェンダー平等へ
 ようこそ、新日本婦人の会（新婦人）のホームページへ。
 1962年に創立し、国連NGOに認定されている女性団体です。

- ★ヒバクシャ国際署名を積み上げ 核兵器禁止条約の実現を→特設ページ
- ★安倍政権退陣を！レッドアクション→特設ページ

Topics 最近のニュース

「共謀罪」阻止連続行動スタート（17.05.10）

今国会で審議されている「共謀罪」法案を阻止しようと、国会周辺で連日の行動がはじまりました。→行動予定はこちら (PDF)



今週の新婦人しんぶん
 5月11日号

新婦人しんぶん
 特別記事 核兵器禁止条約
日本で世界でアクション!

- 核兵器禁止条約 日本で世界でアクション!
- 会員のつぶやまから 公共施設

STANFORD UNIVERSITY LIBRARIES
 Showing <https://www.shinfujin.gsjz/> captured on Nov 10, 2016

New Japan Women's Association
新日本婦人の会 * 新婦人の紹介

女性の願いで行動する会です
平和とジェンダー平等へ
 ようこそ、新日本婦人の会（新婦人）のホームページへ。
 1962年に創立し、国連NGOに認定されている女性団体です。

- ★辺野古新基地建設強行やめて！全国沖縄連帯アクション→特設ページ

Topics 最近のニュース

2017年度予算要求の3署名42万超を提出（17.02.08）

新婦人は2月8日、国会内で2017年度予算要求の「秋の行動」3署名（消費税、社会保障、教育費）42万8289人分を国会議員に託しました。集会には7都県と中央本部など106人が参加。交流では13人が子ども医療費の拡充、就学・・・



今週の新婦人しんぶん
 2月16日号

新婦人しんぶん
 今でも毎月3000万の費用—
**医療・介護の大改善
 もう生活できない**

- ★「戦争法は廃止！」→特設ページ
- ★「消費税増税なんてとんでもない！」→特設ページ

Topics 最近のニュース

【談話】核兵器禁止条約の決議案の採択を歓迎するとともに、反対した日本政府に、満身の怒りをこめて抗議します（16.10.28）

新日本婦人の会 会長 笠井 貴美代は10月28日、談話「核兵器禁止条約の決議案の採択を歓迎するとともに、反対した日本政府に、満身の怒りをこめて抗議します」を発表しました。

今週の新婦人しんぶん
 11月10日号

新婦人しんぶん
 つぶやまの暮らしを大切に
**食品加工室が
 できた!**

MODs metadata

No.1 web_archives_metadata_jpn_update_jc_edits.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Tell me... Sign in Share

Paste Clipboard Font Alignment Number Styles

Calibri 12

General

Conditional Formatting

Format as Table

Cell Styles

Cells Editing

A1 druid

	A	B	C	D	E
1	druid	sourceid	ti:nonSc	ti:title	ti:lang
2	mp548fy3403			自由民主党	jpn
3	qr548ck9856			民進党	jpn
4	my152ng8371			公明党	jpn
5	rz006pq6865			日本共産党	jpn
6	xh105zy8161			日本維新の会	jpn
7	sd199yp6141			自由党	jpn
8	kn444gv1036			社会民主党	jpn

not accessioned yet | authorities ...

ready 100%

How were they archived? What was the process?

- Identify sites to archive
- Send notification letters & respond to inquiries
- Archive-It crawls & quality assurance
- Six month embargo period
- Create metadata
- Ingest into Stanford Digital Repository
- Make archived sites discoverable in SearchWorks & WorldCat

Typical inquiries

- Confirming the URL of the site to be archive
- What other sites were chosen?
- Why was our site chosen?
- We plan to renovate our site, will that cause a problem?
- What will the archived site look like? Please provide an example.
- Will our archived site be viewable after the six month embargo?
- Where will it be possible to view the archived site? Is the below site Stanford's web archive (swap.stanford.edu)?
- We have never received such a notification before. Are you sending it because of some requirement?
- What exactly is “web crawling”(frequency, time, crawler name)
- You mentioned a six month embargo, but can you tell us the exact date when the archived site will be published? Once it is published, could you send us the address to view it?

