

歴史データをひらくこと

ークラウドの可能性ー

橋本雄太（国立歴史民俗博物館）

1. はじめに

大学や博物館、図書館などの文化学術機関において、インターネットを通じて多数の市民の参画を呼びかけ、少数の専門家の手では不可能であった大規模な作業をオンラインで実現する試みが近年始まっています。文献資料のほんこく翻刻、絵図の分類、史料中に現れる地名や人名のタグ付けなど、人文学資料を対象とするさまざまな作業が市民との協働のもと実施されています。さらには、公的資金では賄えない事業資金の寄附をオンラインで呼びかけ、資金調達に成功した事例も増えつつあります。インターネットの分散的な特性を駆使し、世界中の人びとに協力を呼びかけることで問題解決につなげるこうした手法を「クラウドソーシング」と呼びます。

本章では、まず文化学術機関においてクラウドソーシングが普及した背景について説明し、次に人文学資料、特に歴史資料を対象としたクラウドソーシングの主要事例を紹介します。その上で、文化学術領域におけるクラウドソーシングの課題について議論します。

2. クラウドソーシング登場の背景

1990年代から急速な勢いで発展と普及を遂げたインターネットは、遠距離通信にかかるコストを事実上ゼロにただけでなく、人類史上かつてない規模の協働プラットフォームとして成長しました。1990年代から2000年代にかけて、インターネットを基盤とした参加型プロジェクトが次々と実施され、その多くが目覚ましい成果をあげたのです。これらは趣味や慈善といった非営利目的を出発点としつつも、世界中から多数の人びとの協力を引き出すことに成功し、大企業が開発する製品やサービスを遥かに上回る品質や規模を達成しました。その中には世界経済に影響を及ぼしたものも少なくありません。

現在最も高いシェアを有するコンピューターのOS（オペレーション・システム）のひとつであるGNU/Linuxの開発プロジェクトはその代表的事例です。Linuxはもともとヘルシンキ在住のプログラマーであるリーナス・トーバルズ（Linus Torvalds）の個人プロジェクトとして開始しましたが、トーバルズがLinuxの中核部分である「カーネル」のソースコードをインターネット上で公開し、メーリングリストを通じて世界中のプログラマーに開発への参加を呼びかけた結果、1万人以上の開発者が参加する巨大プロジェクトへと成長したのです。LinuxはWebサーバーやスマートフォン、また組み込み機器のOSとして利用されており、現在のICT社会に必要な不可欠な役割を果たしています。

オンラインの百科事典Wikipedia^[01]は、研究者や学芸員などの専門家ではなく、一般のボランティアが記事の執筆を担う参加型プロジェクトです。Wikipediaの前身は、2000年にWeb広告企業のBomis社が設立したNupedia^[02]というオンラインの百科事典サービスでした。Nupediaでは専門家が記事執筆を担当していましたが、厳格な編集方針のため記事公開の遅れが大きな問題となっていました。そこで当時Bomis社に雇用されていた哲学者ラリー・サンガーがwikiエンジンをういたオープンな編集体制を提案し、2001年に

Nupedia と独立のプロジェクトとして公開しました^[03]。誰でも記事編集に参加できる Wikipedia はインターネット上で大きな注目を集め、2001 年の終わりには Wikipedia 上の記事数は 15,000 項目にも上りました。2016 年 8 月時点で Wikipedia 上の公開記事は 520 万件を超え、Wikipedia は世界で 7 番目に訪問者数の多い Web サイトに数えられています。

Linux や Wikipedia は、世界中の人びとに参加を呼びかけることによって大規模な協働を実現しました。インターネット登場以前にこのような大規模協働を実現することは、大企業や政府の主導なくしては不可能だったでしょう。しかしインターネットは人びとの協働の可能性を一変させました。世界中に散らばった興味関心を同じくする人びとが互いに協力し、労働力やリソースを提供することで、最高品質のソフトウェア開発や世界最大の百科事典の執筆が可能になったのです。こうしたプロジェクトの形態は「分散型コラボレーション (distributed collaboration)」や、「ピア・プロダクション (peer production)」と呼ばれます^[04]。同様の事例には、地球外生命探査のためにインターネット経由で計算リソースを共有する SETI@home^[05] や、ボランティアが運営を担うオンラインの電子図書館 Project Gutenberg^[06]、世界中の地図データを作成・共有する OpenStreetMap^[07] などのプロジェクトが知られています。

