

Journal of Japan Society of Information and Knowledge

情報知識学会誌

Vol.35 No.3 (Sep. 2025)

~~~~~目 次~~~~~

**研究論文**

|                                                         |             |     |
|---------------------------------------------------------|-------------|-----|
| 国会会議録の言語モデルを利用した要約の評価 .....                             | 名知浩一郎, 西村昭治 | 341 |
| Floridi の情報倫理フレームワークの応用可能性に関する理論的検討：功利主義の<br>視点から ..... | 大沼太兵衛       | 350 |

~~~~~

部会報告

JAPAN OPEN SCIENCE SUMMIT 2025 開催報告	情報知識学会オープンサイエンス・オープンデータ研究部会編	368
JAPAN OPEN SCIENCE SUMMIT 2025 個人発表セッション予稿 SSJ データアーカイブのメタデータ作成の AI 活用の検討 - LLM アプリケーション を用いた社会調査のメタデータ抽出 -	西村幸浩, 入山浩一	389
Google フォームを用いた航海データマネジメント効率化に向けた試み	高津佳宏, 齊藤千鶴, 佐藤孝子, 福田和代	395

~~~~~

**エッセイ**

|                                                             |                |     |
|-------------------------------------------------------------|----------------|-----|
| コミュニケーションプロセスとしてのデジタルアーカイブの理論的基礎付け：<br>基礎情報学の情報概念を中心に ..... | 椋本輔, 上松大輝, 山中司 | 400 |
|-------------------------------------------------------------|----------------|-----|

~~~~~

お知らせ

2025 年度総会議事録		401
第 30 回情報知識学フォーラム 「文化と社会をとらえるデータサイエンスの最前線」		412
事務局からのお知らせ		413

~~~~~



情報知識学会

<http://www.jsik.jp/>



すべてを突破する。

TOPPAN  
TOPPAN

研究論文

# 国会会議録の言語モデルを利用した要約の評価 Evaluation of summaries using a language model of parliamentary proceedings

名知 浩一郎<sup>1\*</sup>, 西村 昭治<sup>2</sup>

Kouchirou NACHI<sup>1\*</sup>, Shoji NISHIMURA<sup>2</sup>

1 早稲田大学大学院人間科学研究科

Graduate School of Human Sciences, Waseda University

〒359-1164 埼玉県所沢市三ヶ島2丁目579-15

E-mail: nachies@moegi.waseda.jp

2 早稲田大学人間科学学術院

Faculty of Human Sciences, Waseda University

\*連絡先著者 Corresponding Author

国会会議録には専門用語が多く含まれており、文章量も多いため、一般市民がその内容を理解するのは容易ではない。言語モデル(コンピュータが人間の言語を理解し、生成する技術)を活用して国会会議録を要約することができれば、一般市民の理解を促進できると考えられるが、現状では適切な言語モデルを選定することが課題となっている。本研究では、複数の言語モデルが作成した要約文を、自動評価と人手評価を組み合わせることで総合的に評価し、国会会議録の要約に適した言語モデルを選定することを目的とする。まず、2022年の国会会議録データ(約6110万字)のなかから20段落を無作為に抽出し、言語モデルおよび人手によって要約文を作成した。そして、これらの要約文と原文との類似度を、5つの自動評価指標および人手によって評価した。その結果、BERT言語モデルは原文に忠実で適切な単語選択と自然な文章生成を行うことができ、自動評価と人手評価の両方で高い評価を得た。現時点ではBERT言語モデルが国会会議録の要約に最も適切であると考えられる。

It is not easy for the general public to understand the contents of the Diet proceedings because they contain many technical terms and a large amount of text. If a language model (a technology that allows computers to understand and generate human language) can be used to summarize the Diet proceedings, the understanding of the general public would be facilitated, but currently there are challenges in selecting an appropriate language model. This study aims to select an appropriate language model for summarizing the minutes of a Diet session by comprehensively evaluating summaries produced by multiple language models through a combination of automatic and manual evaluation. First, 20 paragraphs were randomly extracted from the Diet proceedings in 2022 (about 61.1 million letters), and their summaries were created using the language models and by human. Then, the similarity of these summaries and the original sentences were evaluated through the five automatic evaluation tools or manual evaluation. The BERT language model generated most natural sentences with appropriate selection of words, and received high evaluation through both of the automatic and manual evaluations. We conclude that the BERT language model is most suitable for summarizing the Diet proceedings at this moment.

キーワード: 国会会議録, 大規模言語モデル, TF-IDFコサイン類似度, ROUGE, BLEU, BERTScore

Keywords: Diet Proceedings, Large language model, TF-IDF Cosine similarity, ROUGE, BLEU, BERTScore

## 1. はじめに

### 1.1 研究の背景と目的

国会会議録は、国会本会議や委員会等における議事の進行経過や発言内容が記録された重要な公的文書である。国会会議録は一般市民に公開されており、誰でも国会ホームページ上で検索・閲覧することができる。一般市民は国会会議録を閲覧することで、政治への関心を高め、その内容や課題について理解を深められる。しかし、国会会議録を理解するには、難しい専門用語を覚えたり、長い文章を読んだりしなければならず、多くの一般市民にとっては容易なことではない。

一方、最近の AI（Artificial Intelligence）技術と自然言語処理の分野では、いくつかの言語モデルが開発されている。これらの言語モデルは、従来困難であった複雑な言語理解や文の抽出、生成を可能とし、人間の言語処理技術に近いレベルに達している。大規模言語モデル（Large Language Model, LLM）を使った自動要約技術は、膨大な量の情報でも簡潔にまとめられるという特徴を持つ。表1に主なLLMの特徴を示す。このLLMの技術を利用すれば、一般市民が国会会議録を容易に理解できるよ

うになることが期待される。

しかし、このアプローチには課題が存在する。一般的な文章を要約するために、さまざまな言語モデルがすでに開発・評価されてきたが、これらはいずれも国会会議録を要約するために開発されたものではない。そのため、生成抽出された要約の内容が国会会議録の主要なポイントを適切に反映しているかどうか、これまでのところ評価されていない。そこで本研究では、複数の言語モデルを使って国会会議録の要約を生成し、自動評価と人手評価を組み合わせることで各言語モデルの要約精度を比較することを目的とする。それによって、国会会議録の要約に最も適した言語モデルを選定する。

### 1.2 大規模言語モデル(LLM)

LLM とは、膨大なテキストデータを学習した自然言語処理モデルである。その特徴として、多様な言語タスクに対応し、高度な文章生成力と文脈理解力を持つ。これにより、さまざまな文章を柔軟に要約することができる。一方、このデメリットとしては、バイアスやハルシネーション（事実に基づかない情報の生成）の可能性、解釈性の低さなどが挙げられる。

表 1 主な大規模言語モデル(LLM)の特徴

| 言語モデル名<br>/開発者                         | 公開年    | 言語モデル  | パラメータ数     | 特徴                                                  |
|----------------------------------------|--------|--------|------------|-----------------------------------------------------|
| BERT[1][2]<br>/Google                  | 2018 年 | 抽出・生成型 | 約 3 億 4 千万 | Transformer エンコーダで構成され、双方向の文脈を考慮できる、幅広い一般的な言語知識を持つ。 |
| Bert-Extractive-Summarizer<br>/オープンソース | 2019 年 | 抽出型    | 約 3 億 4 千万 | BERT を利用した抽出型要約ライブラリ、文単位での抽出型要約が可能。                 |
| GPT-3.5<br>/OpenAI                     | 2022 年 | 生成型    | 約 3550 億   | GPT-3 を改良し、パフォーマンスが向上。                              |
| GPT-4 (Copilot)<br>/OpenAI             | 2023 年 | 生成型    | 非公開        | GPT-3 を上回る規模とパフォーマンス、マルチモーダル入力に対応し、広範で深い知識を持つ。      |
| Claude<br>/Anthropic                   | 2023 年 | 生成型    | 非公開        | 多言語対応で高い言語理解・生成能力、倫理的な側面にも配慮。                       |
| Cyberagent_Calm<br>/CyberAgent         | 2023 年 | 生成型    | 非公開        | 日本語に特化した独自開発 LLM。                                   |
| Gemini<br>/Google                      | 2024 年 | 生成型    | 非公開        | テキスト、画像など多様なデータで事前学習された LLM。                        |

1.3 評価指標

要約の精度を評価する方法は、大きく自動評価と人手評価、およびそれらを組み合わせたハイブリッド評価に分類される。ここではそれぞれの特徴と先行研究の事例を整理する。

まず自動評価とは、コンピュータアルゴリズムを用いた方法であり、要約精度の評価に広く活用されている。この手法には、BERT\_SCORE、QAGS、TF-IDF コサイン類似度、ROUGE、BLEU など、様々な指標が存在する。表 2 に自動評価の主な特徴と先行研究事例を示す。自動評価手法の主なメリットは、効率性、客観性、そして高い再現性である。大量のテキストデータを短時間で処理できるため、人手による評価と比べて効率が良い。

また、評価基準が一貫しているため、客観的な結果が得られ、同じ条件下で同じ結果を再現できる。一方で、デメリットとしては、テキストの意味的理解や文脈の把握に限界があることである。例えば、皮肉や文化のような複雑な文脈依存表現は、適切に評価できない場合がある。そのため、実際の活用現場では、自動評価と人手による評価を組み合わせるのが一般的である。

人手評価とは、人間が文脈や意味を理解しながら行う手法である。この手法の主なメリットは、人間の深い理解に基づく評価が可能なことである。人間は複雑な文脈や微妙なニュアンスを捉えることができるため、質の高い評価が提供できる。一方、デ

メリットとして、時間とコストがかかり、また評価者間でばらつきが生じる可能性がある。先行研究事例として、Nenkova らは、ピラミッド法を用いた深度ある人間評価の有効性を示した[9]。

自動評価と人手評価を組み合わせた手法は、ハイブリッド評価と呼ばれる。

1.4 その他の要約研究

その他の要約手法としては、特定用途向けの要約研究が挙げられる。これらは特定の分野や課題に焦点を当てており、大規模言語モデルなどを補完する。メリットとしては、要約精度が高く、専門知識を活用した要約の生成が可能となることである。特定の領域に集中することで、その分野特有の言語や構造を深く理解し、適切な要約を生成できる。一方、デメリットとしては、汎用性の低さと評価・比較の難しさがある。特定の用途に特化しているため、他の分野への応用が困難な場合があり、異なる研究間での直接的な比較が難しくなることもある。先行研究事例として、Moreno La Quatra らは、精度の高い金融レポートの自動要約を実施した [10]。

2. 研究方法

2.1 原文の抽出

国会会議録検索システム API を使用して、2022 年の衆議院と参議院の国会会議録デ

表 2 自動評価の特徴と先行研究事例

| 評価指標                 | 概要                       | 特徴                                                       |
|----------------------|--------------------------|----------------------------------------------------------|
| BERT Score [3]       | BERT を使用した評価指標           | 埋め込み類似度を計算、意味的類似度も評価可能で、人手評価との相関が高い                      |
| QAGS[4]              | 質問生成と質問回答を組み合わせた評価指標     | 要約の質(情報の完全性、理解度など)を評価、人手の評価と高い相関を示す                      |
| TF-IDF コサイン類似度[5][6] | 単語の重要度を数値化し、ベクトル間の類似度を測定 | 自然言語処理や情報検索で広く使用、単語の出現頻度(TF)と逆文書頻度(IDF)を組み合わせる重要度を計算     |
| ROUGE [7]            | n-gram の重複に基づく要約文の自動評価   | ROUGE-1(単語一致)、ROUGE-2(二語連続一致)、ROUGE-L(最長共通部分文字列)などの種類がある |
| BLEU[8]              | 機械翻訳の自動評価                | 翻訳文と参照文の語彙重複率に基づいて評価、要約評価にも応用可能                          |

ータを選択した。2022 年の国会会議録文字数は、61135329 個だった。これらはいずれも、国会職員が議案の採決結果を正確に記録し、議長の承認を受けた質の高い内容である。

次にオープンソース Python ライブラリで前処理(形態素解析, 正規表現など)を行うため、20 段落をランダムに抽出し、データの分布を確認した。図 1 に、サンプリングした 20 段落の文字数を箱ひげ図で示す。四分位範囲 (IQR) の 1.5 倍を超える値は外れ値と定義し、以降の分析から除外した。その後、Q-Q プロットによりデータが正規分布に従うことを確認し(図 2)、有効な段落の文字数の平均値、標準偏差、最大文字数、最小文字数等の統計量を算出した(表 3)。

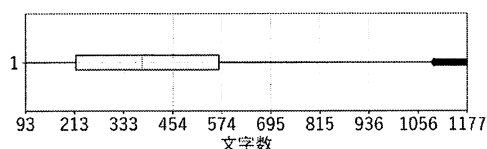


図 1 各段落を構成する文字数の箱ひげ図 (20 段落分)

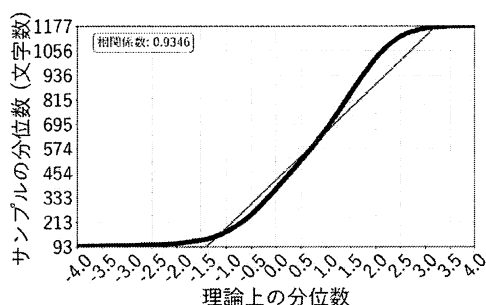


図 2 文字数分布の Q-Q プロット

## 2.2 要約文の作成

### 2.2.1 自動要約

7 種類の言語モデル (BERT, Bert-Extractive-Summarizer, GPT-3.5, GPT-4, Claude, Cyberagent\_Calm, Gemini) を用いて 20 段落の文章を要約し、合計 140 個の自動要約文を作成した。BERT の使用においては、東

表 3 解析した国会会議録の基本統計量

| 項目               | 個数         |
|------------------|------------|
| 2022 年の国会会議録文字数: | 61,135,329 |
| 段落の数:            | 117,057    |
| 各段落の文字数の平均:      | 420.49     |
| 最大文字数:           | 1,177      |
| 最小文字数:           | 93         |
| 中央値:             | 379        |
| 分散:              | 59063.80   |
| 標準偏差:            | 243.03     |
| 分布の幅:            | 1,084      |

北大学日本語 NLP 用 Pre-trained モデル 'cl-tohoku/bert-base-japanese' を使用した。また、生成型言語モデルのプロンプトは、「各段落について 1 番から連番をつけ要約してください」とした。

### 2.2.2 人手要約

人が 20 段落の文章を手作業で要約し、要約文の基準とした。この人手要約文と自動要約文を比較することで、各言語モデルの要約文を評価する。

## 2.3 要約文の評価

本研究では、BERTScore, QAGS (Question-Answering for Evaluating Generated Summaries), TF-IDF コサイン類似度, ROUGE (Recall-Oriented Understudy for Gisting Evaluation), BLEU (Bilingual Evaluation Understudy) の 5 つの自動評価指標を用いた。ROUGE や BLEU などの評価指標は、人手要約と自動要約の間の n-gram の一致度を計算するために開発された。これらは自動要約と人手要約の類似度を示すための指標であり、人手要約文と比較して評価している。

### 2.3.1 自動評価

各要約文 (自動要約および人手要約) に対し、上記 5 種類の指標を用いて評価を行った。評価の比較元としては、BERTScore, QAGS, TF-IDF コサイン類似度については原文 (20

段落)を用い, ROUGE と BLEU については人手要約文(20 段落)を用いた. 各指標について, 20 段落分の要約文を評価し, 20 個の評価点の平均点を, 各指標の最終評価点とした.

### 2.3.2 人手評価

表 4 の人手評価指標に基づき, 各項目について, 5 点(非常に優れている: 要求水準を大きく上回る卓越した内容である), 4 点(良い: 要求水準を満たし良好な内容である), 3 点(標準的: 最低限の要求水準は満たしている), 2 点(やや劣る: 要求水準を一部満たせていない), 1 点(不十分: 要求水準を満たしておらず極めて不十分である)の 5 段階で評価した.

人手評価は, 5 名の評価者が実施した. 各評価者は自動要約文と原文を比較し, 20 個の評価点を付与する. そして, 20 個の評価点の(5 名)平均を最終評価値とした. 5 名の評価者の経歴は, W 大学大学院博士後期課程学生(情報科学), 国立 M 大学 4 年学生(法学), 国立 S 大学 OB(医学保健), A 大学 OB(人間科学), N 大学 OB(経済)である. BLEU, ROUGE 評価用の人手要約文は, 上記評価者のうち 3 名(W 大学大学院博士後期課程学生, 国立 M 大学 4 年学生, 国立 S 大学 OB)が作成した.

表 4 人手評価指標

| 項目     | 内容                                                   |
|--------|------------------------------------------------------|
| 内容の正確性 | 要約が原文の主要な情報をすべて含んでいるかどうかを評価する. 重要な詳細が欠落していないことを確認する. |
| 簡潔性    | 要約が情報を繰り返さず, コンパクトに情報を伝えているかどうかを評価する.                |
| 読みやすさ  | 要約が明確で, 一貫性があり, 文法的に正しいかどうかを評価する.                    |
| 構造的性   | 要約が論理的に組織され, 流れがスムーズであるかどうかを評価する.                    |
| 忠実度    | 要約が原文の意味とトーンを正確に反映しているかどうかを評価する.                     |

## 3. 結果と考察

### 3.1 自動評価

本研究では, 7 種類の言語モデルが生成する自動要約文と原文に対し, 自動評価指標を用いて評価を行った. 表 5 に自動評価の結果を示す. 表の数値は, 原文に対する自動要約文の類似度を示している. これは 0 から 1 の間の値を取り, 1 に近づくほど原文と自動要約文が類似していることを意味する.

#### 3.1.1 要約文の全体的な評価

BERT と Bert-Extractive-Summarizer が多くの評価指標で高スコアを得た. BERT は, ROUGE 系指標 (ROUGE-1: 0.79, ROUGE-2: 0.67, ROUGE-L: 0.66) と BLEU (0.53) で最高スコアを記録し, 生成された自動要約文が原文の内容をよく保持していた. Bert-Extractive-Summarizer は, BERTScore (適合率: 0.88, 再現率: 0.91, F1: 0.89), QAGS (0.76), TF-IDF コサイン類似度 (0.69) で最高スコアを記録し, 抽出型言語モデルの要約文が原文に忠実であることが分かった. 生成型言語モデル (GPT-3.5, GPT-4 など) は, 抽出型言語モデルに比べて, 全体的にスコアが低めであり, 特に原文等との類似度を示す指標が低いスコアとなった. これは, 生成型言語モデルが創造的な要約を生成する傾向があるためと考えられる.

#### 3.1.2 言語モデルの比較

抽出型言語モデル (BERT と Bert-Extractive-Summarizer) は全体的に高スコアを記録した. 生成型言語モデル (GPT-3.5, GPT-4, GEMINI, CLAUDE) は比較的低スコアで, 特に QAGS, TF-IDF コサイン類似度, ROUGE, BLEU は顕著に低かった. Cyberagent\_Calm は全体的に低スコアであった. これらの結果から, BERT と Bert-Extractive-Summarizer が自動評価の精度が高いことが分かった.

#### 3.1.3 要約文と原文間における新出語彙の比較

要約精度を評価するためには, 要約文と原文の語彙の一致性が重要である. 要約文

表 5 自動評価の結果

|   | 項目                   | BERT | Bert-<br>Extractive-<br>Summarizer | GPT-3.5 | GPT-4 | GEMINI | CLAUDE | Cyberagent_<br>Calm | 人手要約 |
|---|----------------------|------|------------------------------------|---------|-------|--------|--------|---------------------|------|
| 1 | BERTScorer<br>P: 適合率 | 0.78 |                                    | 0.65    | 0.64  | 0.64   | 0.61   | 0.56                | 0.75 |
| 2 | BERTScorer<br>R: 再現率 | 0.83 |                                    | 0.73    | 0.73  | 0.74   | 0.71   | 0.64                | 0.83 |
| 3 | BERTScorer<br>F1:F値  | 0.80 |                                    | 0.69    | 0.68  | 0.68   | 0.65   | 0.59                | 0.79 |
| 4 | QAGS:                | 0.40 |                                    | 0.16    | 0.14  | 0.17   | 0.11   | 0.18                | 0.33 |
| 5 | TF-IDFコサ<br>イン類似度:   | 0.47 |                                    | 0.07    | 0.05  | 0.15   | 0.002  | 0.01                | 0.16 |
| 6 | ROUGE -1:            |      | 0.65                               | 0.45    | 0.47  | 0.39   | 0.42   | 0.17                | -    |
| 7 | ROUGE -2:            |      | 0.57                               | 0.18    | 0.22  | 0.16   | 0.13   | 0.03                | -    |
| 8 | ROUGE -L:            |      | 0.64                               | 0.36    | 0.39  | 0.31   | 0.33   | 0.13                | -    |
| 9 | BLEU:                |      | 0.42                               | 0.08    | 0.10  | 0.08   | 0.03   | 0.005               | -    |

※ROUGE , BLEUについては, 人手要約に対する評価となる.

