

2021 EAJRS conference in Saint Petersburg
17th September 2021, Session 8

How to Connect Japanese Resources : Exploring Use of the *Shibusawa Eiichi* *Denki Shiryo*

日本(語)資料をつなぐ：『渋沢栄一伝記資料』の活用方法を探る
TEIエンコーディングの試み

INOUE, Sayaka 井上さやか
Information Resources Center
Shibusawa Eiichi Memorial Foundation

Outline

1. Introduction
2. Digitization Project phase1 -Digital version
3. Digitization Project phase2 -TEI encoding
-Shibusawa Eiichi Diary
4. Future Tasks and Next Goals
 1. はじめに
 2. プロジェクトフェーズ1：デジタル版の公開
 3. プロジェクトフェーズ2：TEIエンコーディングの試み
「渋沢栄一ダイアリー」の公開
 4. 今後の課題と目標



1

Introduction

はじめに

渋沢 栄一

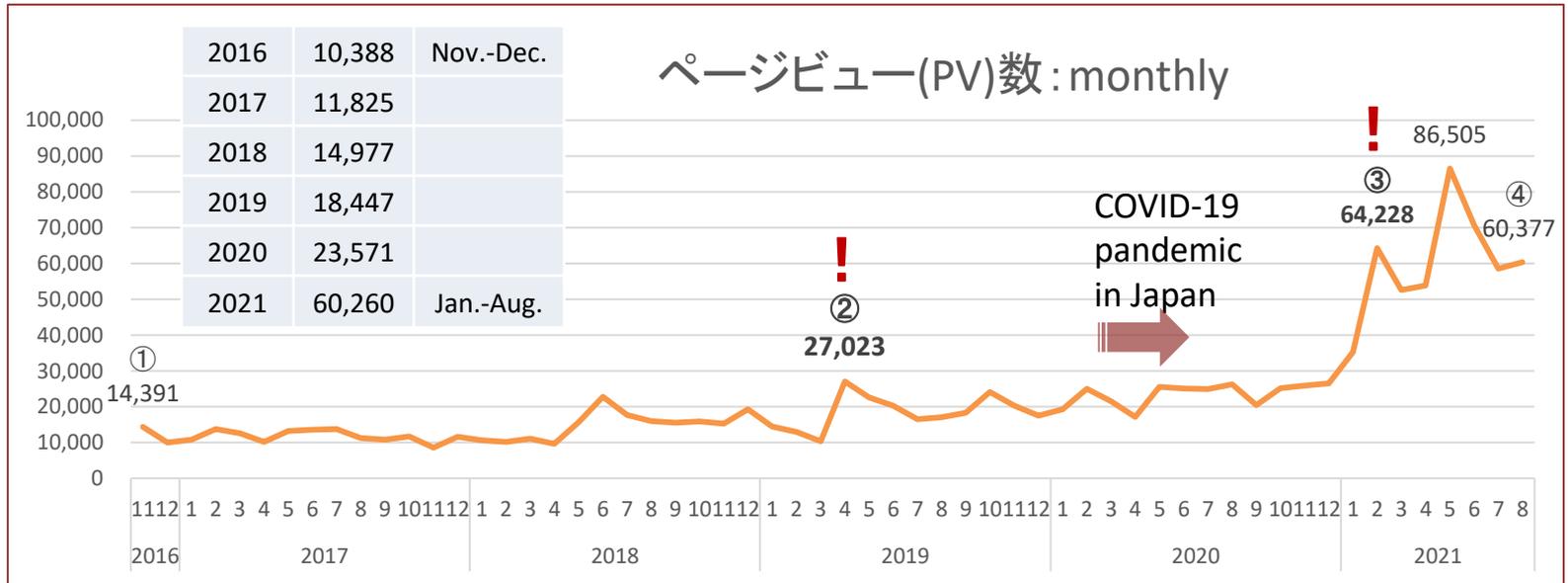
Shibusawa Eiichi
(1840-1931)

a leading figure in the development
of Japan's modern society



Photo : National Diet Library
<https://www.ndl.go.jp/portrait/datas/104.html>

(参考) デジタル版『渋沢栄一伝記資料』アクセス解析



- ① 2016-11 デジタル版『渋沢栄一伝記資料』公開
- ② 2019-04 新1万円札の肖像決定(2024～) !
- ③ 2021-02 NHK大河ドラマ「青天を衝け」放映開始(2～12月放映予定) !
- ④ 2021-08,09 Tokyo 2020 オリンピック&パラリンピック (ドラマ放映はお休み)