非営利プロジェクトである Linux や Wikipedia の成果物はオープンライセンスで公開されており、誰もこれらのプロジェクトの恩恵を受けることができます。一方で 2000 年代以降には、インターネットを通じた大規模参加の仕組みを営利活動に利用する試みも登場しました。さまざまな企業が、単純作業から高度な研究開発に至るまで、かつては従業員に担わせていたさまざまな作業をインターネットを通じて世界中の人びとにアウトソース（外部委託）する取り組みを始めたのです。

2001 年に設立された InnoCentive^[08] は、インターネットを利用する多数の人びとの能力を、研究開発における問題解決のプロセスに活用するためのプラットフォームです。InnoCentive の主要顧客は、自組織では解決できなかった

た困難な問題を抱える企業や政府、非営利団体（seekerと総称されます）です。seekerは未解決問題を賞金付きの「チャレンジ」としてInnoCentiveに登録します。これらのチャレンジには「鉄道路線に繁茂する雑草の管理」から「フグ毒の神経作用を中和する手法の開発」までさまざまな課題が含まれます。InnoCentiveにはおよそ38万人のユーザーが問題解決者（solverと呼ばれる）として登録しており、問題の解決手法を提案することで、seekerから賞金を得ることができます。

2005年にAmazon社が開設したMechanical Turk^[09]は、世界中のインターネット利用者を単純労働のアウトソーシング先として利用するオンラインサービスです。Mechanical Turkの利用者（通常は企業）は、写真や動画に映し出された物体のタグgingや画像・音声の文字起こしなど、機械的な自動処理は困難ながら人間なら比較的容易に実施できる単純作業をオンラインで委託することができます。このタスクは細かく分割され、Amazon社と契約した多数の作業者にオンラインで送信されます。作業者はMechanical Turkが提供するWebインターフェイス上でこれらの作業を実施し、作業量に応じて報酬を得ることができます。作業に対する報酬は少額で、2-3セントに満たない場合もあります。しかし2018年時点で発展途上国や貧困層の人びとを中心に、数万から数十万人の人びとがMechanical Turkプラットフォーム上で作業に従事しているといわれています。

InnoCentiveやMechanical Turkは、インターネットの分散的性質を駆使した新しい問題解決手法として注目を浴びました。2006年6月、テクノロジー関係の記事を中心に掲載する雑誌Wiredに「クラウドソーシングの興隆（The Rise of Crowdsourcing）」というタイトルの記事が掲載されました^[10]。この記事が「クラウドソーシング」という用語の初出です。著者のジェフ・ハウ（Jeff Howe）は、前述のInnoCentiveやAmazon Mechanical Turkなどのオンライン事業を記事中で紹介し、インターネットを介したユーザー参加を問題解決や商品開発に活用するこうした事業の形態を「クラウドソーシング」と名付けま

した。

プロダクトデザインソフトウェアからデジタルビデオカメラに至るまで、あらゆる分野における技術発展が、かつてはアマチュアとプロフェッショナルを分け隔てていたコストの壁を打ち崩しつつある。趣味人、パートタイム従業員、道楽家の人々の前に、突如として彼らの労働力を求めるマーケットが出現したのである。これは、製薬業界からテレビ業界に至るまで、多種多様な業界の中から抜け目ない企業が、群衆（crowd）の潜在的能力を駆使する方法を発見した結果だ。彼らに支払う報酬は必ずしも無料という訳ではないが、伝統的な形態の従業員に支払う報酬よりも、ずっと安く済ませることができる。これはアウトソーシングというより、クラウドソーシングと呼ぶべきだ。

ハウの記事が公開されるやいなや、多数の有力メディアやブロガーが「クラウドソーシング」という用語を取り上げ、この言葉はたちまちインターネット上のミーム（流行語）となりました。記事公開からわずか1週間で、Google 検索で“crowdsourcing”を検索した結果は3件から182,000件にまで増加したと、ハウは自身のブログで述べています^[11]。

一方で、ネット上のミームにはありがちなことですが、「クラウドソーシング」の意味するところは曖昧で、この用語はしばしば濫用されることがあります。一部のブロガーの間では、Linux や Wikipedia など先行する参加型プロジェクトをクラウドソーシングに含めるべきか否か議論が起きました。さらには YouTube や Flickr といったユーザー投稿型の Web サービスをクラウドソーシングの一事例に数える人びとも現れました。そこで情報学や経済学分野の研究者から、「クラウドソーシング」の学術的定義がいくつか提案されています。例えば Estellés-Arolas らは、クラウドソーシングの満たすべき性質として次の8条件をあげました^[12]。