|        |
|--------|
|        |
| 黄: 次点  |
| 緑: 次々点 |

表 6 各言語モデルの要約文における新出語彙数

| 各要約文                         | BERT | Bert-<br>Extractive-<br>Summarizer | GPT3.5 | GPT4 | Gemini | Claude | Cyberagent_<br>Calm | HITO |
|------------------------------|------|------------------------------------|--------|------|--------|--------|---------------------|------|
| 新出語彙数<br>(原文になく要約文に<br>あるもの) | 0    | 0                                  | 70     | 35   | 92     | 85     | 111                 | 0    |

に含まれる単語が原文にも存在することが望ましく, 原文に無い新出語彙は要約の意味を変える可能性がある. 表 6 に各要約文と原文間における新出語彙の比較を示す. 比較方法としては, まず JANOME を使用した形態素解析により, 各要約文を単語に分解した. そして, 各要約言語モデル (BERT, Bert-Extractive-Summarizer, GPT-3.5, GPT-4, Gemini, Claude, Cyberagent\_Calm) および人手要約と, 原文の単語リストを比較し, 新出語彙を特定した. その結果, BERT, Bert-Extractive-Summarizer, および人手要約は, 原文に忠実な要約を生成し, 新出語彙がなかった.

GPT-3.5, GPT-4, Gemini, Claude, Cyberagent\_Calm は, 新しい単語を追加する傾向があった. 特に GPT-3.5 (70) と GPT-4 (35) は比較的少ない新出語彙を生成し, Gemini (92), Claude (85), Cyberagent\_Calm (111) は多くの新出語彙を生成していた (カッコ内の数字は新出語彙数を示す). GPT-4 は他の生成型言語モデルに比べて新出語彙が少なかった.

### 3.2 人手評価

人手評価結果を表 7 に示す. 人手評価により, 要約文脈の把握など, 自動評価では

表 7 人手評価の結果

| 項目/段落                 | 自動要約文 |                                    |         |       |        |        |                     | 人手要約 |
|-----------------------|-------|------------------------------------|---------|-------|--------|--------|---------------------|------|
|                       | BERT  | Bert-<br>Extractive-<br>Summarizer | GPT-3.5 | GPT-4 | GEMINI | CLAUDE | Cyberagent_<br>Calm |      |
| NO1                   | 4.2   | 3.6                                | 4.0     | 3.8   | 4.4    | 4.4    | 2.2                 | 4.4  |
| NO2                   | 3.6   | 3.4                                | 3.4     | 4.0   | 1.8    | 3.6    | 2.4                 | 4.2  |
| NO3                   | 3.8   | 4.0                                | 3.2     | 3.0   | 4.4    | 4.2    | 2.6                 | 4.0  |
| NO4                   | 2.8   | 4.0                                | 3.4     | 3.8   | 4.0    | 4.0    | 3.2                 | 4.2  |
| NO5                   | 3.2   | 3.2                                | 4.6     | 3.6   | 3.8    | 4.2    | 3.2                 | 4.0  |
| NO6                   | 4.4   | 3.0                                | 3.6     | 3.4   | 3.2    | 3.8    | 2.0                 | 4.2  |
| NO7                   | 4.0   | 3.8                                | 2.6     | 3.6   | 2.8    | 3.6    | 2.2                 | 3.6  |
| NO8                   | 3.6   | 3.8                                | 3.0     | 2.6   | 3.6    | 3.8    | 3.2                 | 4.0  |
| NO9                   | 3.6   | 3.2                                | 3.6     | 3.4   | 3.0    | 4.4    | 3.0                 | 4.0  |
| NO10                  | 4.2   | 3.6                                | 3.8     | 3.0   | 2.6    | 3.8    | 1.6                 | 3.8  |
| NO11                  | 4.2   | 3.6                                | 3.6     | 3.8   | 3.2    | 3.8    | 3.0                 | 4.0  |
| NO12                  | 4.0   | 4.6                                | 4.0     | 4.0   | 4.2    | 4.0    | 3.4                 | 4.4  |
| NO13                  | 4.0   | 3.8                                | 4.0     | 3.6   | 3.8    | 4.6    | 2.0                 | 3.8  |
| NO14                  | 4.2   | 3.6                                | 4.0     | 3.4   | 3.0    | 3.6    | 3.0                 | 4.2  |
| NO15                  | 3.8   | 3.4                                | 3.2     | 3.0   | 3.4    | 3.6    | 2.4                 | 3.4  |
| NO16                  | 3.4   | 3.6                                | 3.4     | 3.4   | 3.2    | 3.6    | 2.0                 | 4.0  |
| NO17                  | 4.0   | 3.4                                | 4.0     | 3.8   | 3.6    | 3.4    | 2.4                 | 4.4  |
| NO18                  | 3.8   | 3.6                                | 3.4     | 3.6   | 3.8    | 4.0    | 3.2                 | 4.6  |
| NO19                  | 4.0   | 3.4                                | 3.4     | 4.0   | 3.0    | 3.8    | 2.2                 | 4.0  |
| NO20                  | 4.0   | 2.8                                | 3.8     | 3.6   | 3.2    | 3.8    | 2.2                 | 4.0  |
| 平均評価点<br>(最高5点)       | 3.8   | 3.6                                | 3.6     | 3.5   | 3.4    | 3.9    | 2.6                 |      |
| 合計評価点<br>(最高<br>100点) | 76.8  | 71.4                               | 72.0    | 70.4  | 68.0   | 78.0   | 51.4                |      |

|       |
|-------|
| 黄:次点  |
| 緑:次々点 |

捉えにくい質的側面を適切に評価することが可能となる。

3.2.1 要約文の全体的な評価

人手要約が最高点 (81.2 点) を獲得した。言語モデルでは生成型 CLAUDE (78.0 点) が人手要約に最も近く、僅差で生成・抽出型 BERT (76.8 点) が続いた。Cyberagent\_Calm は全体的に最も低い評価 (51.4 点) となった。

3.2.2 言語モデルの比較

表 8 の結果から、CLAUDE と BERT が他の言語モデルよりも高評価を得ていることが分かった。また、GPT-3.5, GPT-4, GEMINI は比較的近いスコアを示した。

3.2.3. 言語モデルと人手要約の比較

BERT と CLAUDE は人手要約に近い評価を示しており、一部の段落では人手要約と同等の評価を得ていた。この分析から、BERT と CLAUDE が優れた言語モデルであることが分かる。一方で、言語モデル間で評価点の差があり、段落の内容により各言語モデルの強みが異なることも示唆された。

3.3 総合評価と言語モデルの選定

本研究では、国会会議録の要約に適した言語モデルを選定するため、複数の言語モデルから生成される要約文を評価した。その結果、自動評価では、抽出型言語モデル

(BERT, Bert-Extractive-Summarizer) が原文に忠実に重要な情報を要約した。生成型言語モデル (GPT-3.5, GPT-4, GEMINI, CLAUDE, Cyberagent\_Calm) は、新しい単語を追加し、創造的な要約を作成した。人手評価によって、Claude や BERT は人手要約に匹敵する要約を生成することが示唆された。

総合的な評価には、自動評価と人手評価の両方を考慮することが重要である。国会会議録の要約に適した言語モデルには、原文の重要情報を保持することや、生成・抽出型の要約を作成することなどが求められる。

現時点では、BERT 言語モデルが国会会議録の要約に最も適切であると判断した。BERT は抽出型と生成型の両方の特徴を持ち、原文に忠実で適切な単語選択と自然な文章生成を行うことができた。その結果、自動評価と人手評価の両方で高い評価を得た。

#### 4. まとめ

本研究では、国会会議録の要約に適切な言語モデルを選定した。これにより、一般市民が膨大な会議録を効率良く理解する道筋を示すことができた。

今後の課題として、以下の点が重要である。まず、国会会議録の要約評価の効率化が挙げられる。人手による評価は時間、労力、コストがかかるため、言語モデルを活用した自動要約が必要である。そのためには、自動要約の精度を評価する研究が求められる。次に、大規模言語モデルの利用に際しては、公平性やプライバシーなどの倫理的問題に配慮し、社会的影響を慎重に検討しなければならない。また、各言語モデルの長所を活かし、短所を補完する手法の開発も求められる。これらの課題に取り組みながら、国会会議録の自動要約手法の確立を目指していく必要がある。

#### データ利用可能性宣言

本研究中に作成された、分析に用いたデータセットは、妥当な理由があれば責任著者から入手可能である。

#### 参考文献

- [1] Jacob, Devlin; Ming-Wei, Chang; Kenton, Lee; Kristina, Toutanova: “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics, Human Language Technologies, Association for Computational Linguistics, 1, pp.4171–4186, 2019.
- [2] Yang, Liu; Mirella, Lapata: “Text Summarization with Pretrained Encoders, Institute for Language”, Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, pp.3730–3740, 2019.
- [3] Tianyi, Zhang; Varsha, Kishore; Felix, Wu; Kilian, Q. Weinberger; Yoav, Artzi: “BERTScore: Evaluating Text Generation with BERT”, Published as a conference paper at ICLR, pp.1-43, 2020.
- [4] Alex, Wang; Kyunghyun, Cho; Mike, Lewis: “Asking and Answering Questions to Evaluate the Factual Consistency of Summaries”, Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp.5008–5020, July 5- 10, 2020.
- [5] Karen Spärck Jones: “A statistical interpretation of term specificity and its application in retrieval”, Journal of Documentation, Vol.60, No.5, pp.493-502, 2004.
- [6] G. Salton; A. Wong; C. S. Yang: “A Vector Space Model for Automatic Indexing”, Communications of the ACM Published November, Vol.18, No.11, pp.613-620, 1975.
- [7] Chin-Yew, Lin: “ROUGE: A Package for Automatic Evaluation of Summaries”, Text Summarization Branches Out, Barcelona, Spain, Association for Computational Linguistics, pp.74–81, 2004.
- [8] Kishore, Papineni; Salim, Roukos; Todd, Ward; Wei-Jing, Zhu: “BLEU: a Method for Automatic Evaluation of Machine Translation”, Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics,

Pierre Isabelle, Eugene Charniak, Dekang Lin, pp.311–318, 2002.

[9] Ani, Nenkova; Rebecca, Passonneau: “Evaluating Content Selection in Summarization: The Pyramid Method” , Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics HLT-NAACL 2004, Association for Computational Linguistics, pp.145–152, 2004.

[10] Moreno, La, Quatra; Houda, Bouamor: “End-to-end Training For Financial Report Summarization”, Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, Barcelona, Spain, December 12, pp.118–123, 2020.

(2024年 4月17日 受付)

(2024年12月24日 採択)

(2025年 3月14日 J-STAGE早期公開)

研究論文

## Floridiの情報倫理フレームワークの応用可能性に関する理論的 検討: 功利主義の視点から

### A Theoretical Consideration on Applicability of Floridi's Information Ethics Framework from the Utilitarian Viewpoint

大沼太兵衛\*

Tahee ONUMA\*

山形県立米沢女子短期大学

Yamagata Prefectural Yonezawa Women's Junior College

〒992-0025 山形県米沢市通町6-15-1

E-mail: oonuma@yone.ac.jp

\*連絡先著者 Corresponding Author

Luciano Floridiが現代の高度情報社会における包括的な規範倫理として展開する情報倫理 (Information Ethics, IE) の理論は、情報関連分野における倫理をめぐる議論において影響力を強めつつある。本研究では、その理論的フレームワークを図書館等の応用領域において実装するための方法論について考察を行った。まず、応用の観点から解決されるべきIEの理論的課題を、①基本原則の直接適用の困難性、②定量的把握の困難性、③善悪概念の具体的内容の曖昧さ、の3点に特定した。次に、既存の規範倫理理論のうち功利主義に着目し、IEの有する功利主義的側面を踏まえて規則功利主義の考え方を援用することで、上記課題の解決が図られることを示唆した。結論として、IEの倫理原則に、①二次的規則、②定量的観点、③分配の観点を導入することにより、IEの倫理原則を実装に繋げる可能性を提案した。

Information Ethics (IE), which has been developed since the late 1990s by Luciano Floridi as a comprehensive normative ethics for present-day advanced information society, is becoming increasingly influential in various discussions on ethical issues in information-related fields. In this study, we examined the theoretical framework of IE and considered possible methodologies for implementing it in applied fields such as libraries. First, we identified three theoretical issues to be resolved in IE from the perspective of application: (1) the difficulty of directly applying its fundamental principles, (2) the difficulty of quantitative evaluation, and (3) the ambiguity of the specific implications in the concepts of good and evil. Next, focusing on utilitarianism among the existing normative ethical theories, we suggested that the above issues could be resolved by adopting the idea of rule utilitarianism based on

the utilitarian dimensions of IE. In conclusion, we proposed the possibility of implementing the ethical principles of IE to practice by introducing: (1) secondary rules, (2) a quantitative perspective, and (3) a perspective on distribution, after positioning the ethical principles of IE as the primary rules.

キーワード: 情報倫理, 情報の哲学, 功利主義, 規則功利主義

Keywords: Information Ethics, Philosophy of Information, utilitarianism, rule utilitarianism

## 1 はじめに

情報哲学者Luciano Floridiは、現代の情報通信技術の急激な進展と高度情報社会の到来を、コペルニクス革命、ダーウィン革命、フロイト革命に続く、人類史上の「第四の革命」として位置づけている (Floridi [1])。この技術的・社会的変化を背景として、Floridiは情報概念を核とした独自の包括的な哲学体系の構築を1990年代後半に開始した。「情報の哲学」(Philosophy of Information, 以下「PI」)と呼ばれるこのプロジェクトは未だ進行中であるが、その理論の主要部分の提示は、Floridiが三部作と呼ぶ著作 (Floridi [2][3][4]) の刊行によっていったん完了している。

PIは「(a)情報の概念的性質と基本原理についての、その力学、利用、科学も含めた批判的な探究」及び「(b)情報理論的・計算論的な方法論の、哲学的問題に対する精緻化と応用」に関わる領域であるとされている (Floridi [5])。すなわちPIは一方で、独自の存在論に立脚した高度に思弁的かつ包括的な側面、すなわち世界を情報的観点から把握しなおす第一哲学 (philosophia prima) としての性格を有している。他方、PIはその基礎原理から導出される独自の規範倫理を内包している。この倫理理論をFloridiは大文字で「情報倫理」(Information Ethics, 以下「IE」)と呼ぶ。規範倫理であ

るということは、善悪や正不正の判断に係る原則や基準を備えているということであり、これは何らかの現実の対象について倫理評価を行える潜在力を有することに他ならない。前述の三部作にも倫理を主題とする独立した著作 (Floridi [3]) が含まれる事実が示すように、情報に関わる倫理上の諸問題はPIの問題系の重要な一角を占めるものとして位置づけられており、IEは高度情報社会における倫理規範としての援用が想定されている。

Floridi [6][7]は、図書館情報学 (以下「LIS」) とPI・IEとの関係について、PI・IEはLISの哲学的基盤であり、両者はそれぞれ基礎理論と応用領域という形で互惠的な関係にあると主張する (なお、ここでFloridiがLISと呼ぶものには学術研究だけではなく図書館実務も含まれており、本論文でもその用語法に従う)。これは換言すれば、LISがPI・IEをその哲学・倫理学上の基盤として採用することで、何らかの実効的な効用がもたらされるという主張に他ならない。実際、後述するとおり、あらゆる存在を等しく情報オブジェクトと見なし、世界を「情報圏」として捉え直してそのウェルビーイングを評価するIEは、情報の適切な保存・組織化・流通等を通じて全体として健全な情報のライフサイクルの実現を促進するLISとは親和性が高く、また、デジタルデータもアナログ資料も同等の情報オブジェ

クトとして扱うPI・IEは、現代の図書館をめぐるハイブリッドな情報環境にも適合的であると考えられる。そのため、IEの倫理原則（何を善とし、何を悪とするのか）を何らかの形でLISに応用することは、その倫理評価の側面において、Floridi [6][7]が主張するように実効的な効用をもたらすことができる可能性がある。