Digitization Project phase1 - Digital version

『渋沢栄一伝記資料』デジタル化プロジェクト
フェーズ1: デジタル版の公開



『渋沢栄一伝記資料』

Shibusawa Eiichi

Biographical Materials

(hereinafter called 『伝記資料』)

- ✓ 全68 巻
- ✓ 刊行: 1955-1971年
- ✓ のべ38,000 の資料
- ✓ 48,000 ページ

→ **基礎資料集**

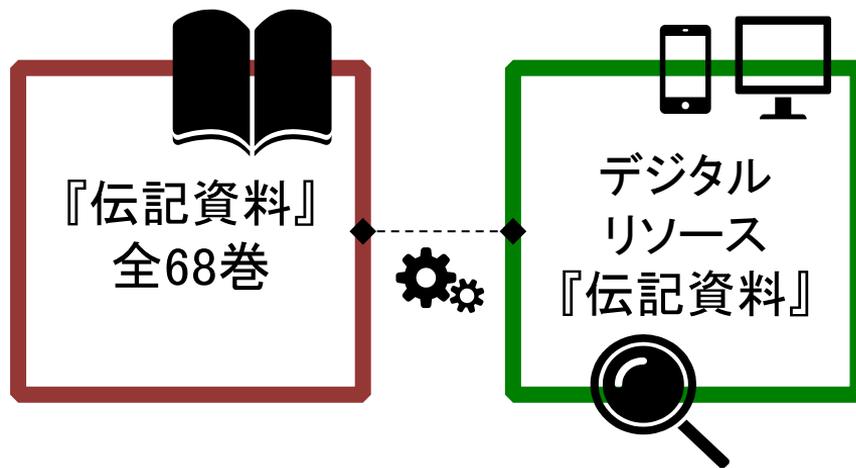
- 1) 渋沢栄一研究
- 2) 近代日本史

→ **膨大なテキスト**



プロジェクトの目的

- 1) 『伝記資料』へのアクセスを改善すること
- 2) 渋沢栄一および日本近代史の情報リソースを創造すること



デジタル化にあたって...

= 書籍同様、基礎資料
集として使えるリソース

+ (デジタルならではの)
利活用の可能性を探る

『伝記資料』の構成

Main Volumes		
Volume	Book	Part Chapters (selected)
1-3	1. Formative Years and Service to the Meiji Government (1840-1873)	1. Formative Years 2. Exile and Years of Service
4-29	2. Building the Japanese Economy and Contributing to Public and Social Works (1873-1909)	1. Business and the Economy 2. Public and Social Works 3. Personal Affairs
30-57	3. Contributing to Social Welfare, International Understanding, and Education (1909-1931)	1. Public and Social Works 2. Business and the Economy 3. Personal Affairs
58	Index	Chronology by Industry; General Table of Contents; Index to the Table of Contents
Supplementary Volumes		
1-2	Diary (1868-1930); Schedules	
3-4	Correspondence; List of Addresses	
5-8	Lectures and Talks	
9	Calligraphy	
10	Photographs	

本編(1-57巻)
事業別編年体
Annals by project

索引(58巻) Index

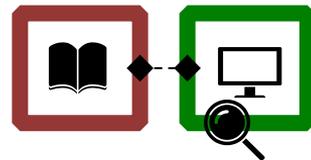
別巻(第1-10)

資料の種別ごとの資料集
Collection by type of materials



デジタル版『渋沢栄一伝記資料』

<https://eiichi.shibusawa.or.jp/denkishiryu/digital/main/>



デジタル版『渋沢栄一伝記資料』

TOP > 各巻リンク > 第1巻【本文】 > 第1巻(DK010001k) 本文

第1巻 (DK010001k) 資料リスト

1冊 在郷及び仕官時代
1部 在郷時代
1冊 幼少年時代

■ 本文

第1巻 p.1-61 (DK010001k) [ページ画像](#)