1. 「クラウド」にあたる多数の参加者が存在すること
2. 目的が明確に定まったタスクが存在すること

3. 作業者が受け取る報酬が明確であること
4. プロジェクトの運営者（個人や団体）が明確であること
5. プロジェクトの運営者が得る利益が明確であること
6. オンラインで作業を割り当てる参加型プロセスであること
7. 公募によって参加者を集めること
8. インターネットを使用すること

この基準に従えば、ボランティアが運営を担う Wikipedia（条件5に反する）や、ユーザーが自由に動画を投稿することができる YouTube（条件2に反する）をクラウドソーシングに分類することは適切とはいえません。

3. 学術領域に広がるクラウドソーシング

2000年代後半には、クラウドソーシングを学術研究や文化学術機関の事業活動に取り入れる試みが次々と現れました。その適用領域は自然科学から人文学まで広範にわたっています。人文学分野では、次のふたつのコミュニティを中心として活発にクラウドソーシングが実施されています。

ひとつめのコミュニティは、美術館（Gallery）、図書館（Library）、文書館（Archive）、博物館（Museum）など、いわゆる GLAM と総称される文化学術機関です。これらの機関はいずれも市民と直接的交流を持つ公的機関であり、インターネットの普及以前から市民との協働プロジェクトを主催してきた歴史を持ちます。またいずれの機関も膨大な人文学資料を抱えており、これらはしばしばクラウドソーシングの対象となります。

ふたつめのコミュニティは、デジタル人文学（Digital Humanities）分野の研究グループです。デジタル人文学分野では人文学研究におけるインターネットの利活用が一個の研究テーマとして確立しており、クラウドソーシングの方法論についても活発に研究がなされています。もっとも、GLAM とデジタル人文学は非常に近い関係にあり、両者を区別することが難しい場合もあります。

表1 文化学術領域のクラウドソーシングの分類 (Oomen 2011 より)

クラウドソーシングのタイプ	説明
テキストの校正・翻刻	資料のデジタル化データの校正や翻刻にユーザーの参加を呼びかける
文脈情報の付与	資料に文脈知識を付与する(ストーリーを付与したり、文脈情報を含む記事や wiki ページを執筆したりする)
コレクション補完	(Web 上の) 展示やコレクションに共同で資料を追加する
資料分類	コレクション中の資料に関連する記述的メタデータをユーザーが付与する。ソーシャル・タギングが良い例
共同キュレーション	(Web 上の) 展示を編成するために、非専門家の発想や知識を活用する
クラウドファンディング	資金やその他のリソースを共同で提供することで、機関の事業を支援する

クラウドソーシングの対象となる作業の種類はさまざまですが、例えば Oomen らは人文学資料を対象としたクラウドソーシングを、表1にあげる6つのタイプに分類しています^[13]。この分類に従って、実際のクラウドソーシングの事例をいくつか紹介しましょう。

3.1. 校正・翻刻

人文学分野のクラウドソーシングの最も一般的な形態は、Oomen らの分類でいう「校正・翻刻」タイプのプロジェクトです。図書館や博物館が収蔵する活字資料のOCRスキャン結果の校正や、デジタル画像化した手書き文字資料の翻刻(文字起こし)などの作業が、多数の市民との協働のもとオンラインで実施されています。

▶ Australian Newspapers Digitization Project^[14]

2008年に開始したオーストラリア国立図書館による Australian Newspaper Digitization Project (ANDP) は、GLAM分野で実施された最初期のクラウドソーシング・プロジェクトです。ANDPは同館の運営するオンライン資料データベース Trove のサブプロジェクトとして実施されています。Troveでは、19

世紀初頭から 20 世紀中葉までにオーストラリア国内で刊行された 1000 以上の新聞がデジタル公開されており、閲覧可能な紙面画像は総計 2000 万ページに及びます。また、紙面の OCR スキャンによって得られたテキストデータをもとに、新聞記事の全文検索サービスも併せて提供されています。しかし、特に古い時代の新聞は不鮮明な印刷のために、OCR スキャン時に誤認識が頻繁に発生します。そこで Trove は、2008 年 8 月に市民参加を利用した記事の校正システムを導入しました。Trove の利用者は、サイト上の記事閲覧画面で OCR スキャンの誤認識を発見した際に、校正システム上で訂正作業を行うことができます (図 1)。2011 年の時点で、2600 万行の新聞記事がボランティアによって校正されたと報告されています^[15]。