一方で現在、LISの議論においては、PI・IEをLISの哲学的・倫理的基盤として支持する言説は一部に存在するものの、具体的な応用に関しては、その方法論が必ずしも明確にされていないこともあり、その実装に係る実質的な議論はあまり進んでいない現状にある。

このような背景を踏まえ、本論文は、IEの規範倫理性に着目し、その理論的フレームワークをLIS、特に図書館実務等へ実装するための検討を行う。具体的には、実装の観点におけるIEの理論上の課題を特定し、さらに当該課題に対する解決方法を提案する。その際、既存の倫理理論の中でも構造的にIEと相似した特性及び課題を抱える功利主義（utilitarianism）をめぐる議論に着目し、功利主義の理論的展開、特に規則功利主義の枠組を参考としてIEの改善案を検討する。以上によって、LIS領域におけるIEの実装につながる可能性が示唆されると同時に、新たな倫理的観点の図書館実務への導入という可能性を開くことが期待される。

本論文の構成は、次のとおりである。第2章では先行研究を概観する。第3章では、IEの理論的フレームワークについて、基盤となるPIの枠組みに言及しつつ、その概要を述べる。第4章ではIEに対する既存の批判及び議論を踏まえ、応用に向けたIEの理論

上の課題の整理を行う。第5章では功利主義の概要を述べた上でIEの持つ功利主義的側面を指摘する。第6章では功利主義的手法の援用によるIEの実装のための改善案を提案する。第7章では全体のまとめを行うとともに今後の展望を示す。

なお、本論文の目的は、IEの応用に向けた課題の整理と改善提案を行うことにあるため、PI・IEの基礎理論に対する、存在論的・認識論的側面にわたる原理レベルでの検討・批判は行わない。また、同じ理由により、本研究の次の段階の課題となる個別具体的な実装に係る検討には踏み込まない。

## 2 先行研究

ここでは、LISにおけるFloridiの理論の受容、及びIEの実社会への応用に関する先行研究をそれぞれ取り上げる。

### 2.1 図書館情報学におけるFloridiの理論の受容

PI及びIEは、基礎情報学、技術倫理、哲学等の幅広い領域にわたって賛否両論の反応を引き起こしてきたが、ここではLISにおけるFloridiの理論の受容状況について述べる。LISコミュニティにおいては、概ね1950年代以降、LISの哲学的基盤をめぐる議論が継続的に行われており、LISにおけるPI及びIEの受容は、主としてこの文脈の中で行われてきた。PIの登場後、Herold [8]がいち早くPIをLISの哲学的基盤として位置づける可能性を示唆した。引き続いてLISとPI・IEとの相補的な関係を述べたFloridi [6][7]が発表されると、LIS周辺の学術雑誌で関連特集[9][10][11]が組まれるな

ど、PI及びIEはLISコミュニティに対して一定のインパクトを与えた。また、Fyffe [12], Martens [13], Bawden & Robinson [14]等、複数のLISの研究者がPI及びIEにLISの基盤哲学・基盤倫理たる可能性を見出してきた。現在の英国の標準的なLISの概説書であるBawden & Robinson [15]では、序文 (Foreword) をFloridiが執筆しているほか、各章の本文記述においてもFloridiの論考への多数の参照が行われており、同書がPI及びIEの一定の影響の下にあることが見て取れる。一方、Cornelius [16]はFloridiのLIS理解の単純性に対して、Yu & Zhang [17]はIEの内的論理に対して、LISの立場からそれぞれ批判を行っている。現在LISコミュニティにおいて、PI及びIEは一定程度認知されており、批判がありつつもLISを基礎づける哲学・倫理として一部の研究者からは強く支持されている状況にあると言える。

日本国内の状況について言えば、すでにFloridiの若下の著作の邦訳 ([18][19][20]) が刊行されているほか、情報学や哲学の分野ではPI及びIEを扱った複数の論考が確認できる。一方、LIS分野のジャーナルにおける言説に限れば、塩崎 [21]でソーシャルメディアの第三者による保存をめぐる倫理的問題に対するIEの援用可能性が検討されたほかは、Floridiについてはわずかに後藤 [22]や書評記事である根本 [23], 緑川 [24]で言及されているのみである。近年、日本図書館情報学会編 [25]において独立した見出しとして「フロリディの情報哲学」が立項されるなど、その認知は徐々に進んでいる状況にあると言えるが、PI・IEの思想的 content が広く十分に理解されている状況にあるとは言えない。

## 2.2 Floridiの情報倫理の応用に関する先行研究

IEは未だ歴史の浅い理論であるが、IEの理論的フレームワークの現実世界への応用を試みた先行研究は、Floridi自身の論考も含め、複数の領域において存在する。いくつか例を挙げるならば、デジタル社会における「コモンズの悲劇」に焦点を当てたGreco & Floridi [26], 情報の透明性をめぐる倫理的考察を行ったTurilli & Floridi [27], デジタル環境での信頼性の理論についてのTaddeo[28], peer-to-peer技術の倫理評価にIEを適用したTaddeo & Vaccaro [29], AIを使った自動意思決定システムの倫理的基盤の検討を行ったMökander et al. [30]等が、個別の課題の解決にIEの理論の援用を試みている。

一方で、LIS分野の先行研究においてIEが言及される場合、LISの哲学的基盤に関する抽象度の高い議論の中で扱われるか、あるいは逆に個別具体的な現実の課題（プライバシー、個人データ保護、AI倫理等）が主題として扱われる論文の中で副次的に触れられることがほとんどである。PIの前提を踏まえたIEの理論的代替案を提示するSpence [31]や、IEの理論と応用とを架橋しようとするYu & Zhang [17]のような中間的な議論もあるが、IEの応用可能性に関わる、より一般的な方法論につながる議論は十分に行われていない。

## 3 Floridiの情報倫理フレームワークの概要

現在、一般に「情報倫理」という用語は、分野や文脈によって異なる意味で使われている。IEは、「既存の倫理学説を現代の

情報通信技術分野に応用する、いわゆる「応用倫理」でも「情報通信技術の利用規範を開発者・利用者に順守させるための「情報リテラシー」を説くもの」でもなく（池田 [32]）、その基盤であるPIの枠組から導かれる独自の包括的な倫理理論である。したがって、IEの理解のためにはまずPIの理論的前提を把握する必要がある。そこで本章では、Floridi [2][3], Martens [13], Bawden & Robinson [14], Yu & Zhang [17]等に基づき、はじめにIEに直接関係する範囲で簡潔にPIの骨子を述べ、次にIEの概要及び倫理原則について確認する。

### 3.1 情報の哲学

#### 3.1.1 情報圏

PIは存在論の観点において「情報構造実在論」（information structural realism）の立場をとる（Floridi [2]）。ここでいう実在論とは、人間の主観的認識とは独立して事物が存在することを前提する哲学的立場を指す。その上でPIにおいては、あらゆる存在は、生物・無生物、自然物・人工物、有体物・無体物といった区別を問わず、全て「情報実体」（informational entity）または「情報オブジェクト」（informational object）としてフラットに扱われる。すなわち、ドキュメント、道端の石、北極星、一人の人間、といったあらゆる存在は等しく情報のキャリアとして理解される。そしてこのような情報実体やその相互関係の集合である総体としての世界全体は「情報圏」

（infosphere）と呼ばれる。なおこれは、世界がデジタル情報から構成されているとする還元主義的な発想に基づくものではなく、あくまで世界を情報という一つの観

点（次に述べる「抽象化レベル」）において記述した結果であるモデルを情報圏と呼んでいる点に注意が必要である。

#### 3.1.2 抽象化レベルの手法

世界を情報圏として記述する際に用いられる方法論が、PIの根幹をなす「抽象化レベルの手法」（method of levels of abstraction, 以下「LoAの手法」）である。LoAの手法とは、端的に言えば「「物事を眺めるための土俵をどのように設定するか」についての方法を提供する」（榎本 [33]）ための方法論である。例えば1台の自動車は、買い手の立場からは盗難防止装置の有無、過去のオーナー履歴といった観点で観察され、修理工の立場からはエンジンの状態といった観点で観察され、エコノミストの立場からは燃費、市場価値といった観点で観察されうる（これらの諸観点は「オブザーバブル（observable）」と呼ばれる）

（Floridi [3]）。これらの買い手・修理工・エコノミストとしての立場が、それぞれ異なる抽象化レベル（LoA）に対応すると考えれば良い（なお、LoAの形式的な定義は、解釈された型変数であるオブザーバブルの集合とされる（Floridi [34]））。異なるLoA同士は互いに疎である（つまり共通のオブザーバブルを持たない）こともあれば、オーバーラップする場合もあり、また階層構造をなすこともある（例えば「バス」「オートバイ」は、より抽象度の高いLoAである「車両」のサブクラスとみなすことができる（Siponen [35]））。

あるLoAがいったん設定されたならば、当該LoAに関係する変数のみが考慮され、それ以外の変数が捨象されたモデルが得られる。そしてPIにおいては、世界のあら

ゆる存在に等しく適用可能な最も抽象度の高いLoAが、情報のレベルのLoAであるとされる。すなわち、物理的な実体であれ、抽象的な概念（例えば個人の意識）であれ、全ての存在を包括的かつ平等に把握できる唯一の共通観点が、情報というLoAである。情報圏において存在が情報実体として把握されることの根拠はここに求められる。そして情報圏とは、LoAの手法によって情報のレベルで世界を記述したときに立ち現れる一つのモデルに他ならない。

## 3.2 情報圏における倫理

### 3.2.1 基本的特徴

以上の前提の下、Floridi は、情報圏に存在するあらゆる情報実体（前述のとおり、これは生物・非生物、自然物・人工物、有体物・無体物等を問わない）は、まさに情報実体であることによって最低限の内在的価値（intrinsic value）を持ち、そのことによって等しく尊重されなければならないと主張する。そして、あらゆる情報実体は自らが存続、改善、繁栄する権利を持つとする。情報圏の倫理であるIEは、この前提に基づいて構築される。

一般的に、標準的な規範倫理の理論は、帰結主義（またはその代表である功利主義）、義務論、徳倫理学の3種類に大別されるが（奈良 [36]）、IEはこれらの人間中心主義の倫理とは異なり、情報として捉えられた存在そのものに基礎を置く存在中心主義的（ontocentric）アプローチを採用する点で、独自のラディカルな立場をとる。既存の倫理理論の中には、IEと同様に人間非中心主義的性格の強い土地倫理（land ethic）や環境倫理学（environmental ethics）等もある。

り、IEとは一定の親和性を持つが（実際、FloridiはIEの基本的性格を、電子情報時代の環境倫理である「e-環境倫理」（environmental ethics）という独自の用語で形容している）、それらと比べた場合、IEはあらゆる存在を情報の観点から包括的かつフラットに扱う存在論的な平等主義をとる点において、より徹底したものであるということができ、既存の倫理理論とは大きく前提を異にしている。また、同様の理由によりIEは、分野依存型の倫理であるミクロ倫理（microethics）ではなく、より包括的かつ一般的なマクロ倫理（macroethics）の一形態として位置づけられている。

### 3.2.2 ペーシエント志向の倫理

情報圏において、情報実体は道徳的行為者（moral agent, 以下「エージェント」）及び道徳的被行為者（moral patient, 以下「ペーシエント」）のいずれにもなりうる。前者は情報圏内に存在する情報を利用して何らかの道徳的行為（moral action）を行う主体であり、後者は道徳的行為の受け手である（ここでいう道徳的行為とは、情報圏内における情報実体の任意の行為を何らかの倫理的な効果をもたらすプロセスとして捉えた呼称であり、日常的な用法における「道徳的な」行為であるという意味ではない）。情報実体同士の不断の相互作用が行われているこの情報圏においては、エージェントの道徳的行為によってペーシエントは何らかの影響を被り、その情報実体としての存在のあり方、さらには内在的価値に変化が生じる。そしてIEにおける倫理評価は、ペーシエントの、ひいては情報圏の変化を把握することによって行われる。なお、IEの倫理評価においては、エー

ジェントの自由意志や心的状態といったものは考慮されない。Floridi [3]が「心なしの道徳」(mindless morality)と呼んでいるように、評価対象となるのはあくまで道徳的行為の結果として生じるペーシエント及び情報圏の客観的な変化である。この意味において、IEはペーシエント志向(patient-oriented)の倫理であるとされる。

### 3.2.3 エントロピーと4つの基本原則

それでは、IEにおいては、具体的にどのような基準によって倫理評価は行われるのか。情報圏においては、個々の情報実体とその内在的価値を十全に発揮し、情報圏の繁栄(flourishing)やウェルビーイング(well-being)が促進されることが善とされる。逆に、「情報オブジェクトとして把握された実体に対する、あらゆる種類の破壊や腐敗、すなわちいかなる形であれ存在(Being)の貧困化」(Floridi [3])は「エントロピー」(entropy)と呼ばれ、IEにおいて避けるべき悪として位置づけられる(なお、エントロピーの語は熱力学や情報理論等における用法とは異なるFloridi独自の定義で用いられており、誤解を避けるために「形而上的エントロピー」(metaphysical entropy)と言い換えられることもある)。具体的にどのようなものがエントロピーとみなされるかは場合によって様々な可能性がありうる。エントロピーは端的に情報の量的な減少や消失だけを意味しているわけではなく、その属性には、例えば「一貫性」(consistency)に対する「非一貫性」(inconsistency)、「安定性」(stability)に対する「不安定性」(instability)、「信頼性」(reliability)に対する「非信頼性」(unreliability)、「利用可能性」(availability)

に対する「利用不可能性」(unavailability)等といったものも含まれる(Floridi [37])。情報圏で生じるこうしたさまざまな形の否定的な状態の総称(すなわち前述の「存在の貧困化」)が、IEにおけるエントロピー概念である。

その上で、Floridi [3]はIEにおける最も基本的な原則として、エントロピー概念を軸とした次の4つの基本原則を立てる。

0. 情報圏においてエントロピーは引き起こされるべきではない。
1. 情報圏においてエントロピーは予防されるべきである。
2. エントロピーは情報圏から排除されるべきである。
3. 情報実体及び情報圏全体の繁栄は、そのウェルビーイングを保ち、育成し、より豊かにすることにより、促進されるべきである。

上記の4原則は0.が最も基本的なレベルであり、数字が大きくなるに従い、より高度かつ望ましい状態に近づいていくとされる。すなわちエントロピーを発生させないことが最低限の要求であり、最終的な目的は、個々の情報実体及び情報圏の繁栄やウェルビーイングの実現に置かれる。そしてIEにおける倫理的評価の基盤はこれらの基本原則に置かれることになるが、情報圏においてはあらゆる存在が情報実体として最低限の内在的価値を持つとされていること、すなわち情報実体の持つ内在的価値が定義上ゼロを下回ることではないことを踏まえれば、IEにおいて倫理的評価の主たる対象となるのは、あくまで道徳的行為であることに注意が必要である。すなわち、

あるエージェントの道徳的行為の結果として生じるエントロピーの増減を把握することで、当該道徳的行為の倫理的評価が行われることになる。

なお、上記の基本原則のうち第4原則(3.)の表現には文献によって若干の異同がある。比較的初期のIEの文献においては「3. 情報の福利は、情報圏を拡大し(情報の量)、改善し(情報の質)、豊かにする(情報の多様性)ことによって、促進されるべきである。」となっているが(Floridi [37])、これは表現をより包括的な形に洗練させたことによる相違であり、趣旨は同一と考えられる。ここでは、情報実体や情報圏の内在的価値の把握に際しては、単純に情報の量のみではなく、質や多様性といった要素もあわせて考慮に入れるべきであるという点を読み取る必要がある。

### 3.3 小括

IEの新規性は、情報実体のもつ内在的価値とエントロピーの概念を措定することで、人間中心主義を脱した存在中心主義に立脚している点に求められると言える。IEは、世界のあらゆる存在を等しく情報の抽象化レベル(LoA)から捉え直した倫理理論として、前述のとおりドキュメントやデータといった情報オブジェクト(物理的な資料であれデジタル形式であれ)を主たる対象とするLISにとって、規範倫理として有効性を持ちうる可能性があると考えられる。続いて次章では、IEの実装に向けた理論面の課題の抽出を行う。

## 4 情報倫理の応用に関する批判及び実装に向けた課題

これまでIEには様々な批判が行われてきた。PIも含めた存在論的、認識論的レベルでの主要な批判は、情報実体の内在的価値に関わる批判、あらゆる情報実体に価値を認める平等主義に対する批判、道徳的判断における相対主義に対する批判の3類型に大別できる(Yu & Zhang [17])。しかし応用的側面を論じる本論文では、IEの基礎理論に係るこれらの原理的な批判は扱わないこととし、その理論体系をいったん所与のものとして扱う。本章では応用の観点に絞った主な批判の論点を確認し、その上で解決すべきIEの理論上の課題を整理する。

### 4.1 応用の観点における主要な批判

IEの一般的な応用可能性に関して、Doyle [38]は、トロツキ問題(複数の人間を助けるために特定の人物を犠牲にできるかという論点をめぐる倫理学上の思考実験)のような倫理的判断を求められる個々のケースに対してIEの枠組では明確な結論を導くことが困難である点を指摘し、IEの倫理原則に基づく善悪の判断において恣意性が介入する可能性を批判している。またStahl [39]は、IEの枠組がプライバシーや人工エージェント(人工知能)の倫理的問題に貢献できる可能性を述べつつも、IEは純理論的性格の強さから、直接的な応用が意図されていない点を指摘している。

また、情報実体のもつ内在的価値の計測・比較に関する方法論の欠如を複数の論者が批判しており、例えばBrey [40]は、IEの存在論的平等主義からは「シェークスピアの作品と三文小説、人間存在と水槽一杯の有毒廃棄物がそれぞれ同程度の価値があることになってしまう」という望ましくない帰結に繋がりを指摘している。

る。また, Siponen [35]は, IE におけるエントロピーは情報理論を援用すれば計量可能であるという Floridi の主張に対して疑義を呈している。

#### 4.2 実装に向けて解決すべき課題

先行研究における批判を踏まえ, IE の実践的な応用にあたって解決すべき課題を次の3点に整理した。

第一に, IE の4つの基本原則は抽象度が高く, 実践とは乖離が存在する点が指摘できる。例えば IE の倫理原則を個々の業務行為の評価や政策決定といった局面において直接適用した場合, 具体的な指標や指針がないために, 必ずしも有用な判断を引き出せるとは想定できない。IE の倫理原則の運用にあたっては, 具体的に適用可能な粒度・形式に落とし込むための作業が別途必要になると考えられる。

第二に, 定量的な比較の問題がある。Floridi [3]は, 倫理的問題の解決においては量的・アルゴリズム的文脈と質的なそれとの間には断絶があるとし, IE は基本的には質の評価のための倫理であると述べている。しかし, これは IE の高度に理念的な性質及びそのことによる IE 単体での直接適用の困難性を踏まえて述べられた言明と見なすのが妥当と考えられ, 現実的な IE の倫理評価は, 情報実体を持つ内在的価値の加算・減算からなる, ある種の計算のアナロジーで発想されており, その最大化が目指されていると見なすのが自然である(情報の非存在や減少が IE におけるエントロピーである以上, 情報の存在や増加はその逆の価値を有するはずである)。そのため, 現実に行うには情報実体の内在的価値及びその変化の定量的な把握が必要

となるが, その計測・計算のための具体的な尺度や手続きを IE は提示していない。

第三に, IE で善とされる情報圏の繁栄やウェルビーイング, 悪であるエントロピーといった概念の意味する具体的な内容は必ずしも明確でない。換言すれば, 現実世界におけるどのような行為や事象がそれらに該当するのかが往々にして曖昧である。例えば, 前述のとおり, 情報圏における倫理評価にあたっては情報の量に加えて質や多様性といった要素も同時に考慮すべきであるとされるが, これら情報の量, 質, 多様性のうちどれをどの程度重視すべきなのか, 相互にトレードオフの関係にあった場合はどの要素が優先されるべきなのか, といった点について Floridi は明確な基準を示していない。例えば, Web サーバの破壊のような情報実体の物理的損失の場合であれば, それがエントロピー(情報圏における悪)の一樣態であることは一応自明と考えられるが, 同じ予算で 2,000 ページからなる 1 冊の辞書をデジタル化することと, 様々な分野から選んだ 150 ページの図書を 10 冊デジタル化することを比べた場合, 単純に情報量(ページ数)の多寡の観点から前者を優先すべきなのか, あるいは多様性の観点から後者の方がより望ましいのか, といった点に関しての結論は自明ではない。

上記の3つの課題はいずれも IE の純理論的な性格に由来しており, そのため IE の倫理原則を直接適用する形で倫理評価を行う場合, 恣意性や主観的判断に依拠する程度が必然的に大きいものとなる。ここにおいて, 実装にあたって過度に恣意的・主観的にならず一定の客観性を担保するとすれば, 何らかの概念的な装置を導入する

ことで、IE の理論的フレームワークを運用に耐える形で補強しなければならない。次章以下では、その解決に向けて、IE の持つ功利主義的側面を指摘し、それを踏まえて IE を実装するための具体的な検討を行う。

## 5 功利主義と情報倫理

IE を既存の主要な倫理理論と比較した場合、IE はあまねく情報実体に対する尊重を共通の義務とみなす義務論的な特性と、道徳的行為の結果に倫理評価の基礎を置く功利主義的な特性の両面を持つ (Yu & Zhang [17])。本論文ではそのうち特に後者に着目し、功利主義に対する批判及びそれに対する功利主義側の理論的応答を参照事例として検討することで、IE の応用上の課題の解決に関する可能性を探る。

### 5.1 功利主義とは

#### 5.1.1 功利主義の概要

功利主義とは「行為の正・不正を判断する際の究極的な判断材料となるのはその行為の結果（帰結）のみである」（伊勢田 [41]）とする帰結主義の、代表的な形態である。歴史的に功利主義には非常に多様なバリエーションが存在するが、その最も古典的な形態は「最大多数の最大幸福」のモットーで知られる Jeremy Bentham の思想に求められる。古典的功利主義においては、人々の幸福のみに内在的価値が存在するとされ、幸福を増進させるものが善とされる。功利主義において何が幸福を増進させるのかについての見解は論者によって異なるが、例えば Bentham の功利主義においては、人々の快楽（pleasure）を幸福の源泉とみなし、社会におけるその総量の最大化

（逆に言えば苦痛（pain）の最小化）が究極的な目的とされる。現代の功利主義においては、快楽ではなく選好（preference）に基礎を置く選好充足型功利主義が主流となっているが、これはあくまで何を幸福とみなすかについての立場の違いにすぎず、功利主義が社会の幸福の総和の最大化を図る倫理理論であることに変わりはない。19 世紀以降、功利主義は様々な批判に応える形で理論の多様化が進むが、本論文においては「功利の原理、すなわち関係者の幸福を最大化するような選択を行うべしという第一原理をベースとしたいろいろな立場」（伊勢田 [41]）を功利主義と総称する。

功利主義は、総和主義を一つの特徴とする。これは、「関係者の効用は加算可能であり、複数の帰結の良し悪しは、関係者が享受する効用の総和を大きい順に並べることで判定できるという考え方」（水野 [42]）である。古典的功利主義においては、例えばある 2 つの行為（または制度、政策等、以下同様）の比較は、それを実行することによって得られる関係者の幸福を単純加算（つまり重みづけを行わない加算）して得られる総和を用いて、社会全体の幸福度を比較することによって行われる。この操作は功利計算 (utility calculus) と呼ばれる。

#### 5.1.2 功利主義に対する主な批判

功利主義に対しては、歴史的に様々な批判が向けられてきたが、ここでは 2 つの主な論点を挙げる。

第一に、功利主義の原則に忠実に従うと直観に反する結論が往々にして導かれることが、繰り返し指摘されてきた（伊勢田 [43]）。例えば、功利主義の原則からすれば、社会全体の幸福度が増す行為であれば全

体のために誰かを犠牲にする（健康な一人の人間を犠牲にしてその臓器を多数の病人に分配・移植する、位置が不明な爆弾を仕掛けたテロリストに非人道的な拷問を行って自白を引き出す等）といった行為が推奨されてしまうが、これらは一般的な社会の支配的な道徳や慣習と齟齬をきたす行為であり、少なくとも異論なく直観的に許容できるものであるとは言えない。