天保十一年庚子二月十三日 (1840年)

武蔵国秩父郡安部新領血洗島村に生れ、幼名市三郎又栄治郎。幼少時代ノ名乗美雄、後栄一と号す。仕官時代一橋次次、昌平塾太郎と称せしことアリ。父ハ通称市郎右衛門、母ハ伊。家ハ世世農ヲ以テ本業トシ、傍ラ美雄ト製紙トヲ兼テ営ム。

■ 資料

渋沢栄一伝記本 第一章・第一頁【大正八——二年】 (DK010001k-0001)

渋沢栄一伝記本 第一章・第一頁【大正八——二年】

青淵先生、氏は渋沢、名は栄一、青淵は其母なり。天保十一年二月十三日武蔵国秩父郡血洗島に生る。血洗島は閑東平野を成る。阿保に於て、一橋次次、昌平塾太郎と稱せしことアリ。父ハ通称市郎右衛門、母ハ伊。家ハ世世農ヲ以テ本業トシ、傍ラ美雄ト製紙トヲ兼テ営ム。

各巻リンク | 第1巻 目次【本文】 | 第1巻(DK010001k) 資料リスト | ▲ページTOP

第1巻 (天保11年～明治2年) (DK010001k-0001)

第1部 在郷及び仕官時代

第1巻 幼少年時代

第2部 仕官及び仕官時代

第3巻 静岡藩仕官時代

第4巻 長岡大藏高倉仕官時代

第5巻 長岡大藏高倉仕官時代

第6巻 長岡大藏高倉仕官時代

第7巻 長岡大藏高倉仕官時代

第8巻 長岡大藏高倉仕官時代

第9巻 長岡大藏高倉仕官時代

第10巻 長岡大藏高倉仕官時代

第11巻 長岡大藏高倉仕官時代

第12巻 長岡大藏高倉仕官時代

第13巻 長岡大藏高倉仕官時代

第14巻 長岡大藏高倉仕官時代

第15巻 長岡大藏高倉仕官時代

第16巻 長岡大藏高倉仕官時代

第17巻 長岡大藏高倉仕官時代

第18巻 長岡大藏高倉仕官時代

第19巻 長岡大藏高倉仕官時代

第20巻 長岡大藏高倉仕官時代

第21巻 長岡大藏高倉仕官時代

第22巻 長岡大藏高倉仕官時代

第23巻 長岡大藏高倉仕官時代

第24巻 長岡大藏高倉仕官時代

第25巻 長岡大藏高倉仕官時代

第26巻 長岡大藏高倉仕官時代

第27巻 長岡大藏高倉仕官時代

第28巻 長岡大藏高倉仕官時代

第29巻 長岡大藏高倉仕官時代

第30巻 長岡大藏高倉仕官時代

第31巻 長岡大藏高倉仕官時代

第32巻 長岡大藏高倉仕官時代

第33巻 長岡大藏高倉仕官時代

第34巻 長岡大藏高倉仕官時代

第35巻 長岡大藏高倉仕官時代

第36巻 長岡大藏高倉仕官時代

第37巻 長岡大藏高倉仕官時代

第38巻 長岡大藏高倉仕官時代

第39巻 長岡大藏高倉仕官時代

第40巻 長岡大藏高倉仕官時代

第41巻 長岡大藏高倉仕官時代

第42巻 長岡大藏高倉仕官時代

第43巻 長岡大藏高倉仕官時代

第44巻 長岡大藏高倉仕官時代

第45巻 長岡大藏高倉仕官時代

第46巻 長岡大藏高倉仕官時代

第47巻 長岡大藏高倉仕官時代

第48巻 長岡大藏高倉仕官時代

第49巻 長岡大藏高倉仕官時代

第50巻 長岡大藏高倉仕官時代

第51巻 長岡大藏高倉仕官時代

第52巻 長岡大藏高倉仕官時代

第53巻 長岡大藏高倉仕官時代

第54巻 長岡大藏高倉仕官時代

第55巻 長岡大藏高倉仕官時代

第56巻 長岡大藏高倉仕官時代

第57巻 長岡大藏高倉仕官時代