Dear Dr Regan Murphy Kao

突然の日本語でのメールで失礼いたします。
私、日本政府の内閣官房TPP政府対策本部でインターネットを担当しております重本（Shigemoto）と申します。

この度は、スタンフォード大学図書館のインターネット・アーカイブにおいて、我々のWebサイトをコレクションに加えていただけるとのこと
で、ご連絡を下さりありがとうございます。大変光栄です。

本件につきまして、何点が質問させていただきたく、メールをお送りいたしました。
以下の内容について質問させていただきます。

1. 該当するWebサイトについて
今回、コレクションに加わることになった我々のWebサイトは、内閣官房TPP政府対策本部が管理・運営をする以下のサイトによろしいでしょうか？
<http://www.cas.go.jp/jp/top/index.html>
2. スタンフォード大学図書館のインターネット・アーカイブについて
今回、コレクションに加わることになった我々のWebサイトは、スタンフォード大学図書館のインターネット・アーカイブにおいて保存され、閲覧が可能になるということでしょうか？
また、スタンフォード大学図書館のインターネット・アーカイブは以下のWebサイトになるのでしょうか？
<https://swap.stanford.edu/>
3. アーカイブされたコンテンツの公開時期について
お送りいただいたレターに記載があるように、公開される前の6か月間は閲覧が差し止められるとのことですが、我々のWebサイトがスタンフォー

Notification letters and project webpage

Snapshot of Japan 2016-2018 | S x +

Not secure | library.stanford.edu/eal/japanese-collections/japanese-collection-news/...

My Account Feedback

Stanford LIBRARIES

About Libraries Using the libraries Collections Research support Ask us

East Asia Library

Today: 9am - 10pm
(650) 725-3435
eastasialibrary@stanford.edu

Home » Libraries » East Asia » Snapshot of Japan 2016-2018

Snapshot of Japan 2016-2018

- East Asia Library + An archive of approximately 500 websites touching on central issues of the time, including politics & economics, employment & labor issues, social issues, child care & pregnancy, academic & humanities, environment & nature, science & technology, internet & computers.
- Chinese collections + The aim of the collection is to give a preserve ephemeral websites that fall between the cracks of other collections and give an important, if complex, perspective on this three year period.
- Japanese collections

E-resources

Japanese collection news

Description

The subcategories are as follows:

Conclusion: challenges

- Time-consuming approach
- Technical issues
- Aiming for accurate picture for posterity, but notification letters may have led to some distortion
- Some site owners perceived selection as endorsement (used Wayback Machine for some more heated topics)

Conclusion: successes

- Provided for posterity a glimpse of issues that were considered central during the two-year period.
- Archived sites that wouldn't have been preserved.

Reflecting on the project

ゆく河の流れは絶えずして、しかも、もとの水にあらず。淀みに浮かぶうたかたは、かつ消えかつ結びて、久しくとどまりたる例なし。世中にある、人と栖と、またかくのごとし⁽¹⁾。

WHILE THE RIVER'S CURRENT IS CEASELESS, THE PASSING WATER CONSTANTLY CHANGES. THE BUBBLES THAT FORM ON ITS SURFACE, FORM AND BURST, NEVER LASTING LONG. PEOPLE AND THEIR DWELLINGS ARE THE SAME⁽²⁾.

(1) KAMO NO CHŌMEI, YOSHIDA KENKO. HŌJŌKI. TSUREZUREGUSA. EDITED BY NISHIO MINORU, NIHON KOTEN BUNGA KU TAIKEI, VOL. 30. (TOKYO: IWANAMI SHOTEN, 1957), P23.

(2) MY TRANSLATION OF THE OPENING LINES OF THE HŌJŌKI.

More information

The screenshot shows a web browser window with the URL current.ndl.go.jp/en/ca1930_en. The page features a search bar at the top and a navigation menu with four items: CA-R, CA-E, CA, and 調査研究. The main content area displays the article title "CA1930e - A Snapshot Model for Web Archiving: Stanford East Asia Library's Japanese Web Archive / Regan Murphy Kao" and a link to the Japanese translation. Below the article title, it indicates "Current Awareness No.336 June 20, 2018" and "CA1930". The article title is repeated in a larger font: "A Snapshot Model for Web Archiving: Stanford East Asia Library's Japanese Web Archive". At the bottom, there is a snippet of Japanese text: "ゆく河の流れは絶えずして、しかも、もとの水にあらず。淀みに浮かぶうたかたは、かつ消えかつ結びて、久しくとゞ

Thank you very much!

REGAN MURPHY KAO
EAST ASIA LIBRARY
STANFORD UNIVERSITY
REGANMK@STANFORD.EDU