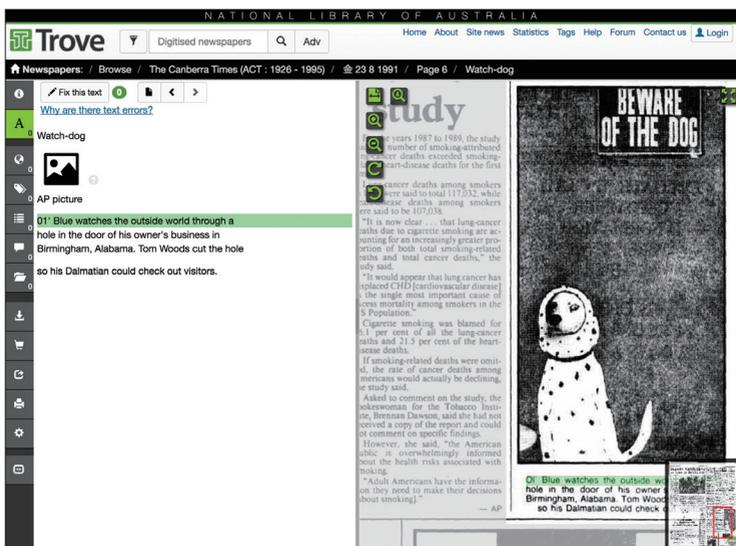


図 1 ANDP のテキスト校正画面

▶ Transcribe Bentham^[16]

ユニバーシティー・カレッジ・ロンドン (UCL) が 2010 年に開始した Transcribe Bentham は、デジタル人文学分野の代表的成果として位置づけられるクラウドソーシング・プロジェクトです。Transcribe Bentham は英国の哲学

者・法学者ジェレミ・ペンサム（1747-1832）の残した6万ページに及ぶ未翻刻の手稿を全文翻刻し、これをもとにペンサム全集の新しい版を刊行することを目的としています。この作業のためにプロジェクトではオンラインで参加するボランティアの協力を募っており、2017年8月時点ですでに18,775ページが翻刻されたと発表されています^[17]。翻刻作業はTranscription Deskと呼ばれるWeb上のエディタで完結するように設計されています（図2）。

▶ みんなで翻刻^[18]

京都大学古地震研究会が2017年に開始した、歴史災害資料の市民参加型翻刻プロジェクトです（図3）。東京大学地震研究所が所蔵する和古書を対象にして、翻刻作業への参加を市民に呼びかけています。「みんなで翻刻」上では2018年末までに4,500名以上のユーザーがサイトに登録し、これまでに翻刻された文字数は500万文字を超えています。大阪大学文学研究科が開発した「くずし字学習支援アプリKuLA」と連携しており、くずし字解読を学習しながら災害史料の翻刻に参加できることが特徴です^[19]。

3.2. コレクション補完

GLAM機関がオンラインで展示するコレクションやデータベースについて、不足している資料情報の追加を市民に呼びかける「コレクション補完」型のプロジェクトも活発に実施されています。

▶ Wir Waren So Frei^[20]

ドイツの映画TV博物館（Deutsche Kinemathek）と連邦政治教育センター（Bundeszentrale für politische Bildung）の連携プロジェクトです。1989年のベルリンの壁崩壊に関する写真や映像資料を収集し、デジタル化することを目的としています。プロジェクトのWebサイトでは、市民から提供を受けてデジタル化された7,000件以上の映像・写真資料が展示されています（図4）。

▶ 寺社・石碑データベース（国立民族学博物館）^[21]

2017年11月、国立民族学博物館のひだかしんこ日高真吾准教授らのグループによって『津波の記憶を刻む文化遺産 - 寺社・石碑データベース -』が公開されました。

UCL Home » Transcribe Bentham » Transcription Desk 60.113.226.166 Talk For This IP Address Create Account Log In

Navigation

- Transcription Desk
- Transcription Guidelines
- Select a Manuscript
- Blog
- Recent changes
- Random page
- Contact Us
- Help

Search

Search []
Go Search

★ [List of un-transcribed material now available](#); [Our FAQ contains a list of common questions](#)

JB/026/015/001

[Click Here to Edit](#)

Feb 1809 3
Press Ch 2 State

27*
Objection. How do you know that these were Ch. Justice's & Atty Gen^{ls} words —
Answer. No. I do not: thence the uncertainty. p.29*

28*
Were these points settled by Statute Law, certainly would or at least could have place. p.30 p.29

29.
Secus under judiciary
Elements of uncertainty enumerated —
1. Reporters & other Bookmakers et quales
2. Judges ubi et quales
p.31

30.