第58巻 長岡大藏高倉仕官時代

第59巻 長岡大藏高倉仕官時代

第60巻 長岡大藏高倉仕官時代

第61巻 長岡大藏高倉仕官時代

第62巻 長岡大藏高倉仕官時代

第63巻 長岡大藏高倉仕官時代

第64巻 長岡大藏高倉仕官時代

第65巻 長岡大藏高倉仕官時代

第66巻 長岡大藏高倉仕官時代

第67巻 長岡大藏高倉仕官時代

第68巻 長岡大藏高倉仕官時代

第69巻 長岡大藏高倉仕官時代

第70巻 長岡大藏高倉仕官時代

第71巻 長岡大藏高倉仕官時代

第72巻 長岡大藏高倉仕官時代

第73巻 長岡大藏高倉仕官時代

第74巻 長岡大藏高倉仕官時代

第75巻 長岡大藏高倉仕官時代

第76巻 長岡大藏高倉仕官時代

第77巻 長岡大藏高倉仕官時代

第78巻 長岡大藏高倉仕官時代

第79巻 長岡大藏高倉仕官時代

第80巻 長岡大藏高倉仕官時代

第81巻 長岡大藏高倉仕官時代

第82巻 長岡大藏高倉仕官時代

第83巻 長岡大藏高倉仕官時代

第84巻 長岡大藏高倉仕官時代

第85巻 長岡大藏高倉仕官時代

第86巻 長岡大藏高倉仕官時代

第87巻 長岡大藏高倉仕官時代

第88巻 長岡大藏高倉仕官時代

第89巻 長岡大藏高倉仕官時代

第90巻 長岡大藏高倉仕官時代

第91巻 長岡大藏高倉仕官時代

第92巻 長岡大藏高倉仕官時代

第93巻 長岡大藏高倉仕官時代

第94巻 長岡大藏高倉仕官時代

第95巻 長岡大藏高倉仕官時代

第96巻 長岡大藏高倉仕官時代

第97巻 長岡大藏高倉仕官時代

第98巻 長岡大藏高倉仕官時代

第99巻 長岡大藏高倉仕官時代

第100巻 長岡大藏高倉仕官時代

書籍
ページ画像
で読む
(png画像)

著作者別
収録資料リスト

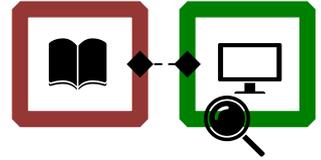
全文検索

全体の構成・目次

テキスト
ページで読む
(html)

デジタル版『渋沢栄一伝記資料』

<https://eiichi.shibusawa.or.jp/denkishiryu/digital/main/>



デジタル版『渋沢栄一伝記資料』

AND OR 検索 詳細検索へ

TOP > 各巻リンク > 第1巻 目次【組文】 > 第1巻(DK010001k)本文

目別 更新履歴 このサイトについて

第1巻 (DK010001k) 資料リスト 第1巻 (DK010001k) 本文 次へ (DK010002k) >

公開日: 2016.11.11 / 最終更新日: 2021.8.26

1冊 在郷及仕官時代

1部 在郷時代

1章 幼少年時代

見出し(index)

■ 網文

第1巻 p.1-61 (DK010001k) ページ画像

天保十一年庚子二月十三日(1840年)

網文(summary)

■ 資料

渋沢栄一伝記資料 第一章・第一頁【大正八——二年】(DK010001k-0001)

第1巻 p.1 ページ画像

資料1(material 1)

渋沢栄一伝記資料 第一章・第一四——一五頁【大正八——二年】(DK010001k-0002)

第1巻 p.1-2 ページ画像

資料2(material 2)