第二に、功利計算に対しては、幸福（快楽・選好）の定量的な測定・比較可能性が議論の対象となってきた（水野 [42]）。すなわち、そもそも幸福は定量的に測ることができるものなのかという問題や、異なる個人間（*interperson*）の幸福と同一個人内（*intraperson*）における異なる選好とを、同一尺度で定量的に比較することはそもそも可能なのかという点に関する批判である。

### 5.1.3 批判に対する功利主義側の応答

このような批判に対応する形で功利主義は次のとおり理論の精緻化・多様化を行ってきた。

第一の批判に係る問題の原因としては、主に2点を挙げることができる。1つ目の原因は、Bentham や Mill の古典的功利主義に見られる、個別の行為に功利原則を直接適用する（これを行為功利主義（*act utilitarianism*）という）点に求められ、それに対する功利主義側の対応が、20世紀後半に明示的に理論化が進められた規則功利主義（*rule utilitarianism*）の立場である。規則功利主義とは、「ある一般的規則の体系が行為の指針として採用された場合に生じる帰結の含む厚生（well-being）の総体を考慮の対象とし規則の評価を下した上で、行為自体の

評価を直接に功利計算によるのではなく、当該行為に対応するルールに従っているか否かで決定する」（安藤 [44]）立場である。つまり、規則功利主義では、直観や既存の道徳規範に適合的であり、かつその規則に従っていれば総体として社会の幸福が増進されるような二次的規則が定められる

（従って個別の局面においては当該行為によって幸福が一時的に減少する事態が起こりうることも許容する）。ここでは、二次的規則は一次的規則である功利原則に基づいて定められ、かつ評価されることになる。そして、行為は二次的規則に基づき、二次的規則は功利原則に基づく、という階層構造の中で、功利計算は個々の行為ではなく二次的規則に対して適用される。前述の例を再び用いるならば、ある社会が「いかなる場合でも拷問を行ってはならない」という二次的規則を採用したならば、それは、（例えば爆弾の位置を特定できないといったデメリットが個別に生じる可能性があるとしても）総体としてはそのような規則を持たない社会よりも長期的にはより多くの福利がもたらされると考えられるためである。もっとも、規則功利主義自体にも多様な立場が含まれるが（Miller [45]）、それらに共通する基本的な考え方は上記のとおりであり、それによって行為功利主義がもたらす問題の解決が図られている。

第一の問題の2つ目の原因は、功利主義の総和主義に求められる、社会全体の幸福量の多寡のみを測るという発想のもとでは各人の幸福は足し合わされて全体に溶け込んでしまうため、その内部における公平性や公正な分配については考慮されないという結果をもたらす（伊勢田[43]）。総

和としての幸福が大きければ、例えば一部の特権階級のために奴隷制を容認するような結論や、幸福な人が少数いる社会よりも不幸な人が多数いる社会の方が望ましいといった結論が導かれうるのである。これは単純な総和主義をとる場合に原理上避けられない難点であるが、この批判に対して功利主義が取りうる主たる対応の一つが規則功利主義の二次的規則において分配的正義を考慮する方法である (Freeman [46])。換言すれば、分配的正義が重視される規則が採用された社会は、そうでない社会よりも幸福の総和が大きくなる、という前提を置いた規則功利主義の立場を取るという戦略である。

第二の定量性に関する批判については、そもそも功利計算は多分に理念的な性格が強い概念装置であり、主唱者の Bentham 自身も含め実際に算術的に功利計算を実行する功利主義者は少数である (児玉 [47]) 点に留意する必要がある。ただし社会実装においては、実際に功利計算に準じた手法が援用される例があり、例えば医療資源の分配をめぐる政策判断において使われる質調整生存年 (Quality-adjusted life year, QALY) (浅井 [48]) の考え方は功利主義と親和性が高いという指摘がある (児玉 [49]) ほか、災害救助時に多数の傷病者の優先順位づけのために行われるトリアージ (東京都保健医療局 [50]) は、時間及び医療資源の制約の中で、医療資源の浪費を避けつつ最大人数の救命を目的とする功利主義的な手法であると理解されている。このように、公衆衛生や災害医療のような領域においては、功利主義的発想に基づく定量的アプローチが採用され、一定の効果を挙げている実例が存在する。

## 5.2 情報倫理の功利主義的側面

IE と功利主義との関係については、Doyle [38]が、IE の恣意性を批判する文脈において両者の対比を行っているほか、その類似点や相違点について、Siponen [35]が考察を行っている。Floridi [3]は、IE において「エントロピーの抑制と情報圏 (の一部) の積極的な保護・増進は最大効用に寄与する」、さらに「行為は、それが (存在論的) 情報を増加させ、かつ (形而上的) エントロピーを減少させる度合いに比例して、善である」と IE を功利主義的に説明することで、両者の間に一定の共通する性格があることを認める。一方で、IE と (功利主義を代表とする) 帰結主義の間に次の 4 つの相違点があることを指摘している。1 点目は「単調性の問題」 (monotonicity problem) であり、善悪の判断に関する限り IE は、その倫理原則に反する行為を正当化しない点において、結果に基づく帰結主義よりも、アプリアリに善悪の判断原理を与える義務論により近い性格を持つという。2 点目は「数学的問題」 (mathematical problem) であり、功利計算の実行可能性に対する疑義が呈される。3 点目は「義務を超える行為の問題」 (supererogatory problem) である。功利主義の原則は行為者に (自己犠牲を含む) 過大な要求を課すことにつながりうるという批判がしばしば行われるが、IE においては情報圏の状態を少しでも改善する行為は善とされるため、この問題は生じないとする。そして 4 点目は、「比較の問題」 (comparative problem) であり、社会全体の総体としての幸福量の増加を評価する帰結主義に対して、IE は行為の影響の及ぶローカルな領域を評価するという違い

があるとする。

上記の 4 点のうち IE と功利主義の原理的な相違に関わるのは 1 点目の「単調性の問題」であり、ここで Floridi は IE の義務論的性格を主張している。しかし、現実には IE に基づく倫理評価を行う場合を考えると、ある行為が情報圏に対してもたらす結果（すなわちエントロピーの増減）は必ずしも予見可能ではなく、かつ結果は所与の条件次第で変わりうると考えられるため、あらゆる行為の善悪をアプリオリに定めることは困難である。この場合、IE における行為の倫理評価は、義務論的ではなく功利主義（帰結主義）的に行う他ない。したがって IE の実装にあたっては、行為のトリガーとなる自由意思や動機といった要素を捨象した上で結果としての幸福／エントロピーの増減に基づいて倫理評価を行う功利主義（帰結主義）的側面を考慮することが有効であると考えられる。そして、前述のとおり功利主義の問題の一部の解決が規則功利主義の理論化によって図られた経緯があることを踏まえれば、規則功利主義の手法を援用することによって、IE の実装に係る問題点の解決を図る可能性を考えることができる。その方針の下、次章において具体的な検討を行う。

## 6 功利主義的手法の導入による情報倫理の応用に向けた検討

4.2 において、IE の実装に向けた理論的課題を、①基本原則の直接適用の困難性、②定量的把握の困難性、③善悪概念の具体的内容の曖昧さ、の 3 点に整理した。前章で確認した功利主義の理論的展開を踏まえ、以下のとおり IE の実装に向けた提案を

行う。

第一に、基本原則の直接適用の困難性の問題については、規則功利主義の枠組の援用が可能であると考えられる。具体的には、まず IE の倫理原則をあくまで二次的規則が準拠するための一次的規則として位置づける。その上で、個別の領域ないしケースに応じた形で、運用可能な二次的規則を定め、実務の遂行は、当該二次的規則に準拠するものとする。ここにおいて、IE の倫理原則が適用される対象、すなわち倫理評価の対象は、二次的規則となる。ここでの二次的規則は、例えば法規、事業計画、事業目標、評価指標等の様々な形をとることが考えられ、個々の情報実体とその情動的価値を発揮するような内容であること（一例としては、合理的なオープンアクセスポリシーを定め、現実的に適正な範囲で資料の利用可能性を高める等）が求められることになる。このような体制を取ることによって、直接適用の困難性の問題を解決するとともに、全体として IE の倫理原則に立脚した、すなわち情報圏の繁栄に資する実践的活動とその評価が可能となる。

この点に関連して、Floridi [3]は「IE は特定の倫理上の問題の解決において直接的に有用なものではない」が、「課題解決の手続きを導く概念的基盤を与えるものである」と述べている。また Taddeo [51]は、IE の理論的インフラはケースバイケースのアドホックな分析のために存在するのではなく、情報環境のデザイン、形成、及び情報環境内における生存を支援するための「倫理的フレームワーク」を構築することを可能にするものである、と主張している。すなわち、個々の現実の課題への IE の応用に際しては、IE の倫理原則に基づく形

で、個別の事情に合わせた倫理評価の枠組を別途構築することが必要とされていると考えられる。これはまさに規則功利主義の発想と親和的であり、IE に立脚した二次的規則を IE の基本原則にいわば連結する形で設け、運用可能な倫理フレームワークとして機能させることは、本来想定されている IE の応用方法として適合的であると言える。

第二に、定量的把握の困難性の問題については、規則功利主義の二次的規則に定量的観点を（少なくとも部分的に）加えることによって解決を図ることが可能であると考えられる。4.2 で述べたとおり、情報圏の「繁栄」や「ウェルビーイング」の含意の中には、情報実体の内在的価値やエントロピーの大小に関わる定量的要素も（たとえば理念的なものであれ）考慮されることを踏まえると、実装の段階において、IE の原則をより具体化する形で何らかの定量的な評価方法を導入することによって、これらの評価が可能となると考えられる。また、倫理評価において恣意性の介入を可能な限り減らすことを前提とするならば、客観的に測定可能な指標を取り入れることはその実現のための一つの方法となりうる。なお、情報量の具体的な計測方法や、特に何をエントロピーとみなして定量的に評価するかという点に関しては、具体的な実装の局面における条件にしたがって個別に検討すべき問題であろう。

第三に、善悪概念の曖昧さ、すなわち情報圏におけるエントロピーの減少とウェルビーイングの増進に係る判断基準に関わる問題については、3.2.3 で述べたとおり Floridi [37]によって例示の形でエントロピーの類型は示されているものの、実装可能

な形でより具体的かつ一般的な定義を行うためには、さらなる検討が必要である。そこで、本論文ではその部分的解決の一つとして、情報の総和だけではなく、多様性を担保するために、当該規則には分配的観点も含めることを提案する。ここでいう多様性とは、情報圏において、一部の情報実体だけでなく多数の情報実体が繁栄している状態、すなわち情報圏を構成する各情報実体が内在的価値を十全に発揮していることと解する。なおこれは功利主義の文脈で言えば、5.1.3 で述べた総和主義で起こりうる内部の不公平性の問題への対応に相当する。一例を挙げれば、あるデジタルアーカイブを構成するコンテンツのうち、少数のものだけに手厚いメタデータや高解像度の画像データ等を作成する（すなわち、より大きな内在的価値を発生させる）一方で、大半のコンテンツには簡易なメタデータとサムネイルしかない（すなわち、比較的貧しい内在的価値しか付与しない）状況を想定した場合、当該デジタルアーカイブ内では内在的価値の偏在が起きていると見なすことができる。ここで仮に、計 10 個のコンテンツのうち 2 コンテンツがそれぞれ 15 の内在的価値を持ち、残り 8 コンテンツが各 2.5 の内在的価値を持つとすれば、その総計は 50 となる。一方、10 コンテンツが仮に全て 5 ずつの内在的価値を持つように等しくコンテンツ整備を行った場合、その総計は同じく 50 となる。結果として両者の総計は等しくなるが、少数の情報実体に内在的価値が集中している前者と比較して、後者では多数の情報実体がそれぞれ一定程度の繁栄を享受している、すなわち内在的価値を発揮していると言え、より高い多様性が達成されているとみ

なしうる。もちろんこれはあくまで、思考実験的な例示にとどまるものであり、IEにおける多様性のより一般的な定義づけはこれからの課題となるが、多様性の担保の一つの可能な様態としてはこのようなあり方を想定することができる。

以上のように、包括的なIEの基本原則に基盤を置いた上で、定量的および分配的観点を含む二次的な規則を導入する、いわば2階建ての倫理評価を行う形で、IEの倫理原則の実装上の課題の解決に繋げることが可能であると考えられる。あくまでIEの倫理原則は二次的規則を支え評価するものとして位置づけ、実際の評価対象となるのは二次的規則である。

## 7 結論

本研究では、まずIEの実装に係る理論的課題を、①基本原則の直接適用の困難性、②定量的把握の困難性、③善悪概念の具体的内容の曖昧さ、の3点に特定した。次いで、規則功利主義の発想に基づく手法の導入を、上記課題の解決のための可能な方法として提案した。またその際、倫理評価の対象とする各応用領域で定める二次的規則（法規、計画、評価指標等）では、エントロピーを減少させ情報圏の繁栄に貢献するような、定量的観点（「量」の増加）かつ分配的観点（「多様性」の担保）も含めるべきであることを示した。そのことによって、恣意性を可能な限り減らし、運用可能かつ実効性のある形でIEを図書館等の業務に実装できる可能性を提示した。

一方で、本研究の限界は次の2点である。第一に、情報の多様性に関するさらなる検討が必要であり、かつ質に関する問題が未

検討である。特に、情報圏のウェルビーイングに資する情報の質の概念についてどのような要件が必要か、意味論的情報の正しさ（真偽）の問題も含めて、さらなる理論面での整理が必要と考えられる。第二に、IEの基本原則に立脚した二次的規則として具体的にどのような要件（何を定量的評価の対象とするのか等）が求められるのかについての考察が不十分である。ただしこの点については、具体的な応用段階において、個別のケースごとに明らかにすべきものであると考えられる。

本研究では、IEの実装に向けた最初の段階として、理論的枠組の整備にフォーカスし、主として理念的な側面における議論を行った。今後の課題としては、本研究で提示した枠組を具体的な応用領域において適用し、提案手法の有効性について検証を行う必要がある。また、その作業を通じて、二次的規則の要件についての検討を行うとともに、情報の多様性及び質の問題に関する検討も併せて行うことが求められよう。

なお、Floridi [52]は、自らの理論が説得性を持って受け入れられるためには、現実世界への適用可能性と、日常生活や経験における有用性を実際に示すことが必要であることを認めている。IEに基づく倫理評価が具体的な効用をもたらすことが明らかになれば、応用面の議論を超えて、LISの基盤哲学・倫理をめぐる問題にも、有用性の観点から一石を投じることになるだろう。

## 8 データ利用可能性宣言

本研究ではデータセットが作成・分析さ

れていないため、共有すべきデータは存在しない。

## 謝辞

本論文の執筆にあたっては、九州大学大学院統合新領域学府ライブラリーサイエンス専攻の石山栄美教授に多大な御指導を賜りました。ここに深く感謝申し上げます。

## 参考文献

- [1] Floridi, Luciano: “The 4th Revolution: How the Infosphere is Reshaping Human Reality”, Oxford University Press, 248p., 2014.
- [2] Floridi, Luciano: “The Philosophy of Information”, Oxford University Press, 405p., 2011.
- [3] Floridi, Luciano: “The Ethics of Information”, Oxford University Press, 357p., 2013.
- [4] Floridi, Luciano: “The Logic of Information: A Theory of Philosophy as Conceptual Design”, Oxford University Press, 240p., 2019.
- [5] Floridi, Luciano: “What Is the Philosophy of Information?”, *Metaphilosophy*, vol. 33, no. 1/2, pp. 123–45, 2002.
- [6] Floridi, Luciano: “On Defining Library and Information Science as Applied Philosophy of Information”, *Social Epistemology*, vol. 16, no. 1, pp. 37–49, 2002.
- [7] Floridi, Luciano: “Afterword: LIS as Applied Philosophy of Information: A Reappraisal”, *Library Trends*, vol. 52, no. 3, pp. 658–65, 2004.
- [8] Herold, Ken: “Librarianship and the Philosophy of Information”, *Library*

*Philosophy and Practice*, vol. 3, no. 2, 2001.

[9] *Library Trends*, vol. 52, no. 3, 2004.

[10] *Ethics and Information Technology*, vol. 10, no. 2–3, 2008.

[11] *The Information Society*, vol. 25, no. 3, 2009.

[12] Fyffe, Richard: “The Value of Information: Normativity, Epistemology, and LIS in Luciano Floridi”, *Portal: Libraries and the Academy*, vol. 15, no. 2, pp. 267–86, 2015.

[13] Martens, Betsy Van Der Veer: “An Illustrated Introduction to the Infosphere”, *Library Trends*, vol. 63, no. 3, pp. 317–61, 2015.

[14] Bawden, David; Robinson, Lyn: “Curating the Infosphere: Luciano Floridi’s Philosophy of Information as the Foundation for Library and Information Science”, *Journal of Documentation*, vol. 74, no. 1, pp. 2–17, 2018.

[15] Bawden, David; Robinson, Lyn: *Introduction to Information Science*, Second edition, Facet Publishing, 386p., 2022.

[16] Cornelius, Ian: “Information and Its Philosophy”, *Library Trends*, vol. 52, no. 3, pp. 377–86, 2004.

[17] Yu, Liangzhi; Zhang, Yao: “In Search of a Coherent Theoretical Foundation for LIS Ethical Principles: An Appraisal of Floridi’s Information Ethics”, *Journal of Documentation*, vol. 80, no. 4, pp. 730–48, 2024.

[18] フロリディ, ルチアーノ (西垣通訳): 『情報倫理の本質と範囲』, 『情報倫理の思想』, NTT 出版, pp. 47–98, 2007.

[19] フロリディ, ルチアーノ (春木良且, 犬束敦史監訳, 先端社会科学技術研究所訳): 『第四の革命: 情報圏が現実をつくりかえる』, 新曜社, 384p., 2017.

[20] フロリディ, ルチアーノ (塩崎亮, 河

島茂生訳)：「情報の哲学のために：データから情報倫理まで」，勁草書房，216p., 2021.

[21] 塩崎亮：「公開サイトを第三者がアーカイブすることに対する個人の意識：質問票調査の集計結果」，図書館界，vol. 73, no. 1, pp. 2-14, 2021.

[22] 後藤嘉宏：「納本制度の思想と中井正一」，情報の科学と技術，vol. 57, no. 11, pp. 519-25, 2007.

[23] 根本彰：『『図書館情報学概論』』，日本図書館情報学会誌，vol. 66, no. 2, pp. 69-70, 2020.

[24] 緑川信之：『『情報の哲学のために：データから情報倫理まで』』，日本図書館情報学会誌，vol. 68, no. 2, pp. 135-36, 2022.

[25] 日本図書館情報学会編：「図書館情報学事典」，丸善出版，754p., 2023.

[26] Greco, Gian Maria; Floridi, Luciano: “The Tragedy of the Digital Commons”, Ethics and Information Technology, vol. 6, no. 2, pp. 73-81, 2004.

[27] Turilli, Matteo; Floridi, Luciano: “The Ethics of Information Transparency”, Ethics and Information Technology, vol. 11, no. 2, pp. 105-12, 2009.

[28] Taddeo, Mariarosaria: “Defining Trust and E-Trust: From Old Theories to New Problems”, International Journal of Technology and Human Interaction, vol. 5, no. 2, pp. 23-35, 2009.

[29] Taddeo, Mariarosaria; Vaccaro, Antonino: “Analyzing Peer-to-Peer Technology Using Information Ethics”, The Information Society, vol. 27, no. 2, pp. 105-12, 2011.

[30] Mökander, Jakob, et al.: “Ethics-Based Auditing of Automated Decision-Making Systems: Nature, Scope, and Limitations”, Science and Engineering Ethics, vol. 27, no. 4,

2021.

[31] Spence, Edward: “Luciano Floridi’s Metaphysical Theory of Information Ethics: A Critical Appraisal and an Alternative Neo-Gewirthian Information Ethics”, International Journal of Technology and Human Interaction, vol. 6, no. 1, pp. 1-14, 2010.

[32] 池田準：「フロリディの「情報倫理」における自立性と責任」，モラリア，no. 23, pp. 34-53, 2016.

[33] 榎本啄杜：「抽象化レベルの手法における2つの「抽象化」をデータベース的に理解する」，関西大学哲学，vol. 41, pp. 24-53, 2023.

[34] Floridi, Luciano: “The Method of Levels of Abstraction”, Minds and Machines, vol. 18, no. 3, pp. 303-29, 2008.

[35] Siponen, Mikko: “A Pragmatic Evaluation of the Theory of Information Ethics”, Ethics and Information Technology, vol. 6, no. 4, pp. 279-90, 2004.

[36] 奈良雅俊：「倫理理論」，『入門・倫理学』，勁草書房，pp. 27-50, 2018.

[37] Floridi, Luciano: “Information Ethics: On the Philosophical Foundation of Computer Ethics”, Ethics and Information Technology, vol. 1, no. 1, pp. 37-56, 1999.