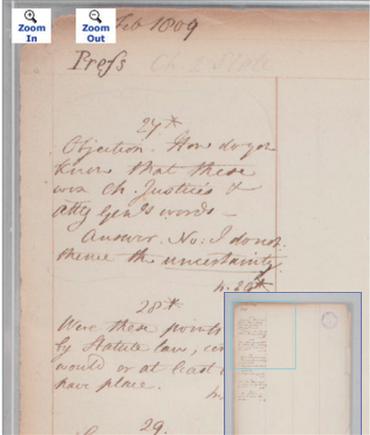


図2 Transcribe Bentham の Transcription Desk

みんなで翻刻【地震史料】

はじめに

- ホーム
- ごあいさつ
- ブログ・SNS・ML

地震史料を翻刻する

- 翻刻ガイドライン
- 史料一覧
- お気に入り・作業履歴
- フォーラム

くずし字解読を学ぶ

- まなぶ
- つながる

ライセンス/クレジット

- お問い合わせ

さんが318文字翻刻しました 1時間前 添削希望

浅間大焼騒動記 中一コマ26
も動れはて夫より種月然へ押寄大勢ふミ入れれ共人悪人も見へず唯本堂計ぞおわしける然共委...

あっぱれ!

さんが424文字翻刻しました 2時間前 添削希望

江戸大地震井二出火場所細見録一コマ3
一新吉原江戸町京町角町揚屋町伏見丁仲の丁すべてくるのは
こらずゆりつづれ? : てみへ出...

あっぱれ!

さんが679文字翻刻しました 2時間前 添削希望

図3 みんなで翻刻

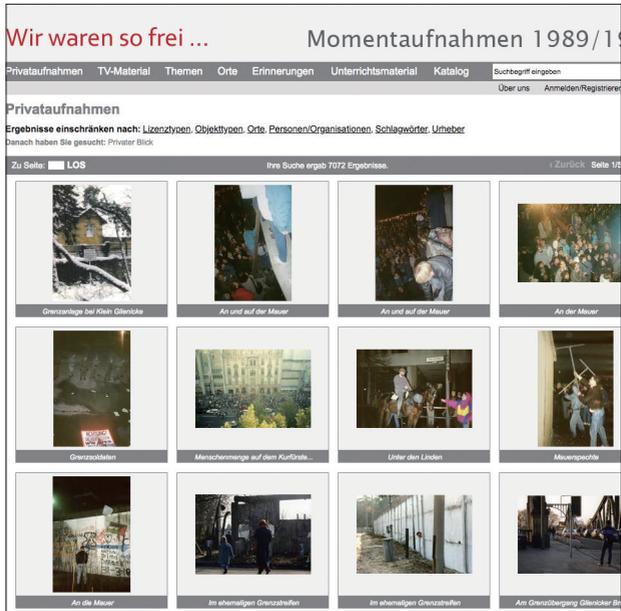


図4 Wir Waren So Freiに寄せられた写真

海嘯溺死精霊塔 津波の年代：明治29（1896）年

岩手県上閉伊郡大槌町吉里吉里第1地割21 塚鼻神社前



撮影日：2012-05-10 15:09:21



碑文 海嘯溺死精霊塔
 建立時期 明治30（1897）年5月

緯度・経度情報



関連する寺社・石碑を探す

詳細 大槌町

図5 寺社・石碑データベース

た。このデータベースは地震や津波災害の記憶を伝える、各地の寺社や石碑についての情報を公開するものです(図5)。しかし日本各地に点在する寺社や石碑の情報を、少数の研究者グループが網羅的に把握することは現実的に困難です。そこで同データベースでは災害に関する寺社や石碑についての情報提供を市民に呼びかけています。市民から寄せられた情報は、データベースの拡充に役立てられます。

3.3. 資料分類

すでにデジタル化されオンラインで公開されている資料について、ユーザーが資料を分類するタグなどのメタデータを追加する形態のプロジェクトです。付与されたメタデータは、資料の内容把握や検索性向上に役立てられます。

▶ Operation War Diary (英国帝国戦争博物館)^[22]

イギリスの帝国戦争博物館(Imperial War Museum)が主催する市民参加型のプロジェクトです。第一次世界大戦に従軍した英国軍兵士によって書かれた150万ページに及ぶ日誌記録について、日時や場所、人名といった重要情報を多数の市民の手によってタグ付けし、検索可能にすることを目的としています。プロジェクトは市民科学のポータルサイト Zooniverse でホストされており、すべての作業は Web で完結します(図6)。