渋沢栄一伝記資料 第一章・第一頁(外題)【大正八——二年】
先生の名、幼少の時は市三郎といひ、又栄治郎と改め、実名を是類とつけたるは十二才前後の事なりしが、後父伯父徳富誠賢の命名によりて栄と改め、之を通称小なせり。安政三年六月先生十七才の時、長高齋に就いて多藝を習ひしに、藝者一足下の通称栄一は好き字なり、孔の言、吾一以貫之といふことより、仁とは仁なり、仁とは諸の推行の統ななり、且名を以て通称するは古の礼なれば、栄一名を以て、字を仁とせしむべし、百行為すに足るものなれど、仁とは足なく仁の一を榮とせしむべし、百行為すに足るものなれど、是より榮一を榮とせしむべし、百行為すに足るものなれど、仁とは一とせしむべし、一稱を出仕の初めに用ふるは、出仕後平岡四郎改めて徳富太夫と称せしむ。明治二年母國に仕へし頃、太夫・新門等の名を改むべき朝命ありしより、再び改めて篤太郎の稱を用たり。後もなく朝任官の後、正式には藩制第一といひ、ヒモカズと訓まはれるが、彼といはなく吾れとせしむるにあらざりし。号を吾れといへるに當り、此の地名の地名に沿ふりて、此辺の地名を湖上といへるに當り、(以上は自筆の題詞)

族類は藩士、初は陸田屋士族なりしに、明治四年六月、東京に移住せし時、請ひて東京府平民に編入、後族類に列せらるゝ、これらの事

天保十一年 二月十三日

第一部 在郷時代

第一章 幼少年時代

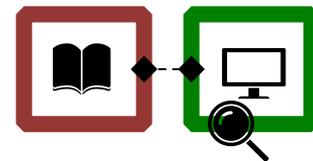
武蔵国様郡那都郡領血洗島村二生ル。幼名市三郎
又栄治郎 幼少時代ノ名乗美濃。仕通稱ヲ栄一ト
名乗ヲ栄一ト改メ、青洲ト号ス。後通稱ヲ栄一ト改メ、
太夫、尋ノ篤太郎 称セシコトアリ。父ハ通称市
郎右衛門 名乗美濃。晩香ト号ス。母ハエ。家
ハ三世世傳リ以テ本草トシ、傍ノ養蚕ト製紙トヲ業
ス営ム。

渋沢栄一伝記資料 第一章・第一頁(外題)【大正八——二年】

吾洲先生、氏は此類、名乗一、吾洲は其号なり、天保十一年二月十三日武蔵国様郡那都郡領血洗島村二生ル。直後以後は開天字號を流る、利根川流域の小村にして、いま大里町八五村に居す。

○ 族類は藩士、初は陸田屋士族なりしに、明治四年六月、東京に移住せし時、請ひて東京府平民に編入、後族類に列せらるゝ、これらの事

いくつかの問題



- 1) 「別巻」「索引巻」の公開ができていない
- 2) 基盤となるテキストが独自ルールのXMLファイル



問題解決の方法を模索

2017～2018年 TEIの存在を知る&入門講習

2019年 未公開「別巻」テキストを対象に勉強会

...難しい

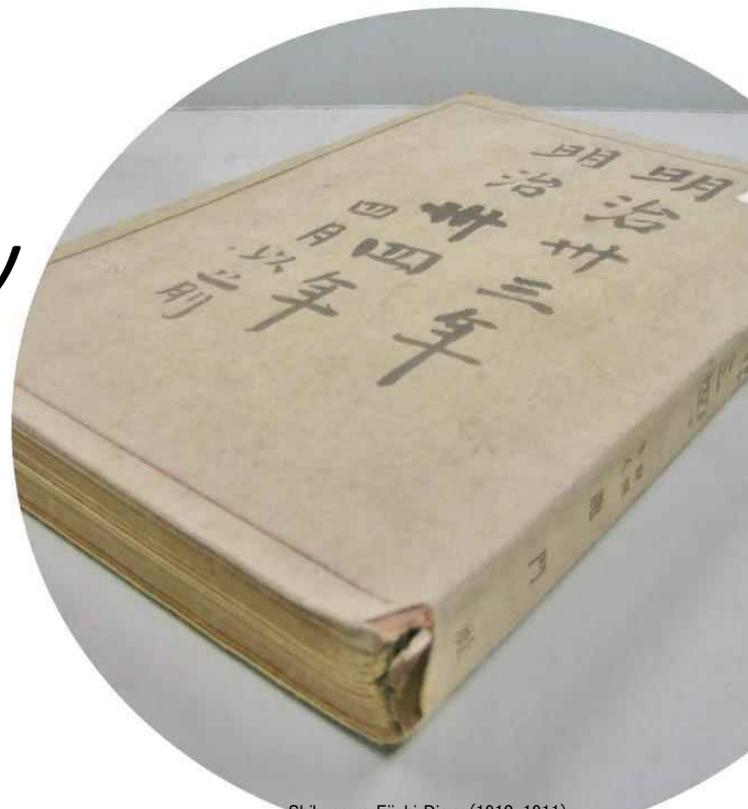
(参考)Text Encoding Initiative (TEI)