[38] Doyle, Tony: “A Critique of Information Ethics”, Knowledge, Technology & Policy, vol. 23, no. 1, pp. 163-75, 2010.

[39] Stahl, Bernd Carsten: “Discourses on Information Ethics: The Claim to Universality”, Ethics and Information Technology, vol. 10, no. 2-3, pp. 97-108, 2008.

[40] Brey, Philip: “Do We Have Moral Duties towards Information Objects?”, Ethics and Information Technology, vol. 10, no. 2-3, pp.

109–14, 2008.

[41] 伊勢田哲治: 「功利主義とはいかなる立場か」, 『生命倫理学と功利主義』, ナカニシヤ出版, pp. 3–25, 2006.

[42] 水野俊誠: 「功利主義」, 『入門・倫理学』, 勁草書房, pp.91-104., 2018.

[43] 伊勢田哲治: 「功利主義をめぐる論争」, 『生命倫理学と功利主義』, ナカニシヤ出版, pp. 26–47, 2006.

[44] 安藤馨: 「統治と功利: 功利主義リベラリズムの擁護」, 勁草書房, 315p., 2007.

[45] Miller, Dale E.: “Rule Utilitarianism”, *The Cambridge Companion to Utilitarianism*, Cambridge University Press, pp. 146–65, 2014.

[46] Freeman, Samuel: “Utilitarianism, Deontology, and the Priority of Right”, *Philosophy & Public Affairs*, vol. 23, no. 4, pp. 313–49, 1994.

[47] 児玉聡: 「実践・倫理学: 現代の問題を考えるために」, 勁草書房, 257p., 2020.

[48] 浅井篤: 「QALY と医療資源配分」, 『生命倫理学と功利主義』, ナカニシヤ出版, pp. 193–217, 2006.

[49] 児玉聡: 「医療資源の配分」, 『入門・医療倫理. 1』, 勁草書房, pp. 333–49, 2017.

[50] 東京都保健医療局: 「トリアージ」, <https://www.hokeniryo.metro.tokyo.lg.jp/iryo/kyuukyuu/saigai/triage.html> (2024 年 10 月 7 日参照) .

[51] Taddeo, Mariarosaria: “The Moral Value of Information and Information Ethics”, *The Routledge Handbook of Philosophy of Information*, pp. 361–74, 2016.

[52] Floridi, Luciano: “A Defence of Informational Structural Realism”, *Synthese*, vol. 161, no. 2, pp. 219–53, 2008.

(2024年10月17日 受付)

(2025年 4月10日 採択)

(2025年 5月23日 J-STAGE早期公開)

オープンサイエンス・オープンデータ研究部会報告

## JAPAN OPEN SCIENCE SUMMIT 2025開催報告

### Report on JAPAN OPEN SCIENCE SUMMIT 2025

情報知識学会オープンサイエンス・オープンデータ研究部会編  
JSIK, Special Interest Group on Open Science and Open Data, ed.

E-mail: sig\_os-od@ml.kanazawa-u.ac.jp

#### 1 はじめに（芦野俊宏、東洋大学）

今年度から公的資金を得た研究の成果発表論文については即時のオープンアクセスが義務化された。また本年度は多くの学術機関にとってデータポリシー策定の期限とされている年度でもあり、多くの研究分野においてオープンサイエンスへの取り組みを具体化することが求められている。

このような中で、今回のJOSS2025では協賛セミナーを含めて30のセッションが開催され、平日5日間毎日6つのセッションが行われるというスケジュールとなった。今回の特徴はやはり、AIに関わる開発の開発が劇的に進む中で、AIの活用、AIとデータに関わるセッションが複数あったことであろう。FAIR原則（Findable, Accessible, Interoperable Reuseable）の提唱者の1人であるBarend Monsが数年前にFAIRを”Fully AI-Ready”とも述べていたが、これが重要性を増しつつある。

AIは学術研究にも大きな進歩をもたらすものと期待されるが、一方でデータポリシー、

データ倫理の議論にも多大な影響を及ぼしている。国際的には、700以上の機関が加盟するCoARA(Coalition for Advancing Research Assessment)にもワーキンググループERIP, Research Integrity Policy in Responsible Research Assessment for Data and Artificial Intelligenceが設けられるなど、今後ともデータ活用において議論が必要となるものと考えられる。

AIに関する議論は今回のJOSSにおいて新たに見られた話題であるが、やはり主要な話題は研究データマネジメントに関わる経験の共有やインフラに関わるものであった。

ほとんどのセッションがオンラインにて開催されたため、参加者数の算定は難しいが、zoomによる参加者は延べ2000程度、youtubeでの視聴(リアルタイム、動画視聴含む)もほぼ同数となった。30のセッションのうち9つについてセッション報告原稿が寄せられ、個人発表2件についてはいずれも予稿が投稿された。これらを部会報告として掲載する。なお、本報告の作成にあたって部会のメンバーである林正治氏、高田良宏氏、堀井洋氏のご協力を頂いたことを付記しておく

## **2 B4 JST-CHORUS Forum: Strategies to Ensure Transparency and Digital Tracking of Research Outputs (Mark Robertson, CHOR Inc.)**

### **2.1 Overview**

This JST/CHORUS Forum explored the evolving open science requirements for researchers. Under Japan's new open science policy, researchers seeking certain national awards must make their publications and research data openly accessible. Globally, funders are implementing similar mandates. To uphold the integrity of the scholarly record, researchers must confirm adherence to research protocols and disclose funding sources and grant details. During submission and publication, they must also include links and supporting data to ensure accessibility and transparency. This forum discussed strategies during the publication process to encourage compliance to ensure accuracy and to enable digital tracking of outputs.

#### **Panelists**

(Moderator) Kazuhiro Hayashi, Principal Senior Fellow and Director, Research Unit for Data Application, NISTEP

Haruki Nakamura, Professor Emeritus, Institute for Protein Research, Osaka University and EiC *Biophysics and Physicobiology*

Julie Glover, Executive Director, National Health and Medical Research Council, Australia (NHMRC)

Amanda French, ROR Technical Community Manager, Crossref

Tetsushi Ueta, Vice-Director and Professor,

Center for Administration of Information Technology, Tokushima University and EiC *Nonlinear Theory and Its Applications (NOLTA)*, IEICE

Jeff Lewandowski, Director of Publishing, American Physical Society

### **2.2 Presentations**

Haruki Nakamura began by considering open data and science from the viewpoint of all stakeholders. He outlined the introduction of research in computational science in the 2000s to data-driven science in the 2010s, to AI-driven science now. For researchers he commented on some reluctance to pay APCs for open access, the current evaluation by publication metrics, and emphasized that research data must be open. For funders, universities, research institutes, and publishers he highlighted how research is evaluated, using as an example his own journal *Biophysics and Physicobiology*. Nakamura suggested that the public could become the largest future stakeholder in the evaluation of research output.

Next Julie Glover explained that NHMRC introduced a policy in 2014 requiring open access after a 12-month embargo and encouraging sharing research data was added in 2018. In 2022 immediate open access with CC BY licenses was made a requirement provided Aboriginal and Torres Strait Islander cultural rights are respected. She showed how open access publication by NHMRC-funded Australian-authors was above 70% in 2023. Currently NHMRC is considering how to incorporate more aspects of UNESCO's recommendations on open science policy, and

how to reward and recognize open science practices as a member of the Coalition for Advancing Research Assessment. Glover showed the Australian National Persistent Identifier (PID) Strategy to accelerate research impact through universally connected PIDs.

Next Amanda French explained why the Research Organization Registry (ROR) is a key part of open science. Open metadata enables tracking and analysis of open science and ROR as a global, community-led registry of open PIDs for research organizations has been established as a widely adopted standard. ROR makes information about research organizations clean, consistent, and easy to exchange among software systems. Not only can research outputs be reliably associated with those organizations, but also ROR is designed for research systems to be able to track research organization activity and for the analysis of open science. French highlighted how many scholarly systems are using ROR including CHORUS and NISTEP.

The improvement of open access and open data in journals published by academic societies in Japan was illustrated by Tetsushi Ueta with his *NOLTA*, *IEICE*. It was first published in 2010, is on J-STAGE, mandates ORCID IDs, publishes under CC BY-NC-ND licenses, and received its first Impact Factor in 2023. In its development the Information for Authors was updated. Open access requirements were implemented resulting in listing in the Directory of Open Access Journals (DOAJ) in 2024 and indexing in Scopus in 2025. Next will be an open data policy. Ueta summarized that the role of academic societies and publications is to increase the visibility of the research outputs

and to do so open access journals need to be properly structured with clear policies.

The final presentation addressed how the American Physical Society (APS) is using open science to maximize the return on investments in scientific research. Jeff Lewandowski introduced APS which has more than 50,000 members. Of its 17 journals 7 are fully open access and nine are hybrid. The Physical Review journals have been supporting open science for over 30 years since the launch of arXiv in 1991, the first fully open access journal in 1998, and a revised open access policy in 2009 allowing posting of accepted manuscripts after a 12-month embargo. Lewandowski said that further investment in open science will be with PIDs, enhanced metadata, open data policies, platform upgrades, and open peer review, with sustainable open access whilst maintaining quality and trust in peer review and workflows.

## 2.3 Discussion

Kazuhiro Hayashi showed a PID map as a potential strategy for Japan and moderated a discussion about the integration of PIDs globally and locally with collaboration involving all stakeholders across the scholarly research community. AI was raised as a risk to integrity and its value in tracking metadata.

Presentations and video recordings are available on the CHORUS event page here: <https://www.chorusaccess.org/events/jst-chorus-forum-strategies-to-ensure-transparency-and-digital-tracking-of-research-outputs/>

### 3 B4 JST-CHORUS Forum: 研究成果の透明性確保とトラッキングのための戦略(浅野光基, 科学技術振興機構)

#### 3.1 セッション概要

日本の新しいオープンサイエンス政策では、2025年度から新たに公募を行う競争的研究費によって創出された学術論文や根拠データの即時オープン化が義務づけられる。世界的にも、研究成果のオープンアクセス義務化の動きは加速している。このようなオープンサイエンスの流れにおいて、研究公正の維持が重視されており、研究者には、研究プロトコルの遵守、独創性の表明、資金の詳細の開示や研究に使用された設備などについて明示することが求められている。また、研究成果論文の投稿や出版の際には、研究成果の透明性やアクセス性を確保するために、リンク、永続的識別子 (PID)、根拠データの記載が求められている。

本フォーラムでは、国内外から講演者を招き、出版プロセスにおける研究者の責任について議論し、研究成果の透明性の確保と研究成果情報のトラッキング促進に向けた戦略を探った。

#### 登壇者

林 和弘 (モデレーター)、文部科学省科学技術・学術政策研究所 (NISTEP) データ解析政策研究室長

中村 春木、大阪大学蛋白質研究所 名誉教授、*Biophysics and Physicobiology* 編集委員長

Julie Glover, Executive Director, National Health and Medical Research Council,

Australia (NHMRC)

Amanda French, ROR Technical Community Manager, Crossref

上田 哲史、徳島大学情報センター 副センター長・教授、*Nonlinear Theory and Its Applications (NOLTA)*, *IEICE* 編集委員長

Jeff Lewandowski, Director of Publishing, American Physical Society (APS)

#### 3.2 講演

中村春木氏は、多様なステークホルダーの視点からオープンデータとオープンサイエンスについて考察した。冒頭では、2000年代の計算科学、2010年代のデータ駆動型科学、現在のAI駆動型科学という、科学の変遷について紹介した。続いて、オープンアクセス化のために論文掲載料 (APC) を支払うことに対する研究者の抵抗感や、論文出版に基づいた研究業績評価の現状を述べ、ステークホルダーが公正に評価できるように、研究成果はオープンにすべきであると強調した。将来的には、市民が研究成果の評価における最大のステークホルダーになる可能性を示唆した。

Julie Glover氏は、NHMRCのオープンアクセスポリシーの歩みを紹介した。2014年に12カ月のエンバゴ後に論文をオープンアクセスにすることを義務化し、2018年には研究データの共有を推奨する方針を追加。さらに2022年には、先住民族の文化的権利を尊重することを条件に、CC BYライセンスによる即時オープンアクセスを義務付けたと説明した。また、NHMRCから助成を受けたオーストラリア人著者による論文のオープンアクセス率が2023年には70%を超えたことを報告した。現在NHMRCは、UNESCOの「オープンサイエン

ス勧告」を政策に反映させる取り組みや、CoARAの一員として、オープンサイエンスの評価と促進のあり方を検討している。最後に、研究活動の加速を目指す国家PID戦略を紹介した。

Amanda French氏は、Research Organization Registry (ROR)がいかにオープンサイエンスにとって重要な役割を果たしているかについて説明した。French氏は、RORが、研究機関を識別するためのオープンなPIDレジストリとして、幅広い領域で採用されていること、ソフトウェアシステム間で相互運用できるように提供されていること、研究成果と研究機関の紐付けだけでなく、オープンサイエンスに関する追跡・分析が可能となるように設計されていることを述べた。また、CHORUSやNISTEPを含む多くの学術システムがRORを使用していることを強調した。

上田哲史氏は、日本の学協会が発行するジャーナルにおけるオープンアクセスやオープンデータの改善の取り組みについて、自身が編集委員長を務めるNOLTA, IEICEを例に紹介した。2010年に創刊し、J-STAGEに掲載されている同誌が、論文投稿時にORCID IDの登録を必須としていること、CC BY-NC-NDライセンスで論文を公開していること、2023年に初めてJournal Impact Factorを取得したことを紹介した。その後、投稿規定の改定により、2024年にはDOAJ、2025年にはScopusに登録されたことを紹介した。次のステップとしてオープンデータポリシーの策定を挙げ、学協会と出版物の役割は研究成果の可視性を高めることであり、そのためにはオープンアクセスジャーナルが明確な方針で適切に設計・運営される必要があるとまとめた。

JeffLewandowski氏は、科学研究への投資効果を最大化するために、APSはオープンサイエンスを活用していると説明した。APSが、1991年のarXivの立ち上げ時からarXivでのプレプリント投稿を推奨していること、1998年に初の完全オープンアクセスジャーナルを出版したこと、2009年に12カ月のエンバーゴ後に原稿の公開を許可するオープンアクセス方針を導入したことなど、30年以上にわたりオープンサイエンスを支援してきたことを紹介した。最後に、今後のオープンサイエンスへの投資は、PIDの導入、メタデータの強化、オープンデータポリシーの策定、プラットフォームの改善、公開査読モデルに向けられると述べ、あわせて、査読やワークフローの品質と信頼を維持しつつ、持続可能なオープンアクセスを実現していくことの重要性を述べた。

### 3.3 ディスカッション

5名の講演後に、林和弘氏の司会によるパネルディスカッションが行われた。林氏は冒頭で、日本において将来的に戦略となり得るものとしてPIDマップを示した。続いて、学術コミュニティが協力してPIDをグローバルおよびローカルに統合していくことについて議論した。AIについては、研究公正に対するリスクと、メタデータの追跡における有用性の両側面が指摘された。

講演資料および録画は、以下のCHORUSイベントページから閲覧できる。

<https://www.chorusaccess.org/events/jst-chorus-forum-strategies-to-ensure-transparency-and-digital-tracking-of-research-outputs/>

## 4 B6 AI×文学研究の可能性を探る (村尾優子, 国立国会図書館)

### 4.1 セッション概要

近年、デジタル人文学の分野で、大規模言語モデル (LLM) をはじめとするAI技術の活用が進められている。国立国会図書館 (NDL) においても、AI技術を用いたOCR処理プログラムの研究開発を行うとともに、OCRによるデジタル化資料のテキスト化、本文検索サービスの提供等を進めている。これらの技術やその成果物が文学研究でも活用されつつあり、研究の効率化や深化の観点で、今後さらに大きな役割を果たしていく可能性もある。本セッションでは、デジタル技術を活用して文学を研究されている方々から、AI技術やLLMの可能性、研究のアイデア等を紹介いただいた。本セッションの資料は、当館の「NDLラボ」のページ (<https://lab.ndl.go.jp/event/joss2025/>) で公開している。

### 4.2 講演

最初に青池亨 (NDL) から、NDLOCR, NDL古典籍OCR等の技術や、次世代デジタルライブラリー, NDL Ngram Viewer等のテキストデータ関連サービスについて説明した。続いて、4人の研究者が順に登壇し、AI技術を活用した文学研究について紹介した。

近藤泰弘氏 (青山学院大学) からは、NDLが公開しているOCR全文テキストデータを取得し、生成AIを用いて活用する事例が紹介された。著作権保護期間が満了した『大日本国語辞典』初版を例に、テキストデータと位置を示す座標からなるOCRデータから本文テキストを見出し語と語釈が対応

する形式に復元する過程や、復元したテキストをMCP (Model Context Protocol) サーバに配置して生成AIから利用することでハルシネーションを避け、出所を明確にしつつ生成AIの古典知識の正確性を高められることが語られた。『大日本国語辞典』をAIで利用できることは後継辞書の改訂作業等に役立つ。またこの手法は辞書以外の資料についても有用との見解が示された。

北本朝展氏 (国立情報学研究所) からは、現代日本語での質問に対して江戸時代の古文風テキストで回答するチャットボット「からまる」について紹介された。現代日本語テキストを学習したLLMに、江戸時代の書物から構築したデータセットを継続学習させることで、江戸時代には存在しなかった事物についても古文風の表現に言い換えて回答できる。回答には江戸時代の世界観が反映されており、ある種の創造性や、元のLLMにはない謙虚さも見られる。AI技術により生成されたテキストを従来の作品と対比できる状況において、LLMとは何か、文学とは何かという根源的な問いにつながっていくことが指摘された。

黄晨雯氏 (大阪大学大学院) からは、Transformerに基づく深層学習モデルを用いて、小説投稿サイト「小説家になろう」のオンライン小説のジャンルを分類・可視化して分析する試みについて紹介された。BERT (Bidirectional Encoder Representations from Transformers) ベースの日本語に特化した深層学習モデルに、「小説家になろう」から取得したジャンル・あらすじ情報を学習させ、主成分分析等の手法で可視化した結果、ジャンルの特徴を捉えやすくなった。学習量の増加や、ジャンル・あ

らすじ以外の情報の学習により、考察の幅を広げられる可能性も示唆された。

日比嘉高氏（名古屋大学大学院）からは、NDLが公開している戦前期の小説を対象に、小説のジャンルを判定する試みが紹介された。OCR全文テキストデータから小説本文を復元した上で、本文データから小さな場面（チャンク）を切り出し、場面ごとにLLMで主題を自動抽出・分析した。この結果、トピックモデルLDAを用いた従来の分析に比べ、主題の抽出精度と解釈のしやすさが向上した。この手法は、従来の作品単位での研究とは異なる、場面主題の単位、間テキストの文学研究と位置付けられる。一方で、AIによる多数の作品の分析結果をどう検証するか、既存の知とどう交差させていくかといった課題も指摘された。

#### 4.3 ディスカッション

近藤氏をモデレータとして北本氏、黄氏、日比氏の4人でディスカッションを行った。まず近藤氏から、適用しているAI技術が4人とも異なっていた点が指摘され、4人の研究者からAI技術選択の背景や意図について説明された。研究に関するデータの公開について参加者から質問が寄せられたことに対して、Hugging FaceやGitHubでの公開状況、今後の公開予定が紹介された。また、NDLが公開しているOCR全文テキストデータ（著作権保護期間満了分）から本文を復元して再配布することの可否について質問があり、NDLから、特に問題ない旨を回答した。

日比氏からは文学研究における作品テキストデータの精度が重要であることや、データ整備・共有が研究者の責務になると考えられることが指摘された。また、NDLが

公開しているOCR全文テキストデータが、レイアウトの構造化や読み順序の整序等の面で改良され、使い勝手の向上したOCRによるものに更新されることへの期待が語られた。

文学に活用できるAI技術の可能性や文学へのアプローチの多様性が感じられるセッションとなった。

#### 参考情報

- [1] 次世代デジタルライブラリーのAPIについて、<https://lab.ndl.go.jp/service/tsugidigi/apiinfo/>（2025年7月10日参照）。
- [2] AI中納言を作る—ClaudeによるSQLコーパス操作—、<https://note.com/yhkondo/n/nac7f19127162>（2025年7月10日参照）。
- [3] 江戸時代の古文風テキストで会話できるチャットボット「からまる」を公開、<https://sakana.ai/karamaru/>（2025年7月10日参照）。
- [4] なろう小説APIから全ての作品情報を一括取得するPythonスクリプト、<https://github.com/karupoimou/NarouDataAllDownload>（2025年7月10日参照）。
- [5] 日比嘉高：トピックモデルを利用した近代日本小説のサブジャンル判定の試み、*跨境*, 19(1), 59-83, 2024, <https://doi.org/10.22628/bcjjl.2024.19.1.59>（2025年7月10日参照）

## 5 C1. 今、「オープンさ」を改めて問う(井出和希, 榎本啄杜, 長門裕介, 神崎隼人; 大阪大学)

### 5.1 セッション概要

近年, 「オープンサイエンス」の学術研究政策上の重要性が増し, 基盤整備と共に推進の機運が高まっている. 公的資金による研究に基づく学術論文や付随するデータの「即時オープンアクセス」も実現されようとするものの, 「オープンさ」のそもそもを考え語らう機会は少ない.