3.4. 共同キュレーション

日本ではあまり事例がありませんが、国外の博物館や美術館には展示のコンセプトやデザインに市民のアイデアや知識を取り込む「共同キュレーション」型のプロジェクトに取り組む機関もあります。

▶ Chicago Authored (シカゴ歴史博物館)

1856年に創設されたシカゴ歴史博物館は、シカゴの歴史と文化を伝える米国を代表する歴史博物館のひとつです。同館は米国の歴史博物館としてはじめてクラウドソーシングの要素を展示に取り入れたことで知られています。同館は2015年に新展示室の開設を予定していましたが、そのコンセプト

も登場しています。

▶ 正倉院文書複製事業（国立歴史民俗博物館）

国立歴史民俗博物館は、日本最古の紙史料群である正倉院文書しょうそういんもんじょの複製を35年にわたり制作してきました。しかしながら近年の予算減の影響で、複製事業に充てる資金の確保が難しい状況が続いていました。そこで国立歴史民俗博物館は、2018年1月よりクラウドファンディング・プラットフォームのReadyfor^[27]上で正倉院文書複製事業のための支援募集を開始しました。当時の目標金額は350万円でしたが、同年3月末の募集終了までに1,064万円もの支援金が538人の人びとにより寄せられました^[28]。

▶ 『新版 緒方洪庵と適塾』刊行事業（大阪大学適塾記念センター）

適塾てきじゅくは幕末の蘭学者・緒方洪庵おがたこうあん（1810-63）が1838年に大阪に設立した蘭学塾です。1942年に当時の大阪帝国大学に寄贈され、現存する唯一の蘭学塾遺構として国の重要文化財にも指定されています。大阪大学は長年適塾の管理運営にあたってきましたが、開塾180年にあたる2018年にあたって洪庵と適塾の歴史的役割を伝える図録の刊行を企画し、上記のReadyfor上で出版費支援を募集しました。このプロジェクトで大阪大学は募集終了の2018年末までに目標金額の1.5倍にあたる307万円の寄附の獲得に成功しています^[29]。

4. 市民との協働手段としての「クラウドソーシング」

ここまで本稿では文化学術機関の主催するプロジェクトについても「クラウドソーシング」という呼称を使用してきました。しかしこの用法をネガティブに捉える識者もいます。例えば米国議会図書館デジタルコンテンツ管理室の室長 Trevor Owens は、この用語の問題を次のように指摘しています。

「クラウド（群衆）」という言葉は、いくぶん誤解を招く言葉である。というのも、最も成功したクラウドソーシング・プロジェクトは大規模な匿名の人間の集団によって成し遂げられたものではないからだ。これら

のプロジェクトは、公共社会を形成するメンバーの参画を促すことによって成功したのである。その成功は、公共利益の創成と発展に寄与してきた、長きにわたるボランティア活動と市民参加の伝統の上に築き上げられている。

(中略)

「ソーシング」という用語の問題は、この言葉が労働と深く結びついていることだ。これは Wikipedia にある「クラウドソーシング」の定義に端的に示されている。「クラウドソーシングとは、分散した人々のグループに対して、タスクをアウトソーシングすることである」。この定義のキーワードは「アウトソーシング」だ。クラウドソーシングはビジネス世界において発明・定義された概念であり、この用語を文化資料の世界に持ち込む際には、よくよく物事を検討しておく必要がある^[30]。

人文学領域の「クラウドソーシング」プロジェクトは、文化学術機関が長年市民との間に育んできた協力関係の延長線上にあり、ビジネスの世界で生まれた「クラウドソーシング」と同一視すべきではないと Owens は主張します。

彼の指摘の通り、ボランティア参加者によって支えられる文化学術領域の「クラウドソーシング」は、企業の営利事業とは性質を異にするものです。自然科学分野の参加型プロジェクトにおいてもクラウドソーシングという呼称は忌避される傾向があり、「市民科学 (citizen science)」など市民との連携を強調した用語が使用される傾向にあります。

一方で「クラウドソーシング」という言葉が非常に便利なラベルであり、すでに学術用語として定着していることも事実です。本稿では「クラウドソーシング」の呼称を採用しましたが、用語の扱いをめぐって上記のような議論が存在することには注意を払う必要があります。