<https://tei-c.org/>

TEI エンコーディングの適用を試みる

Digitization Project phase2 - TEI encoding - *Shibusawa Eiichi Diary*

『渋沢栄一伝記資料』デジタル化プロジェクト
フェーズ2: TEIエンコーディングの試み
「渋沢栄一ダイアリー」の公開



Shibusawa Eiichi Diary (1910-1911)
From Japan Industrial History Museum Collection
(National Institute of Japanese Literature)

TEI 適用の目的



- 1) 既存xmlファイルの課題の解消
標準化されたTEI/XML形式
原資料の構造や来歴の記述、長期保存
- 2) 利活用の可能性をさぐる
人文情報学、そのほか研究対象・素材
渋沢関連リソースの利用価値向上
- 3) 別巻・索引巻の公開



人文情報学分野の研究者の協力を得る

TEI 適用の研究



令和2(2020)年度国立歴史民俗博物館総合資料学奨励研究
「TEIを用いた『渋沢栄一伝記資料』テキストデータの再構築と活用」

メンバー: 財団 デジタルキュレーター、人文情報学研究者
(協力) 財団 司書 デジタル化プロジェクト担当

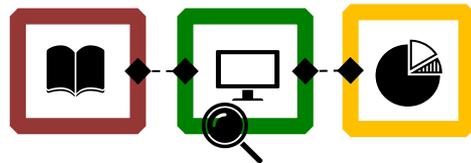
期間: 2020年5月～2021年3月

研究成果

- ・「渋沢栄一ダイアリー」 <https://shibusawa-dlab.github.io/app1/>
- ・”Reconstruction and Utilization of Text Data Using TEI:
Case study of the Shibusawa Eiichi Denki Shiryo”, JADH2021.
https://www.hi.u-tokyo.ac.jp/JADH/2021/Proceedings_JADH2021_rev0905.pdf#page=126

ほか

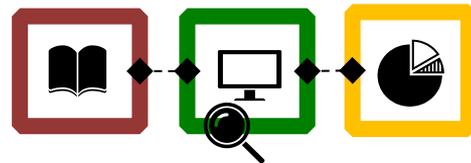
TEI 適用の研究



- ・対象 : 『伝記資料』別巻 第1-2
 渋沢栄一の「日記」「集会日時通知表」
- ・課題 : 1) 近現代日本語資料に汎用性のあるTEIマークアップ手法を提案
 2) 可視化および分析を用いた多角的な研究アプローチを提示
 3) TEIの特徴を活かしたテキストデータの構造化とモデル化、
 アーカイブズ資料への応用可能性を探る
- ・方法 : 1) TEIマークアップ
 原書籍の特徴を反映させたマークアップ方針の設定
 pythonを利用した半自動処理＋手動確認
 2) 可視化と分析
 マークアップ情報を利用 & RDFで外部情報とリンク
- ・成果 : 「渋沢栄一ダイアリー」公開

「渋沢栄一ダイアリー」

<https://shibusawa-dlab.github.io/app1/>



渋沢栄一伝記資料別巻第一日記 (TEI/XML版)