オープンであることは常によきことであろうか. オープン概念そのものを整理しながら, 結果としての「オープンさ」だけでなく, その過程についても議論を深めるべく本セッションを企画した.

注目され易い知的財産権や安全保障上の問題をこえて, 「社会への還元」といった論点に閉じることなく, これからの学術研究のあり方を考える機会とし, 参加者(最大111名)からのコメントも踏まえた討論を行った.

### 5.2 講演

最初に井出から導入として学術政策の動向を紹介すると共に問題提起を行った後, 「「オープンさ」の所在と倫理的・社会的課題」(榎本, 長門), 「「オープンさ」と知識の経済: ラテンアメリカから考える」(神崎)と題した話題提供に進んだ.

学術政策の動向を紹介するにあたっては, 前提にあるデータの扱いに関するFAIR原則[1]に触れ, G7科学技術大臣会合コミニケや即時オープンアクセス方針の根底にある「すべきこと」としてのオープンさ向上を論じた. 近い将来これにインセン

ティブ付与された場合, 「よきこと」としてのオープンさが強化されることに対する懸念(=オープンでないことが悪しきことになっていく)を示すと共に, フレキシブルさの重要性を指摘した[2]. この背景には, 検証に要する時間の長さ, センシティブなトピックを扱うことによるハレーションの懸念, フィールドに住まう人々の生活への影響といったことが「オープンにしにくい研究を避ける」という動機を生み出し, 延いては学問の豊かさを失わせかねないといったことを想像することの肝要さがある.

所在と倫理的・社会的課題では, そもそもオープン概念とは一体何なのかということに注目し, オープンソースソフトウェアやオープンデータ, オープンアクセス, オープンサイエンスといったことが一文脈依存的ではあるものの「アンチ囲い込み」という理念的モチーフとしての共通性を有することを述べた. その上で, アリストテレス的な中核依存的同名異義性(*core-dependent homonymy, CDH*)の考えを用いることで秩序ある多様性を捉えることを提案した[3]. これを踏まえ, オープンさの部分(要素)や段階(程度)を許容することの重要性を強調しながら概念を定義した.

倫理的・社会的課題としては, 悪意ある利用, 情報過多と品質管理の困難さ, 公共的利益と個別的利益のバランス, 不平等・格差の拡大といったことを挙げた.

「オープンさ」と知識の経済では, 人類学を背景にもつ神崎がラテンアメリカの社会, 経済, 文化からそのあり方を再考した.

サイエンスが西欧近代由来の特異的・特権的な知識形態の経済を形成してきたなかで, 他なる知識形態と価値から一とすれば

無批判になり得る一オープンさを考える意義がある。独自に発展した／せざるを得なかった背景には、商業的・学術情報出版の弱さ、リソースの乏しさ、スペイン語／ポストガル語の基盤といったことが挙げられる[4]。

そこでは、アマゾン先住民の人々が自分たちに関する非／先住民研究者による学術論文を閲覧し、自分たちの権利要求のための証拠として活用するといった実践も為されている。これは、草の根運動と国家との部分的な結合（＝中央集権化に向かうと分散を志向し国家に抗すること）を示唆するものであり、「私たち」に対しても「義務化」のような強い力に対して自らその営みを問い直すことの意味の大きさを想起させる。

### 5.3 参加者からのコメントと応答、総括

以下に参加者からのコメントの内容を一部抜粋して応答を付す。

1) オープンアクセス比率は様々な課題を解決したり落としどころを探ったりしながら「100%を目指していくべき」かと問われたが、これについては「100%」であることを目的化したり、オープンであることが望ましいと決めつけたりせずに研究者の自律性に依拠した柔軟さを残すことが重要であると考えている。

2) 「市民科学」における上から目線ののにおいに対して「もうひとつのサイエンス」（註：神崎発表に依る）はそれを乗り越える視点を感じたとの意見があった。これは正にオープンさ（の促進）においても同様に生じている課題であり、私たちがどのように在りたいかを示していくことの不可欠さを示唆するものでもあろう。

3) もっと人文社会科学の分野の人間が参画して多様なデータと価値観をそこ（オ

ープンサイエンスやリポジトリの促進）に組みこむべきとの指摘については、促進を前提におくかは留保するが、重要である半面、実務上で専門性を発揮してもよいと認識できるような施策が重要である。

他にも、オープンさが固着化する傾向を実感すると共にさまざまな具体の場におけるオープン化について語る場を広げていく必要があるように感じるとの見解も共有された。本企画がそのような場づくりと実践を行き来する機会の端緒になれば幸甚である。

### 謝辞

本報告の一部は、日本財団・大阪大学 感染症対策プロジェクト、科研費 若手研究（23K12845）の支援を受けて推進した研究に基づくものである。

### 参考文献

- [1] GO FAIR : FAIR principles , <https://www.go-fair.org/fair-principles/>, (2025年7月19日参照) .
- [2] Ide, Kazuki: “Japan can embrace open science — but flexible approaches are key”, *Nature*, Vol.629, No.8011, p.286, 2024.
- [3] 榎本啄杜, 長門裕介: 「多義的なオープン概念を整合的に理解する」, *Jxiv*, 2025, preprint, <https://doi.org/10.51094/jxiv.1299> (2025年7月19日参照) .
- [4] 村井友子: 「ラテンアメリカの学術情報プラットフォームの活動」, *ラテンアメリカ・レポート*, Vol. 38, No. 2, pp. 86-92, 2022.

## 6 D1「地域と連携した人文系学術研究データ生成の可能性と課題」 (堀井 洋, 合同会社AMANE)

### 6.1 セッション概要

古文書や民具など地域資料を対象とした学術研究データ生成については、これまで様々な議論・取り組みが行われてきた。特に、昨今の機械可読な学術研究データに対する需要の高まりにより、より大規模かつ効率的なデータ生成手法の確立が急務となっている。その一方で、地域の過疎化・人口減少が急速に進行する現状においては、市民の地域資料に対する興味・関心を喚起させ、地域文化に対する理解を継続的に促すことも重要な課題である。

合同会社AMANE（以下、AMANE）・長野県伊那市・みんなで翻刻プロジェクト（以下、みんなで翻刻）は、2024年9月に「地域資料デジタルアーカイブの公開にむけた包括的連携協定」を締結し、伊那高遠地域の古文書を対象とした「第1回古文書解読コンテスト」を実施した[1][2]。本セッションでは、本事例を中心に市民参加型の人文系学術研究データ生成の可能性と課題について、話題提供と議論を行った。

### 6.2 話題提供

本セッションでは、セッション提案者であるAMANE堀井による趣旨説明の後に、登壇者3名からの話題提供があった。

#### 第一回：古文書解読コンテストについて (伊那市地域おこし協力隊 前田和弘)

前田氏からは、2024年10月から2025年1月末まで開催された古文書解読コンテストについて、話題提供があった。本コンテ

ストは、高遠藩の藩主内藤家やその家臣に関する古文書・古記録等資料834点を対象に、みんなで翻刻およびAMANE Archivesを利用してオンライン上で実施された[3]。その特徴としては、産学官連携に基づく実施体制や、上位優秀者が賞金を獲得する点などが挙げられる。加えて本発表では、地域おこし協力隊に所属し、かつクリエイターである前田氏が自治体と学術界を繋ぐ役割を担っていたこと、根幹となるコンセプトが「参加するだけで日本の課題解決になる知的な遊びを作ろう。」であったことなどが紹介された。前田氏は、今後に向けてイベントの全国展開や事業化を構想している。

#### みんなで翻刻と古文書解読コンテスト@伊那市（東京大学地震研究所 加納靖之）

加納氏からは、みんなで翻刻の紹介と同コンテストでの活動について報告があった。みんなで翻刻の初期は、京都大学古地震研究会に所属するメンバーが中心となり、地震や火山噴火等に関する史料を対象として遠隔地の専門家による翻刻（くずし字で記された文字を活字化）の実現を目指していた。その後、学術クラウドソーシングの考え方にに基づき、市民参加型史料翻刻プロジェクトとして2017年1月からサービスを公開し、2025年6月現在28プロジェクト、通算参加人数10,844人、翻刻文字数4,848万文字の成果を挙げている。現在もAI技術を適用した新バージョンのリリースやNDL古典籍OCRへのデータ提供など、積極的なデータおよびサービスのオープン化に取り組んでいる。市民参加による大量な翻刻データの獲得と、その成果のオープン化によるAI等の技術発展への貢献・フィ

ードバックは、“史資料・翻刻データの好循環”をもたらすことが期待される。

「地域資料デジタルアーカイブの公開にむけた包括的連携・協力に関する協定」におけるAMANE Archivesの役割（合同会社AMANE 小川歩美）

小川氏からは、多様な組織が連携する地域資料デジタルアーカイブの視点から、古文書解読コンテストにおける3組織の役割と連携の意義について発表があった。AMANEが2023年5月から社会貢献事業の一環として構築・公開しているAMANE Archivesは、2025年6月現在、6アーカイブ8コレクション約12,000点の資料について、IIIF形式で画像およびメタデータを公開している。アーカイブ公開の主な目的としては、学術的な試行および新たな知見の獲得や調査研究事業成果の逐次公開、資料公開を通じた地域・社会との連携が挙げられる。古文書解読コンテストでは、伊那市から提供されたメタデータおよび資料画像データをもとにAMANE ArchivesにおいてIIIF形式画像データを生成・公開し、みんなで翻刻がそれらを利用することで、イベントを実現した。従来の多くの事例では、資料所蔵機関がデジタルアーカイブの公開から活用事業の実施までの全てを担うことが一般的であったことに対して、本事例では多様な組織が互いに連携し役割分担をすることで、新たな取り組みを実現した。本連携形態は今後の発展・普及が期待される一方で、コンテスト開始時からの急激に増加するサーバーアクセスへの対応など、解決すべき課題が複数存在することも本発表において示された。

## 6.3 ディスカッション

ディスカッションパートでは、学術データ作成に参加する市民のハードルや、市民を学術プロジェクトに巻き込むための手法やブランディングについて、議論が行われた。さらに、生成AI技術が市民参加型学術プロジェクトに与える効果・影響はどのようなものか？という問いに対しては、加納氏からは、参加者がそれぞれにレベルに応じた翻刻をAI技術の支援により実現すること、前田氏からは、コンテストを実施する際のルールをAI技術の利用を前提としたものを検討する必要があること、などの発言があった。さらに、小川氏からは、デジタルアーカイブについても、対人向けのサービスと併せて、AI経由での利用も想定した構成を目指すことなどが示された。

## 6.4 まとめ

市民参加型の人文系学術研究データ生成については、学術研究分野のみならず一般社会からも非常に大きな関心と期待が寄せられている。その一方で、AI処理との役割分担など、新たに検討を進めるべき課題が山積している。今後もこれらの話題・課題について、議論する機会を設ける所存である。

### 参考文献

- [1] みんなで翻刻, <https://honkoku.org>
- [2] 古文書解読コンテスト, <https://www.komonjo-contest.com>
- [3] AMANE Archives, <https://ourarchives.amane-project.jp> (2025年7月31日参照)

## 7 E2 若手セッション:オープンサイエンスについて話そう!(大月希望, 東京大学史料編纂所)

### 7.1 セッション概要

JOSSは今回の開催で7回目を迎えるが、参加者、発表者とも若手が少なく、そもそもオープンサイエンスに関心を持つ若手も少ない。そこで、オープンサイエンスについて若手や若手を育てる人(教員等)の関心対象、取り組みを共有し交流したいという考えから、本セッションを企画するに至った。成果報告だけでなく経験や悩みを共有できる場をつくることを目的とした。

はじめに企画者の大月より趣旨説明を行った。若手セッションは今回初めての試みであり、まずは企画者の研究分野である人文系分野を中心に、発表、紹介、議論及び情報交換を行うことを目指した。理工系や社会科学系に比べ人文系のオープンデータは少なく、研究資源の独占も散見される。論文のオープンアクセス化など研究成果のオープン化も進みつつあるが不十分である。研究過程や研究データは一部でオープン化の動きは見られるものの、形式が定まっておらず参照可能性が低い。本セッションでは、オープンサイエンスに関わる実践や研究の内容と課題について、3組の登壇者より話題提供いただいた。

### 7.2 発表、紹介

第一に、合同会社AMANE(以下、AMANE)の小川歩美氏と佐々木紫帆氏から、AMANEが学術研究・社会貢献活動として構築・維持管理している地域と連携するデジタルアーカイブ「AMANE Archives」[1]についてと、

具体的な取り組みとして「奥州市学校資料」の調査・公開の流れの紹介が行われた。

第二に、総合研究大学院大学、国立情報学研究所他に所属する前山和喜氏[2]から、日本のコンピューティング関連資料について発表が行われた。資料が未保存のまま失われ、何が重要か不明なために残らないというオープン化以前の課題が指摘され、具体例として「使われた状態のパンチカード」が挙げられた。また、企業とNDA(秘密保持契約)を締結して研究を行う際の機密資料や流通に乗らない「グレー資料」の扱いについて、情報の信頼性確保やセキュリティ面での困難、公開ルールが未整備であることが課題として示された。

第三に、東京大学の小川潤氏から、「歴史研究プロセスの構造化による暗黙知の可視化と共有に向けて」と題した発表が行われた。小川氏は、歴史研究において膨大な作業やデータが明示されず、研究プロセス自体を記録する仕組みがないことを指摘した。この課題に対し、情報資源の生成過程を記述するためのデータモデル「PROV Ontology」を歴史研究プロセスに特化する形で拡張したデータスキーマの事例研究が紹介された[3]。

### 7.3 ディスカッション

ディスカッションでは、資料のオープン化にかかわる課題について議論が行われた。企業所有データに関して完全なオープン化の前に研究目的に限定した公開を行うことは可能かという質問には、コンピュータ関連企業と歴史研究者がNDAを締結して研究する事例がほぼなく、実践例を作り議論の促進を試みていることが前山氏から

示された。関連して AMANE の小川歩美氏、佐々木氏から、「AMANE Archives」での具体的な工夫が紹介された。資料群全体を把握しつつも公開可能な資料からデジタル化して公開する方針をとり、個人情報を含む資料については表紙のみの撮影やマスク処理、管理用と公開用のメタデータの分離などの工夫を行っている。一方で、建築図面などのように企業が権利を有している場合やセキュリティの問題で公開が難しい資料もある点が指摘された。

データ流通に向けた取り組みに関する質問には、特定地域に限定されない連携型アーカイブやメタデータについて、既存の横断型アーカイブでの扱われ方を検討している旨がAMANEから返答があった。

また、研究プロセスの構造化に関して、特定の史資料を直接的に用いた研究の前段階の、史資料収集・整理といった行為についても詳細な記述が可能なのかという質問が出された。これに対し小川潤氏からは、問いが明確でない段階での史資料収集や収集・整理を主目的とするケースについても記述し、コンテキストも含めて歴史研究のプロセスの中に取り込んでいくことができると考えていることが示された。

## 7.4 まとめ

本セッションでは、人文系分野におけるオープンサイエンスの推進には、資料の特性に応じた保存・公開ルールが必要であること、研究プロセスの構造化により分析過程の再現性を高め、研究行為の分析を可能にする仕組みが示された。また若手ならではの視点で、人材不足の原因として計算機系分野で学芸員資格を取得できる大学が

少なく、理工系資料の保存公開にはまず専門的な教育の基盤が必要なこと、歴史研究よりも最新技術が重視され就職も有利な分野であることが指摘され、それを踏まえて分野間の相互理解を深めたいという意見が出された。理工系分野における歴史・文化的資料の保存と公開には、人文系分野とはまた異なる困難があることが示唆された。今回紹介された逐次公開や研究プロセス構造化が何らかの形で役立つかもしれない、今後も交流ができればと思う。

本セッションでは若手として高校生や高校教員の発表・参加も期待されたが、今回は定期テストや学校行事により叶わなかった。また大学学部生にも積極的に参加していただきたいと考えている。来年度はより広く発表を呼びかけ、参加者獲得に向けた広報の拡大を試みたい。本稿をお読みの方で関心をお持ちの方はご連絡をいただければ幸いである。

## 参考情報

[1] 合同会社AMANE: AMANE Archives, <https://ourarchives.amane-project.jp> (2025/07/29参照)。

[2] 前山和喜: 【連載】計算の歴史学とジェンダー—誰が計算をしていたのか?, <https://bungaku-report.com/computing.html> (2025/07/29参照)。

[3] 小川潤: 「『行為 (Activity)』概念に基づく歴史研究プロセス構造化オントロジーの検討: 歴史研究の再現性・再検証性の確保に向けて」, *Library and Information Science*, No. 93, pp. 1-22, 2025, DOI: 10.46895/lis.93.1 (2025/07/29参照)。

## 8 E4 シビックテックにおける参加のデザイン～オープンソース・オープンデータが拓く共創～（瀬戸寿一・白松俊・白澤美幸，東京大学CSIS）

### 8.1 セッション概要

近年，オープンソース技術やオープンデータの活用が進み，行政や市民活動の現場においても新たな可能性が広がっている．そこで本セッションは，大学・行政・企業・市民団体との協力を通じて，オープンデータやオープンソース技術などを活用した新しい参加と共創の場をどのように形成できるか，シビックテック分野での活動実績を多数有する話題提供者の実践事例を通じて，そのあり方や課題を議論した．

### 8.2 話題提供

話題提供に先立って，瀬戸寿一（駒澤大学/東京大学CSIS）より，2024年度に開設された東京大学「シビックテック・デザイン学創成寄附研究部門（CTDI）」の活動概要と本セッションの趣旨説明が行われた．シビックテックは「市民が主体的に情報技術やデータを通じて身近な地域課題や社会課題の解決を目指す活動」と位置づけられ，日本でも東日本大震災やオープンデータ政策の推進を一つの契機として，全国各地で様々な実践型活動が行われている．

陣内一樹（一般社団法人コード・フォー・ジャパン）は，「シビックテックにおける場づくりと共創」と題して，2013年に設立されたコード・フォー・ジャパンの活動経過について振り返ったうえで，シビックテ

ック活動の実践事例として，これまで延べ70回以上の実績がある「ソーシャルハックデー（SHD）」における運営の仕組みや工夫について紹介し，継続的な活動の場を提供する重要性を強調した．SHDが契機となった事例として「アフターピル検索サービス」や，流山地域の市民主体であるエコポイント「もりポ」，さらには偽情報対策ツールとして開発された「BirdXplorer」など，シビックテックの社会実装モデルの諸要素を説明した．

榎本真美（パシフィックコンサルタンツ株式会社/一般社団法人 社会基盤情報流通推進協議会）は，「土木・まちづくり分野におけるデータドリブンな共創の挑戦」と題して，おもに社会インフラに関わるデータ利活用や，デジタルを駆使したコミュニケーションについて，広島県インフラマネジメント基盤「DoboX」などの例を基に言及した．社会インフラに関わるデータ・知識がよりオープンになることで，多様なステイクホルダーの協働が生まれ，透明性・再現性のある意思決定やエンパワーメントにつながる可能性がある一方で，専門性が高く大容量データという特性上，データ操作に関わる障壁や，専門家や行政と市民との意識ギャップをシビックテック的な取り組みからどう埋めていけるかが課題であると説いた．

小泉和久（シビックテック・ウラシマ）は「地域と一緒にシビックテック活動」と題して，浦安市を対象に様々なオープンデータを駆使しながらデジタルマップ（視覚化することで，地域活動へ活用し対話のツールになる可能性を紹介した．この活動では，地元でのまち歩きや各種ワークショップを経て立ち上がってくる新たな（地理空

間) データや事例も多く、行政と住民という関係性だけでなく、地域のいろいろな諸団体とのコラボレーションが、シビックテック活動から創発されている例として「景観を巡るまち歩き」・「親子de通学路探検隊」・「うらやすドリームクエスト」などが例示された。

白松俊(名古屋工業大学/東京大学CSIS特任教授)は「LLMが拓く市民共創の新しい可能性」と題して、主に対話による合意形成や市民同士のコラボレーションを支援するためのAI技術として、LLM(大規模言語モデル)を活用したシステムの研究および実践例を紹介した。LLMを対話の場に導入することで、多様な立場のペルソナを生成することができ、例えば議論をシミュレーションすることや、対話の場に足りない観点(ステイクホルダー)からの意見を生成できるなどの具体的な効果を挙げた。これらは、人同士の合意形成を行う上で課題となる対立感情を極力抑えた形で、LLM等の技術が「社会の結び目」になりうる可能性を示唆するものである。

### 8.3 ディスカッション

4件の話題提供を受けて、白澤美幸(東京大学CSIS/一般社団法人コード・フォー・ジャパン)の進行により、主に参加者から登壇者への質問に答える形で「オープンな参加型プロジェクトを通じた社会変革の可能性と課題」を議論した。

このパートでは、「Slido」(オンラインのQ&A投稿ツール)を介して、参加者から様々なコメントが寄せられた。その一例は以下のとおりである：

・シビックテックに取り組むには何(ハード・ソフト・環境)が必要か？

・(登壇者が)シビックテックに興味を持ったきっかけはなにか？

・シビックテック始めたいけど、ハードルがまだあるなと思います。どうしたらいいですか？

これらの意見については、シビック「テック」という用語が付くことで、参加の心理的障壁が発生しうる点をどのように軽減できるか、各登壇者が実践を行う上での創意工夫について紹介し合った。この点はシチズンサイエンス活動とも共通する部分が多く、活動の場やきっかけをどのように開いたものにできるかが鍵となる。また、シビックテックを通じた参加のあり方を問う意見は次の通りであった：