5. クラウドソーシングの諸課題

本節では、研究者や文化学術機関がクラウドソーシング型のプロジェクト

を実施するにあたっての実際上の課題をあげます。

5.1. システムの構築と運用

クラウドソーシングを実施するためには、参加者がデータを入力するための何らかの Web システムが必要です。しかしユーザーとのインタラクションが発生する Web システムは、静的コンテンツのみで構成された Web システムと比べて複雑で、構築にかかるコストも高価になる傾向があります。また、大規模なプロジェクトの場合には、参加者間のトラブルなどに対応するスタッフを置く必要もあります。

こうした費用負担を個人の研究者や文化学術機関が引き受けることは容易ではありません。ただしオープンソースのシステムを利用したり、運営者がスタッフを兼務したりすることでコストを低く抑えることは可能です。

5.2. 参加者の動員

クラウドソーシングを成功させるためには、第一の段階としてプロジェクトを多数の人びとに認知してもらう必要があります。成功したクラウドソーシングの多くは、プロジェクト初期の段階でマスメディアに取り上げられ注目を受けていることも事実です。このためプロジェクト実施に十分な数の人びとの参加を得られるか否かは、主催機関の知名度に左右されてしまう傾向があります。

このような外部要因にとらわれず十分な数の参加者を確保するためには、マスメディアのみに頼らず、地域の学習サークルや市民団体に参加を呼びかけるなど、オフラインのつながりを駆使することが効果的かもしれません。例えば文献資料の翻刻プロジェクトであれば、地域で活動する古文書解読サークルなどと連携することも可能です。

5.3. 参加者の動機づけ

仮に多数の人びとをプロジェクトに誘導することに成功したとしても、そこから継続的にプロジェクトに参画する人びとが現れない限りは、大量のデータ収集やタスクの処理は困難です。プラットフォームの訪問者が継続的

にプロジェクトに参加するためには、そのための動機が存在していなければなりません。

Rogstadius らは、クラウドソーシング参加者がプロジェクトに参加する動機を、金銭的報酬や社会的地位など具体的な見返りのもとに設定されるものを「外的動機 (extrinsic motivation)」、参加者自身の知的関心や社会貢献欲求などに訴えるものを「内的動機 (intrinsic motivation)」に区分し、両者が成果物の品質に与える影響を分析しました^[31]。金銭報酬などの「外的動機」を与えることの難しい文化学術分野のクラウドソーシングでは、参加者の「内的動機」に訴える何らかの要素をプロジェクトに導入する必要があります。

5.4. 作業の難易度

文化学術分野のクラウドソーシングでは、作業を遂行するにあたって高い能力や背景知識を必要とするものがあります。特に日本語の歴史資料を対象としたクラウドソーシングでは、資料解読にかかる難易度の高さが参加者の作業の障害となることがあります。アルファベットで書かれた欧米圏の資料とは異なり、くずし字で書かれた前近代の日本語資料の解読は、相当の訓練を積まない限り非常に困難です。

この課題への対処方法のひとつは、参加者のスキル向上の仕組みを提供することです。前述の「みんなで翻刻」プロジェクトでは、くずし字解読の教育プログラムをシステムに組み込むことでこの課題を解決しています。

5.5. 成果物の品質

作業の難易度は成果物の品質にも直結します。クラウドソーシングを通じて膨大な資料群の翻刻や分類に成功したとしても、誤刻箇所や誤分類の数が多ければ、その成果物に学術的価値を認めることは難しくなります。しかし多数の非専門家が介在するクラウドソーシングにおいて成果物の品質を担保することは容易ではありません。

前述の Transcribe Bentham などの一部プロジェクトでは、専任のスタッフを複数人配置し、成果物のすべてをレビューすることで品質を担保して

います。しかしこの作業は多大な労力と時間を必要とするため、より省コストな方法がいくつか考案されています。例えば資料翻刻プロジェクトの Shakespeare's World^[32] では、ひとつの資料を複数人が独立に翻刻し、結果が一致した場合にのみ妥当な翻刻文として採用する方式を採用しています。

6. おわりに

本稿では、まず2000年代にインターネットを駆使した新しい問題解決手法としてクラウドソーシングが登場した背景と、それが文化学術領域の事業に普及した経緯を説明しました。次に、国内外の機関が主催するクラウドソーシングの主要な事例を紹介し、文化学術領域において「クラウドソーシング」の用語の是非をめぐる議論についても紹介しました。最後に、文化学術機関がクラウドソーシングを実施するにあたっての実際上の課題を5点あげました。

人文学の長い伝統の中で、クラウドソーシングはここ数十年の間に登場した歴史の浅い新規手法です。残念ながら日本国内の文化学術機関ではクラウドソーシングはほとんど浸透していません。しかしながら日本国外に目を向けると、大規模な資料翻刻やタグ付け、資金調達など、少数の研究者グループの手では実現が困難であった数々の成果が、クラウドソーシングを通じて生み出されています。加えてクラウドソーシングは、場所的制約を超えて文化学術機関が市民と協働する手段としても位置づけることができます。市民との協働は、社会にとって「開かれた」人文学の実現にも貢献するはずです。歴史学を含む日本国内の人文学諸分野においても、クラウドソーシングの可能性と課題について本格的に議論を進める時期に来ていると考えるべきでしょう。

——注（Web ページはいずれも 2018-12-1 参照）

[01] Wikipedia, <https://www.wikipedia.org/>.