西暦 八月二日 我 六月十四日

朝八時御出発、八時半サンラサア場、九時五十分ニユート、同所御発ダニールと名ク橋セイヌに架ス、ヲウルと云ふ鉄の類ニ而作ル、重廿の捻手三百十七ネチ、直ニBezons上ニ架するなり、川上に架す橋九三十メートル宛あり、汽車道の左置、石を掘出し、巴里に運輸す、其を貯、諸の野菜を作り出し、これを土地ニ而景色も稍佳なり、村落毎の入手能行届けり

巴里より半時ニ而汽車サンセルマンのpremiere station de Paris. ou il y a laprovisionnement 《(provision)》 de bicherie 《(boucherie)》 de Paris. il y a une maison de detention des hommes condamnaient 《(condamnes)》 a morts. pont sur @a Seine 37 arche, les grands qui ne sont pas egales. 十時十五分 Mantesと云ふ所着、巴里の十五里 tunnel de Manvoisin 地道十四メートルと云ふ、十時四十五分 Bueil といふ所着、十一時十五分 Evreuxといふ一小市街に添て過る、汽車中ニ而午餐、しかも貯整たしハ、毎事餐着の差支なし、十一時四十五分 Conches 着、十二時 Beaumont-le-Roger 着、十二時十五分 Serquigny 着、小

渋沢栄一伝記資料. 別巻第1 日記 (慶応4年-大正3年)

書籍
ページ画像を読む
(IIIF Viewer)

テキストを読む (TEI Viewer)

「渋谷栄一ダイアリー」

<https://shibusawa-dlab.github.io/app1/>

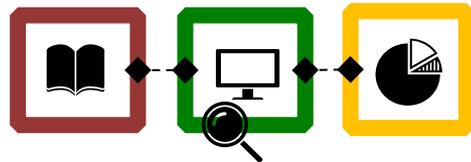


TEIマークアップした要素を活用した、さまざまな可視化と分析
 RDFで外部リソース (Wikipedia) とリンク表示

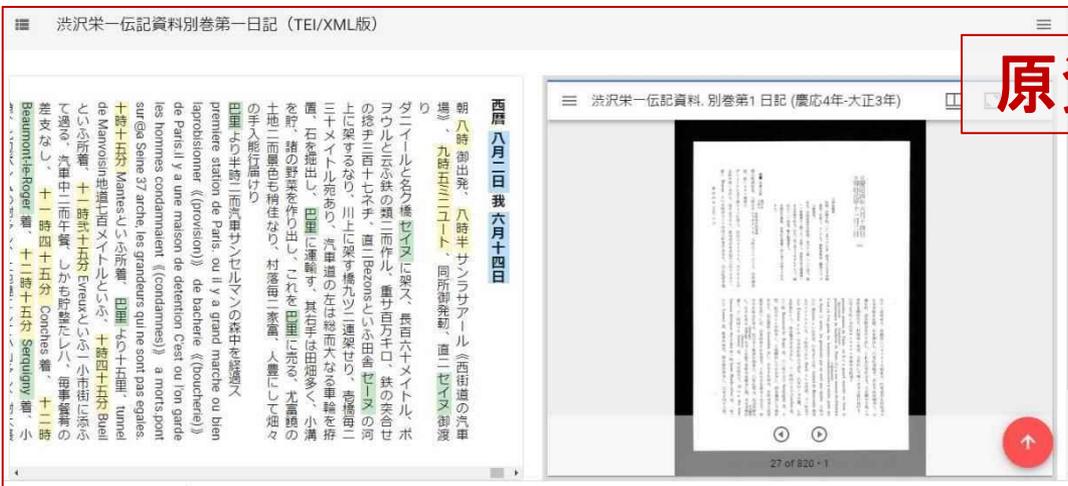
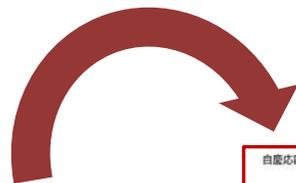


「渋沢栄一ダイアリー」

<https://shibusawa-dlab.github.io/app1/>



さまざまな可視化と分析
原資料情報とリンク



原資料情報

TEI 適用の成果



- 1) 標準化されたテキストデータ
- 2) 利活用の可能性の幅をひろげる
 - テキストそのものを再利用可能な形で公開
 - さまざまな可視化と分析、外部リソースとのリンク
- 3) 別巻の公開を開始