・シビックテックのような参加型の文化では、意見をうまく表現できる人や十分な時間に余裕のある人など、いわゆる恵まれてる層の意見が吸い上げられやすい構造にありませんか？

・まちの課題を持っていない市民も参加できるような環境があればいいですね。いまは、まちの課題を持っている人たちだけの集まりのような気がします。

これらの意見に対しては、シビックテック自体の認知度・普及度が日本で途上である一方で、オープンデータやオープンソース文化といった活動の「透明性」を重視して取り組まれた点を踏まえて、継続的な活動を行う重要性が強調された。シビックテックを担う人材の発掘・育成については、東京大学「シビックテック・デザイン学創成寄附研究部門

(CTDI)」の一環として講座の試行や教科書を制作予定であることも付言する。

## 9 F1 研究データをより見つけやすくするためのメタデータ変換と学術機関リポジトリへの登録(能勢正仁、名古屋市立大学)

### 9.1 セッション概要

研究データの検索は、一般的に専門性の高いホームページからのみ利用が可能な状況であり、専門外のより広い研究コミュニティや一般市民に対しては、「Findability」はそれほど高くはない。研究境界領域における融合研究や市民科学を促進するためには、メタデータデータベースの被検索性を向上させる必要がある。こうした問題に対し、(1)専門分野のメタデータスキーマで記述されたメタデータを、広く利用されている汎用メタデータスキーマに従うメタデータ形式に変換し、(2)変換したメタデータを、学術機関リポジトリに登録する、ことで解決を試みた。このセッションでは、上記取り組みの報告に加え、メタデータ互換性に関わる話題を提供し、研究データを見つけやすくし、その流通・利用を促進するための方策について議論を行った。

### 9.2 講演

当日は、セッションの趣旨説明 (doi: 10.18999/2013002) の後、次のような5つの講演が行われた。

- 新堀 淳樹、大平 司、直江 千寿子 (名古屋大学) 「名古屋大学における研究データの可視化・検索向上を目指したメタデータ変換と機関リポジトリへの登録について」 (doi:10.18999/2013003)

大学間連携事業IUGONETでは、太陽地球系物理学分野でよく用いられるSPASEデー

タスキーマを採用し、現在では1200以上のメタデータを作成している。名古屋大学はIUGONETに参画し、284件のメタデータを登録してきた。これらのメタデータをIUGONETや宇宙地球科学分野内にとどまらず、広く検索されるようにし、研究データがより発見・利用され易くなるように、メタデータ形式をSPASEからJPCOARに変換した。そして、変換したメタデータを名古屋大学機関リポジトリ (NAGOYA Repository) に登録した。これらのメタデータは、他のメタデータデータベースからハーベストされ、NII学術機関リポジトリデータベースやGoogle Dataset Searchなどでも検索可能となっている。また、メタデータスキーマのバージョンアップに対応して、メタデータ対応表を改修し、継続登録ができるようにメンテナンスを行っている。今後は、開発したツールの共用や異分野データベースの調査など、学術機関リポジトリへの登録に関して学術分野コミュニティ内の情報共有や連携を強化したい。

- 阿部 修司、芦北 卓也 (九州大学) 「研究データの発見・検索性向上に向けたメタデータ変換と学術機関リポジトリへの登録：九州大学における適用事例」 (doi:10.18999/2013004)

九州大学国際宇宙惑星環境研究センターでは、世界各地で取得した地磁気観測データについてSPASEメタデータを作成し、IUGONETに登録を行ってきた。これらのメタデータを、名古屋大学と同様の手法でJPCOARメタデータ形式に変換し、九州大学の機関リポジトリ (QIR) に180件登録した。QIRでは、国内の多くの機関リポジトリで採用されているJAIRO Cloudとは異なるシステムを用いているが、そうした機関でも

展開可能であることを実証した。また、変換にあたってはGakuNin RDMやChatGPT等のツールを独自に活用し、迅速でスムーズな実施を行った。

- 三上 絢子（JPCOARコンテンツ流通促進作業部会/室蘭工業大学）「JPCOARと名古屋大学の研究用メタデータマッピングプロジェクトでの協働について」（doi:10.18999/2013005）

名古屋大学と九州大学で行っているメタデータ形式の変換と機関リポジトリへの登録に関して、JPCOARコンテンツ流通促進作業部会では、メタデータスキーマバージョンアップデータの情報提供とプロジェクト事例紹介により、協働を行ってきた。

研究データの利活用の促進は、JPCOARの活動でも意識されてきており、研究データ公開の研究者側のメリットを伝えるのは重要である。今回の事例紹介 (<https://schema.irdb.nii.ac.jp/ja/case-space-nu>) では、実際に分野メタデータを公開した事例について、具体的な公開方法や公開によるメリットについて紹介できる貴重な機会であった。今後も様々なニーズに対応するJPCOARスキーマ活用事例の共有を行う計画である。

- 藤原 一毅（国立情報学研究所）「未公開のメタデータを流用したい～NBDCヒトDBから包括的未病DBへ～」（doi:10.18999/2013006）

ヒトに関する様々なデータを共有するためのプラットフォームであるNBDCヒトデータベースでは、BioProject形式のメタデータが作成される。一方、ムーンショット目標2「超早期に疾患の予測・予防をすることができる社会の実現」では、包括的未病データベースのメタデータがGakuNin

RDM上で作成されている。研究者がNBDCにメタデータを登録する際には、基本的な情報をエクセルシートに記入し、それをNBDCの登録担当者がBioProject形式のメタデータに変換するが、同時に、そのエクセルシートを使ってGakuNin RDM上の未病データベースへ登録する手順を開発した。変換プログラムは、pythonにより実装し、Jupyterノートブック上で実行する。メタデータの変換および登録をad hocなハードコーディングで実施した例であり、こうしたアプローチもありうる。

- 叢 艶、青木 学聡（名古屋大学）「分野別・汎用メタデータ互換性のスクリーニング調査」（doi:10.18999/2013007）

「研究データエコシステム構築事業」においては、ルール・ガイドライン整備の一つとして、機械可読データの標準化が挙げられている。異なる分野間でメタデータの変換効率を高め、相互流通を拡大していくには、様々なメタデータスキーマの共通項目の抽出と比較が不可欠なためである。そこで、分野別メタデータの仕様書のあり方を調査し、各仕様書に記載されているタイトル、作成者、作成日時などの記述項目を抽出し、それらとDataCiteメタデータスキーマとのマッピング対応関係の構築を目標とした。

125件の分野別メタデータスキーマから、対応関係が導けそうな8件のメタデータスキーマを対象とし、DataCiteメタデータスキーマの必須項目へのマッピングの可否を判断した。今後は、これに基づき、メタデータの互換性に関する作成方針の策定、マッピングにおける不足項目の補完手法、および変換ツールの開発について、検討を進めていく予定である。

## 10 F2 異分野の研究データをつなぐ: 統合検索システムの構築とメタデータの課題 (朝岡誠, 国立情報学研究所)

### 10.1 セッション概要

近年、オープンサイエンスの推進に伴い、政策や研究資金提供機関の要請、データ駆動型研究の発展などを背景に、研究データの公開と共有の重要性が高まっている。しかし、研究データは機関や分野ごとに異なるメタデータ標準で管理されており、横断的な検索や利活用は容易ではない。

本セッションでは、東京大学社会科学研究所の南山氏による司会のもと、異なる機関・分野の研究データを統合的に検索可能なデータカタログの運用をテーマに、統合検索システムの構築意義とその効果について議論を行った。メタデータの標準化などの技術的課題にとどまらず、多言語化やデータカタログの利用実績評価、そして異分野連携における法的・倫理的課題についても検討した。

### 10.2 講演

最初に国立情報学研究所の朝岡からは、人文学社会科学分野における統合データカタログ「JDCat」の事例が紹介された。社会科学では、構造化された量的データ(例: 調査票に基づく数値データ)が中心であり、変数情報や回答選択肢を含むメタデータの整備によって、再利用や二次分析が比較的容易に行える。一方、人文学分野では、非構造的な一次資料、特に画像データが中心であり、OCRによる文字認識や人物・地名のタグ付けといった精緻なメタデータ

付与が必要とされる。JDCatでは、DDI (Data Documentation Initiative) に基づいた独自のスキーマを採用し、31の項目を設定し、多言語対応(日本語・英語)、統制語彙による専門用語検索、Jupyter Hubによる分析環境の連携などが実装されている。今後の課題として、知名度向上や分析ツールとのさらなる統合、関連文書やコードのリンク、データ提供機関の拡大が挙げられた。

次に、国立遺伝学研究所の有田氏から、生命科学・生物学分野におけるデータカタログ整備の課題が報告された。すでに整備が進んでいる同分野においても、著作権や分類体系の変化といった運用上の課題が存在する。特に、和名と学名の対応付けや、学名の動的な変更への対応が実務上の大きな困難となっている。INSDC (国際塩基配列データベース連携) では、世界各国の論文に掲載される全塩基配列データが日米欧の連携により公開されており、現在20PBを超える大規模データが共有されている。一方で、生物多様性条約(CBD)による「利益配分(ABS)」の義務が、従来の無償公開を原則とするデジタルデータにも適用されることになった点は、学術データベースの運用方針に大きな影響を与える。特に、生物資源が国家の資産と見なされることで、国際的な合意形成と法的調整が求められる状況にあることが共有された。

最後にROIS-DS極域環境データサイエンスセンターの田中氏から、超高層大気データを対象とする大学間連携プロジェクト「IUGONET」と、その知見を活用した異分野データ統合への展開として「AMIDER」プロジェクトについて紹介された。IUGONETでは、太陽地球系科学における多

様な観測データを統合し、SPASE (Space Physics Archive Search and Extract) メタデータスキーマを拡張することで、学際的な解析に対応するデータ基盤を整備した。AMIDERではさらに、極域環境に関する生物・地質・気象データなどを統合的に扱うため、ISO 19139形式のメタデータにも対応したデータベースを構築した。ユーザーフレンドリーなインターフェース、キーワードによる検索支援、関連データのスコア表示など、多機能な設計が特徴である。今後は、ExcelからXMLを生成するメタデータ登録ツールや、自然言語処理を用いた多言語化支援、テキストマイニングによる関連性抽出の仕組みなども開発予定である。

### 10.3 ディスカッション

ディスカッションでは、能勢氏のモデレーションのもと、有田氏、田中氏、朝岡の3名が参加し、データカタログ運営に関する3つの課題について討論が行われた。

1つ目は、多言語化および翻訳の課題である。人文社会系では、文書タイトルや歴史的概念の翻訳が困難であるとの意見があり、むしろローマ字表記など原文保持が望ましい場合もある。一方、生命科学では学名と和名の厳密な対応が求められるなど、翻訳の要件そのものが分野ごとに異なる。いずれにせよ、専門知識を要する用語の揺らぎ、文化的背景の違いをどう扱うかが共通の課題として浮かび上がった。

2つ目は、データカタログの利用実績評価と維持に関する課題である。JDCatではアクセスログの大半が機械的（スクレイピング等）であるため、実利用者数の把握が困難であることが報告された。IUGONETでは1万人規模のユニークユーザーや論文で

の謝辞記載という実績がありながらも、安定的な財政支援の確保は難しく、データ基盤整備の価値が定量的に評価されづらいという構造的問題が指摘された。

3つ目は異分野連携の可能性と技術的課題である。生物学分野のデータカタログでは、学名を軸にして考古学文書や環境データとリンクさせるといった応用が議論される一方で、メタデータスキーマの多重管理やバージョン追従の煩雑さが連携の障壁となっている。特に、生物多様性条約により導入された利益配分の原則は、従来の「自由な学術情報流通」とのギャップを際立たせ、異分野連携において予期せぬ法的・倫理的課題があることが浮き彫りになった。

本シンポジウムを通じて明らかになったのは、研究データ統合が単なる技術的課題ではなく、文化・制度・法制度を横断する複雑な営みであるということである。多様な分野の実践者が互いの前提を理解し、共通の語彙や技術を開発していくためには、継続的な対話と基盤整備のための支援体制が不可欠であることが、改めて認識された機会となった。

## 11 G2 データ管理専門職のペルソナ設計（南山泰之，東京大学社会科学研究所）

### 11.1セッション概要

近年，政策的な側面から大学・研究機関によるデータ管理の要請が強まりつつある。2024年2月に公表された「学術論文等の即時オープンアクセス実現に向けた基本方針」では，所定の競争的研究費によって生成された学術論文及び根拠データは，学術雑誌へ掲載後即時に機関リポジトリ等の情報基盤への掲載が義務づけられており，機関リポジトリ実務者にとっては根拠データの登録が現実問題として迫ってきている。一方で，日本ではデータライブラリアン，データキュレーターなどの専門職が根付いておらず，機関リポジトリの担当者はデータ管理や公開といったタスクに対して，どのような態度で臨めばよいのか具体的なイメージを持てていない。さらに，データ管理は分野によっても流儀が異なるため，個別対応がどの程度発生するのか手探りの状況が続いている。

このような現状把握をもとに，本セッションではワークショップ形式で各分野のデータ管理者のペルソナを作成し，データ管理に求められる専門家としての人物像を明らかにすることを試みる。本セッションは，小林裕太氏及び熊崎由衣氏（千葉大学附属図書館）との共同企画である。

### 11.2話題提供

本セッションの前半では，データ管理に明るい各分野の研究者として，八塚茂氏（独立行政法人製品評価技術基盤機構），福田陽子氏（国立研究開発法人国立環境研究所），杉原孝充氏（国立研究開発法人海

洋研究開発機構）の3名から，ご自身の日常的なデータ管理の実態に関する話題提供があった。八塚氏からは，生命科学系データベースの構築・運用やメタデータの付与といった業務経験に基づく話題提供があった。業務に必要な知識・スキルとしては，データベースの基礎知識やプログラミングスキルが挙げられた。また，バイオ専門でなくとも，利用者目線でのデータベース構築や，不明点の調査・コミュニケーション能力があれば対応可能な業務も多いこと，専門用語については利用者に分かりやすく伝える能力が重要であること，の2点が強調された。

福田氏からは，環境科学分野におけるデータ公開支援や基盤整備，問い合わせ対応といった業務経験に基づく話題提供があった。業務に必要な知識・スキルとしては，情報工学やプログラミングの知識に加え，ワークショップ参加者の分析からはコミュニケーション能力，業務における協働や分担，時間管理，学習能力といった総合的な側面が挙げられた。加えて，マニュアルに縛られない柔軟な対応や，ニーズに応じた情報提供など，研究経験に基づいたデータ利用者の視点が強みとなる旨言及があった。

杉原氏からは，海洋科学分野におけるデータ管理部署のマネジメント業務経験に基づく話題提供があった。業務に必要な知識・スキルとしては，全体フローの理解や分析方法に関する深い知見のほか，チーム内のリソースを俯瞰して作業を割り当てるマネジメント能力が挙げられた。また，データ管理業務をより客観視する観点から，過去に経験した近隣分野（宇宙探査）の知見も役立っている旨言及があった。

### 11.3 ディスカッション

セッションの後半では、各グループが作成したペルソナを会場全体で共有し、参加者を交えてディスカッションが行われた。

ディスカッションでは、大学図書館が多様な分野の研究データを扱う際のドメイン知識の深さや、マネジメント層の戦略的なアプローチの重要性など、実務的な関心を中心にした議論が展開された。前者の話題となる専門用語の理解に関しては、組織やコミュニティの段階によって必要なレベルが異なるとの意見があった。扱う専門分野が限定され、データ管理がルーチンとして確立された組織では一定以上の専門用語に対する理解が求められる一方で、幅広い分野を扱う組織では緩やかな連携やハブ的な役割のほうが重要視されるのではないかと、との指摘があった。

関連して、「研究データ管理業界」で用いられる専門用語を研究者に説明するためには、一定程度の経験値が必要であるという課題意識も共有された。この点、大学図書館員が研究データ管理に今後関わっていく際には、情報処理技術への関心や、レファレンスやガイダンスで培ったコミュニケーションスキルが生かせるのではないかという可能性が示唆された。

### 11.4 まとめ

本セッションでは、複数分野におけるデータ管理者の実務経験を題材に、データ管理に求められる専門家を明らかにすることを試みた。「専門用語を解釈し、利用者へ分かりやすく伝える」業務はどの話題提供者からも言及があり、データ管理専門職の重要な職務と言えそうである。データ提供者と利用者の間に存在するギャップを理解し、個人としてのドメイン知識や研

究経験、及び組織やコミュニティが保有しているノウハウの両方を駆使して課題解決にあたる態度が求められよう。本セッションでの議論が、大学・研究機関におけるデータライブラリアン、データキュレーターの在り方を探る契機になることを期待したい。

### 11.5 謝辞

本セッションの企画に当たっては、研究データ活用協議会（RDUF）ジャパンデータリポジトリネットワーク推進部会（JDARN）から多大な支援を受けた。特に、白井知子氏（国立研究開発法人国立環境研究所）、福田和代氏（国立研究開発法人海洋研究開発機構）からは企画趣旨に関する貴重な助言をいただいた。ここに記して謝意を示す。

### 11.6 参考情報

- [1] 八塚茂. 生命科学におけるオープンデータの理想と現実. <https://doi.org/10.6084/m9.figshare.5682847.v3>
- [2] 福田陽子. NIES における研究データ管理の支援業務. <https://doi.org/10.34462/0002000238>
- [3] 杉原孝充. 海洋研究開発機構（JAMSTEC）でのデータ管理と体制 -データ管理専門職のペルソナ設計：JAMSTEC の場合-. <https://jir.repo.nii.ac.jp/records/2000344>
- [4] 小林裕太, 熊崎由衣. 研究データ管理専門職のペルソナ検討ワークシート. <https://opac.ll.chiba-u.jp/da/curator/900123436/>

JOSS2025予稿

## SSJデータアーカイブのメタデータ作成のAI活用の検討 -LLMアプリケーションを用いた社会調査のメタデータ抽出-

### Exploring the Use of AI in Metadata Creation for the Social Science Japan Data Archive: Metadata Extraction from Social Surveys with an LLM Application

西村幸浩<sup>1\*</sup>, 入山浩一<sup>1</sup>

Yukihiro NISHIMURA<sup>1\*</sup>, Koichi IRIYAMA<sup>1</sup>

1 東京大学

The University of Tokyo

〒113-0033 東京都文京区本郷7-3-1

E-mail: nishimura@iss.u-tokyo.ac.jp

\*連絡先著者 Corresponding Author

本研究では、社会科学分野のデータアーカイブ運営におけるメタデータ作成業務の効率化を目的として、大規模言語モデル(LLM)を活用したメタデータ作成支援のノーコードアプリケーションを開発し、その有効性を評価した。本アプリは、クラウドLLMとローカルLLMをGUI上で選択でき、非エンジニアの職員でもマウス操作のみで調査票・報告書からメタデータを自動作成できる。評価では、LLMと専門家作成のメタデータとのコサイン類似度、処理時間、コストを比較した。その結果、全ての項目で高い一致度を示し、約6万字規模の資料を対象とした処理時間は最大251秒、コストは最大0.28ドルと安価であった。以上より、LLMを用いたアプリはメタデータ作成工数の削減に有効であることが示された。

In this study, we developed and evaluated a no-code tool that leverages large language models (LLMs) to streamline metadata creation in social-science data archives. The tool features a graphical user interface that allows users to switch between cloud-based and local LLMs, enabling non-engineer staff to automatically extract metadata from questionnaires and reports using only mouse operations. We evaluated the tool by comparing (i) the cosine similarity between metadata generated by an LLM and that created by experts, (ii) processing time, and (iii) computational cost. The LLM-generated metadata exhibited a high degree of similarity to expert work across all items. For documents of approximately 60,000 characters, processing time was up to 251 seconds, and the maximum cost was USD 0.28. These findings demonstrate that an LLM-powered no-code application

can effectively reduce the labor required for metadata creation.