[02] Nupedia は 2003 年 9 月に閉鎖された。Nupedia という名称は、リチャード・ストールマンの

GNU プロジェクトに影響されたものだといわれる。

- [03] Larry Sanger. The early history of Nupedia and Wikipedia: A memoir. In *Open Sources 2.0: The Continuing Evolution*, pp. 307–338. "O'Reilly Media, Inc.", October 2005.
- [04] Yochai Benkler. The wealth of networks: How social production transforms markets and freedom. *Yale University Press*, 2006.
- [05] SETI@home, <http://setiathome.ssl.berkeley.edu/>.
- [06] Project Gutenberg, <https://www.gutenberg.org/>.
- [07] OpenStreetMap, <https://www.openstreetmap.org/>.
- [08] InnoCentive, <https://www.innocentive.com/>.
- [09] Amazon Mechanical Turk, <https://www.mturk.com/>.
- [10] Jeff Howe, The Rise of Crowdsourcing, <https://www.wired.com/2006/06/crowds/>, June 2006.
- [11] Jeff Howe, Crowdsourcing: Birth of a meme. http://www.crowdsourcing.com/cs/2006/05/birth_of_a_meme.html, May 2006.
- [12] Enrique Estellés-Arolas and Fernando González-Ladrón-De-Guevara. Towards an Integrated Crowdsourcing Definition. *Journal of Information Science*, Vol. 38, No. 2, pp. 189–200, April 2012.
- [13] Johan Oomen and Lora Aroyo. "Crowdsourcing in the cultural heritage domain: opportunities and challenges." *Proceedings of the 5th International Conference on Communities and Technologies*, pp.138-149, 2011.
- [14] Australian Newspapers Digitization Project, <https://www.nla.gov.au/content/newspaper-digitisation-program/>.
- [15] Rose Holley, Crowdsourcing: How and Why Should Libraries Do It?, *D-Lib Magazine*, Vol. 16, No. 3/4, March 2010.
- [16] Transcribe Bentham, <http://blogs.ucl.ac.uk/transcribe-bentham/>.
- [17] Transcribe Bentham 運営チームのブログ記事より, <http://blogs.ucl.ac.uk/transcribe-bentham/2017/08/21/transcription-update-22-july-to-18-august-2017/>.
- [18] みんなで翻刻, <https://honkokur.org/>.
- [19] 橋本雄太 「市民参加型史料研究のためのデジタル人文学基盤の構築」、博士論文、2018年7月、京都大学文学研究科。DOI: <https://doi.org/10.14989/doctor.r13199>.
- [20] Wir Waren So Frei. <https://www.wir-waren-so-frei.de/>.
- [21] 津波の記憶を刻む文化遺産－寺社・石碑データベース－, <http://sekihi.minpaku.ac.jp/>.
- [22] Operation War Diary, <https://www.operationwardiary.org/>.
- [23] Chicago Authored, <http://chicagoauthored.com/>.
- [24] Kickstarter, <https://www.kickstarter.com>.
- [25] CAMPFIRE, <https://camp-fire.jp/>.
- [26] academist, <https://academist-cf.com/>.
- [27] Readyfor, <https://readyfor.jp/>.
- [28] 国立歴史民俗博物館「正倉院文書」複製製作プロジェクト, <https://www.rekihaku.ac.jp/others/news/crowdfunding/index.html>.
- [29] 適塾 緒方洪庵と門人たち, <https://readyfor.jp/projects/handai-tekihuku>.
- [30] Trevor Owens. Digital Cultural Heritage and the Crowd. Curator: *The Museum Journal*, Vol. 56, No. 1, pp. 121–130, January 2013. 翻訳は筆者による。

- [31] Jakob Rogstadius, et al. An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets. *In Proc. ICWSM'11*, January 2011.
- [32] Shakespeare's World, <https://www.shakespearesworld.org/>.