TEI エンコーディングに必要なこと

- ・対象テキスト(資料)の内容・構造の理解 - 当然ながら
- ・DHの知識(TEIの基本知識) - できれば専門家ともコネクション
- ・実際のマークアップ作業には人力と時間 - どうしても

4

Future Tasks and Next Goals

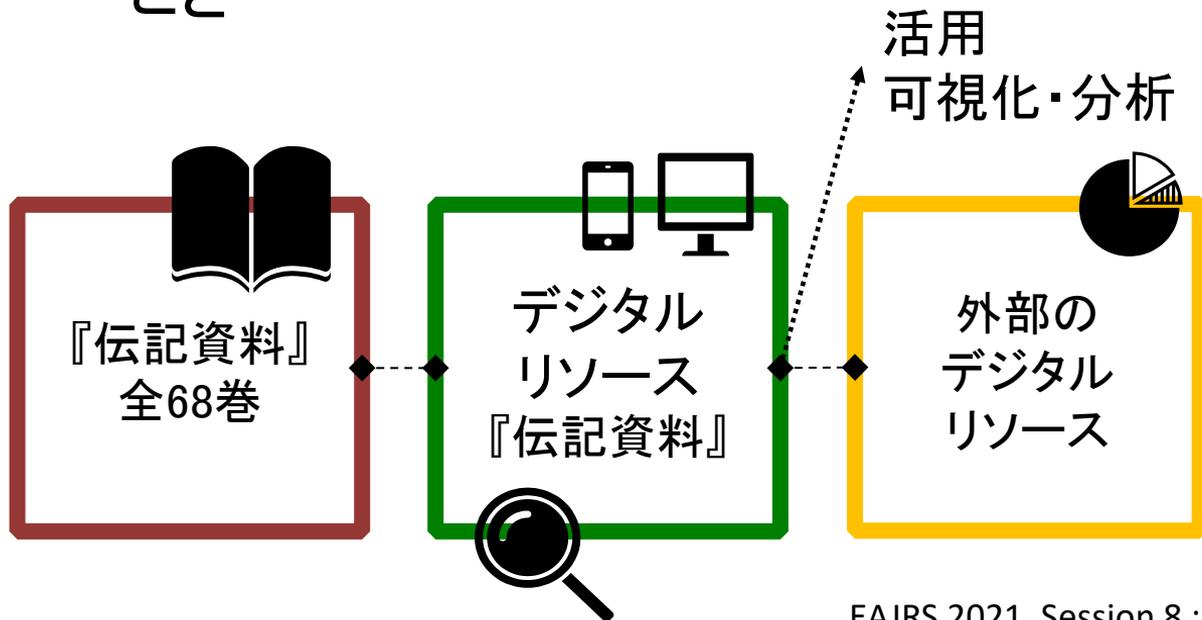
今後の課題と目標



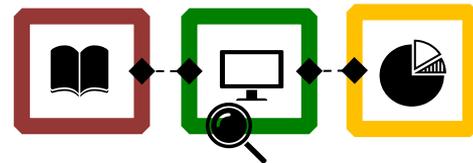
Dai-Ichi Kokuritsu Ginkō
from “Denki Shiryo” supplement, Vol. 10

プロジェクトの目的

- 1) 『伝記資料』へのアクセスを改善すること
- 2) 渋沢栄一および日本近代史の情報リソースを創造すること



今後の課題と目標



- 1) 『伝記資料』別巻の公開(続き)
- 2) 既存リソース(『伝記資料』本編ほか)へのTEI適用
- 3) 外部リソースとの連携



→『伝記資料』全容の公開

→渋沢栄一情報のニーズに応えられているか？

Shibusawa Eiichi
Digital Archive 渋沢栄一デジタルアーカイブ

「見つけよう、
渋沢栄一」。



渋沢栄一記念財団がインターネットで公開するデジタルアーカイブ

渋沢栄一 検索

https://www.shibusawa.or.jp/eiichi/

専責：国立文庫館蔵

Please
visit our
website!

All digital resources
are free to use
and do not require
registration.



Thank you for your kind attention.