キーワード: メタデータ, 大規模言語モデル (LLM), データアーカイブ, ノーコード, 人工知能 (AI)

Keywords : Metadata, Large Language Models (LLM), Data Archive, No-Code, Artificial Intelligence (AI)

## 1 はじめに

本稿は、大規模言語モデル (Large Language Models, 以下LLM) を用いた社会調査のメタデータ作成支援アプリケーションの設計・評価を通じ、LLMがもたらす運用効率化の可能性について検討するものである。

オープンサイエンスの潮流のもと、研究データの共有と利活用の促進は学会全体で強く求められている。東京大学社会科学研究所附属社会調査・データアーカイブ研究センターは、SSJデータアーカイブ (Social Science Japan Data Archive, 以下SSJDA) を構築、労働調査・社会調査の個票データを収集・保管し、その散逸を防ぐとともに、学術目的での二次的な利用のための個票データの提供を1998年より行ってきた。

2025年には、SSJDAはデータリポジトリの国際的な認証基準であるCoreTrustSealの認証を取得した[1]。これは、CoreTrustSealの現行基準移行後、日本国内で初の認証事例であり、SSJDAが国際的な品質基準に適合していることを公的に示すものである。

こうしたデータアーカイブ運営において、労力を要する業務の一つが「調査年」「主要調査事項」などのメタデータの作成である。メタデータの質はデータ内容の理解と適切な利用を支える極めて重要な業務であるうえに、高度に専門的な知識やスキルが要求されるため、現在SSJDAの内部

スタッフによって行われている。調査票・報告書には数百ページに及ぶものもあり、公開データ数の増加に伴い、メタデータ作成の負担は年々増大している。メタデータ作成プロセスの効率化はアーカイブ運営における喫緊の課題である。

そこで、SSJDAでは、2023年10月より、「人文学・社会科学データインフラストラクチャー強化事業」の一環として、人工知能 (AI)、とりわけLLMを活用したメタデータ作成の効率化に向けたシステムの開発を開始した[2]。

初期検証ではメタデータ作成において一定の有効性を確認した (西村・入山 2024) もの、Pythonなどのプログラミング知識を要する点や、OpenAIなどのクラウド型LLMへの機密データの送信が困難であるといった、現場実装上のハードルが残存していた[3]。

そこで、本研究では上記の課題を解消するために、ノーコードGUIで非エンジニア職員でも操作可能で、ローカルLLM/クラウドLLMを用途に応じて選択可能なメタデータ作成支援アプリケーションを新たに開発し、その有効性を検証した。

本稿では、開発中のアプリケーションの概要を説明するとともに、LLMを活用したメタデータ作成の可能性と課題について論じる。これらの知見は、データアーカイブの運営効率化に寄与するだけでなく、より広範な学術データマネジメントにも示唆を与えるものと期待される。

## 2 アプリケーションの概要

### 2.1 ユーザーインターフェイス

本システムの GUI は、調査票や報告書のメタデータを、非エンジニアの職員でもマウス操作のみで生成できるよう設計した。図1に示すように、画面上部には後述する4種類のLLMを選択するプルダウンとファイル追加ボタンを配置し、追加済みファイルはスクロール可能な一覧に即時反映される。生成結果は画面下部のテキストエリアに表示し、Excel形式でのエクスポートもワンクリックで可能である。以上の設計により、データ投入から結果共有までをクリック操作だけで行えるユーザーフレンドリなインターフェイスを実現している。

### 2.2 LLMのモデルとデータソース

本アプリでは、OpenAIのクラウド上のLLMと、秘匿性の高いデータの運用を想定してローカルLLMを選択可能である。クラ

ウド上で処理可能なデータでは、精度・コストに応じてo3, o3-mini, gpt-4.1-nanoの3つのモデルから選択でき、秘匿性の高いデータでは、ローカルLLMとして株式会社ELYZAが2024年6月に発表した、Meta社のLlama3(8B)をベースとする日本語能力強化の調整をしたLlama-3-ELYZA-JP-8B（以下ELYZA）を使用できる[4]。なお、将来的には一般職員が使用するPCでの動作を想定してELYZAについては4bitに量子化した。本研究の開発環境ではNVIDIA RTX A5000のGPUで計算を実行した。

メタデータ作成フローは、西村・入山（2024）が提案するフローにて社会調査のデータからメタデータを作成する。このプロセスは、図2に示すように、調査票・報告書からのインデックス作成、プロンプトとクエリの生成、およびLLMによるメタデータ生成の3つの主要ステップから構成される。プロンプトの作成にあたっては、OpenAIのプロンプトジェネレーターを用

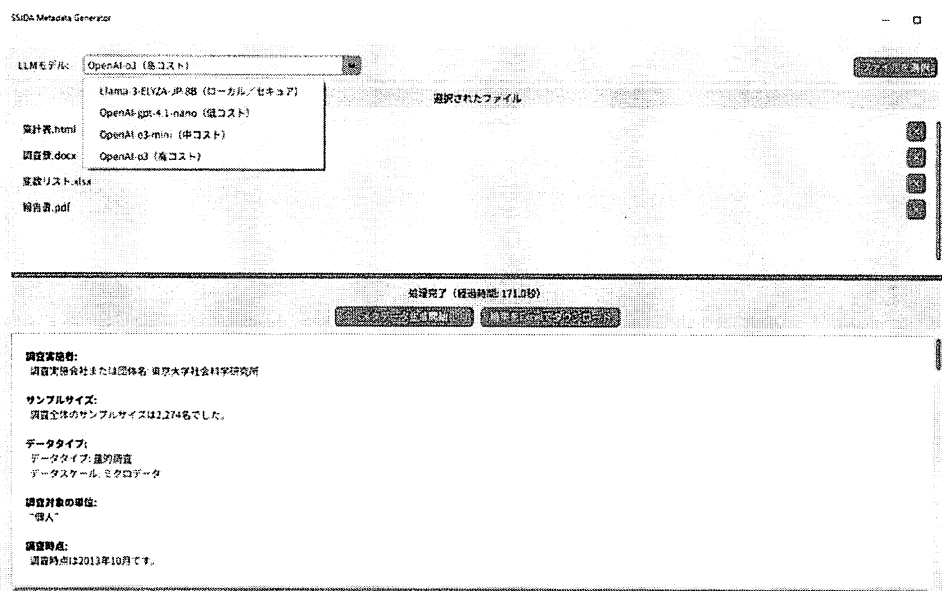


図1 LLMによるメタデータ作成アプリケーション

いて作成したプロンプトを実際の運用に即した修正を加えた。プログラミング言語にはPythonを使用し、主要なライブラリとしてLangChainを使用した[5]。

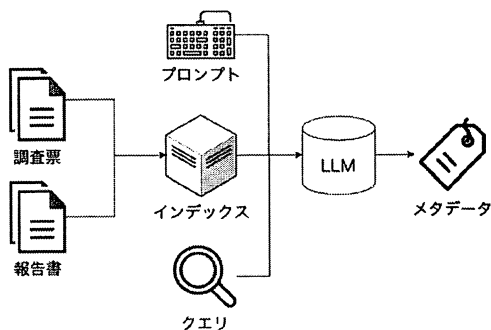


図2 メタデータ作成フロー

なお、本研究で実施したメタデータ作成にあたっては、個票データは使用せず、SSJDAのウェブサイトなどで一般公開されている調査票・報告書を使用した。これらのデータには個人を特定できる情報は含

まれていないことを確認している。

### 3 結果

#### 3.1 モデル間のメタデータの比較

本節では、LLMのモデル間の比較を行う。パイロット調査として、SSJDAにて公開されている10件の調査データに関する調査票・報告書を対象に、LLMでメタデータを作成した。

図3はOpenAIのo3とローカルLLMのELYZAによって生成されたメタデータと、SSJDAの専門家によって作成されたメタデータとのコサイン類似度を図示したものである。いずれのメタデータにおいても高い一致度が確認された。

具体的には、「調査時点」はo3が0.806でELYZAが0.800、「調査対象」はo3が0.904でELYZAが0.903といった形で、多くの項目で同程度の水準の結果が得られた。

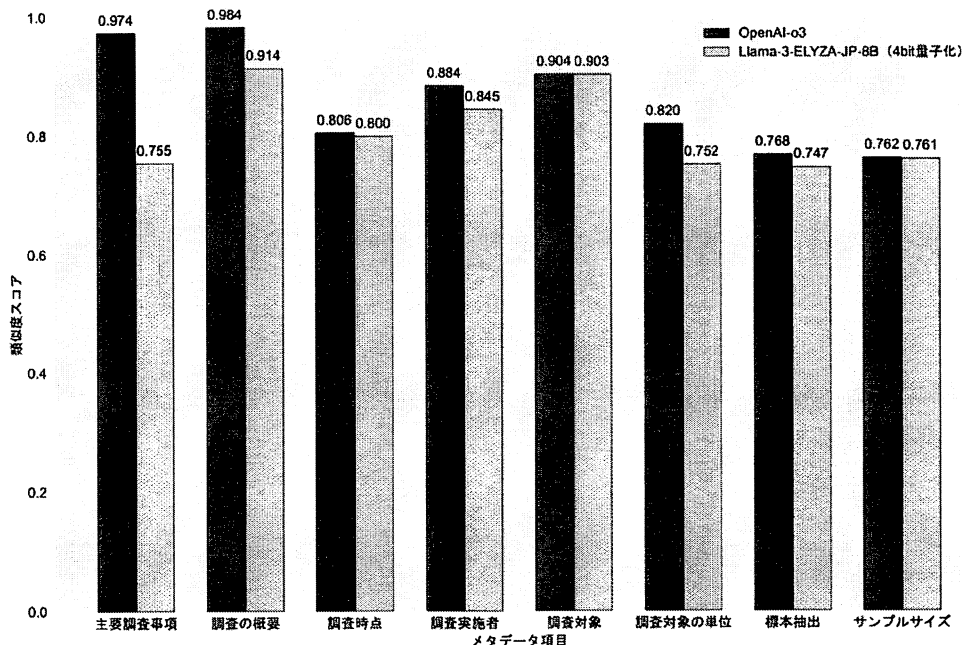


図3 LLMモデル間のメタデータ作成結果のコサイン類似度の比較

このような、「調査時点」や「調査対象」といった、調査票と報告書から該当箇所を抜き出す形式の項目については、いずれのモデルも良好なメタデータを作成する。一方でモデル間の違いが顕著に出たのが、「主要調査事項」と「調査の概要」である。

「主要調査事項」はo3が0.974でELYZAが0.755, 「調査の概要」はo3が0.984でELYZAが0.914と、調査全体の文脈を把握する必要があるメタデータについては、精度に開きがある結果が得られた。

こうしたモデル間の違いはあるものの、最も低い値を示している「標本抽出」においてもo3が0.768でELYZAが0.747と、いずれの項目においても人手によるメタデータ作成時に参考に値する結果が得られている。

### 3.2 メタデータの作成の処理時間・コスト

処理時間とコストにおいて、人手を上回る良好な結果が得られた。図4が示すよう

に、約6万字の調査データに対するメタデータ作成プロセスにおいて良好な結果が得られた。最も時間がかかったローカルLLMのELYZAでも251秒と人間による作業と比較して大幅に高速化している。最も高速なgpt-4.1-nanoでは24秒である。また、コストに関しては、最も高額なo3でも0.28ドル、最も安価なgpt-4.1-nanoは0.06ドルと人手によるコストを大幅に下回っている。

これらの結果は、人手による方法と比較して顕著な効率性の高さを示しており、本プロジェクトの主要な目的であるメタデータ作成の効率化に向けたシステム構築が十分に達成できていると評価できる。

## 4 まとめと今後の展望

本研究では、LLMを用いたノーコードアプリを開発して社会調査データアーカイブにおけるメタデータ作成の効率化を図る方法を検討した。

その結果、社会調査のメタデータ作成に

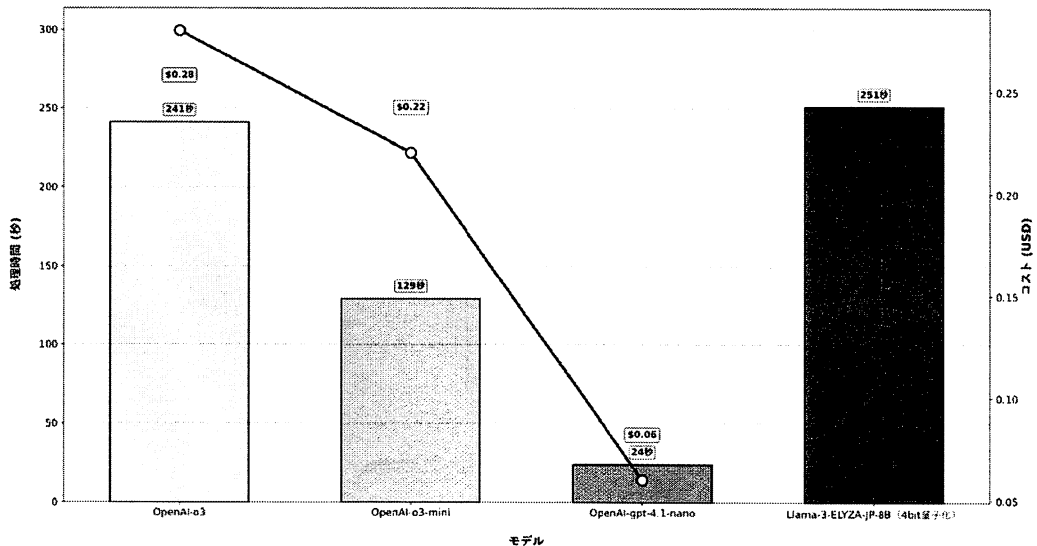


図4 約6万字データからの11項目メタデータ生成の処理時間とコスト比較

において担当者の作業工数の削減の有効な手段となり得ることが確認された。先行研究ではOpenAIのクラウド型LLMに依存していたのに対し、本研究ではローカルLLMを併用する構成を採用したため、非公開の報告書などからメタデータを作成する運用も可能となった。

一方で、3点課題が明らかになった。第一に、社会調査領域に特有の専門用語を正確に扱うためのドメイン適合性である。汎用LLMでも高い類似度を示したが、メタデータ作成上の形式的ルールをプロンプトの制約のみで遵守させるのは限界があり、継続的なファインチューニングが必要である。

第二に、インフラの面では、業務用の10万円前後PCでは処理性能に限界があり、ローカルLLMの運用が困難である。特に、機密データを扱う場合はオンプレミスが必須のため、より小規模なLLMなどを検討する必要がある。

第三に、保守の課題として、LLMの仕様変更により、将来の動作保証が難しい点があげられる。

今後はドメイン特化チューニングとインフラ最適化を進めることで、更なる効率化とスケーラビリティの向上が期待される。

## 5 データ利用可能性宣言

本研究で使用した調査票および報告書などのデータは、SSJDA[6]などのウェブサイト入手可能である。

## 謝辞

本研究はJSPS人文学・社会科学データインフラストラクチャー強化事業（JPJS0032

0231001）によるものである。記して感謝申し上げる。

本研究の遂行にあたり、アプリケーションの開発および分析を西村が担当し、入山がデータ整備を担当した。

## 参考文献

- [1] 東京大学：「社会科学研究所SSJDAが国際認証CoreTrustSealを取得」, [https://www.u-tokyo.ac.jp/focus/ja/articles/z0204\\_00089.html](https://www.u-tokyo.ac.jp/focus/ja/articles/z0204_00089.html)（2025年7月25日参照）。
- [2] 東京大学史料編纂所／社会科学研究所（日本学術振興会委託事業・中核機関）：「人文学・社会科学データインフラストラクチャー強化事業」, <https://www.hi.u-tokyo.ac.jp/di/>（2025年7月25日参照）。
- [3] 西村幸浩, 入山浩一：「SSJデータアーカイブのデータ公開におけるAI活用の検討-LLMを用いた社会調査のメタデータ抽出-」, 情報知識学会, Vol34, No.3, pp.292-297, 2024.
- [4] Hirakawa, Masato, Shintaro Horie, Tomoaki Nakamura, Daisuke Oba, Sam Passaglia, and Akira Sasaki: Elyza/Llama-3-ELYZA-JP-8B, <https://huggingface.co/elyza/Llama-3-ELYZA-JP-8B>（2025年7月24日参照）。
- [5] Chase, Harrison: LangChain, <https://www.langchain.com/>（2025年7月19日参照）。
- [6] 東京大学社会科学研究所附属社会調査・データアーカイブ研究センター：「データ検索」, <https://ssjda.iss.u-tokyo.ac.jp/Direct/datasetsearch.php>（2025年7月10日参照）。

JOSS2025 予稿

## Google フォームを用いた

## 航海データマネジメント効率化に向けた試み

# Improving the efficiency of cruise data management using Google Forms

高津佳宏<sup>1\*</sup>, 齊藤千鶴<sup>1</sup>, 佐藤孝子<sup>1</sup>, 福田和代<sup>1</sup>

Yoshihiro TAKATSU<sup>1\*</sup>, Chizuru SAITO<sup>1</sup>, Takako SATO<sup>1</sup>, Kazuyo FUKUDA<sup>1</sup>

<sup>1</sup> 国立研究開発法人海洋研究開発機構 地球情報科学技術センター

Center for Earth Information Science and Technology, Japan Agency for Marine-Earth Science and Technology

〒236-0001 神奈川県横浜市金沢区昭和町 3173 番 25

E-mail: takatsuy@jamstec.go.jp

JAMSTEC の船舶を用いた研究航海等で取得している各種データには研究資金配分機関や共同研究等との個別契約が存在するケースがあり、取得したデータは契約等に基づく適切な管理が必要である。JAMSTEC 情報管理部署では、データサイトからの公開において、JAMSTEC データポリシーを始めとする各種契約の有無の確認を行っている。これら航海データマネジメントのさらなる効率化に向け、Google Forms を用いた試みについて紹介する。

Data obtained through research cruises using JAMSTEC vessels may be subject to individual contracts, such as those related to a funding agency or joint research. The data must be managed appropriately in accordance with these contracts. The JAMSTEC Information Management Division confirms the existence of these contracts, including the JAMSTEC Data Policy, when publishing data from the data site. This article introduces an initiative to further streamline data sample information management of using Google Forms.

キーワード: 研究データ, データ管理, DOI, Google フォーム, オープンサイエンス

Keywords : Research data, Data management, DOI, Google forms, Open science

## 1 はじめに

海洋研究開発機構（JAMSTEC）では、これまで研究航海等を実施し、数多くのデータ・サンプルを取得してきた。これらのデータ・サンプル情報等は、JAMSTEC のデータポリシー[1]に則り、「JAMSTEC 航海・潜航データ・サンプル探索システム（DARWIN）」[2]から航海・潜航のメタデータとともに継続して公開を行っている。

取得されたデータやサンプルについては規程類等の機構内ルール、条約や保護区等の機構外のルールだけでなく、研究資金配分機関（FA）や共同研究等の個別の契約が存在する等、公開要件の複雑化が進んでいる状況である[3]。なかでも、学術論文等の即時オープンアクセスの実現に向けた基本方針[4]では FAIR 原則[5]に沿ったオープンサイエンスの推進が求められるとともに、公的資金による研究データの管理・利用についてはオープン・アンド・クローズ戦略に基づいた管理・利活用[6]が必要となっている。

このような状況の下、JAMSTEC 情報管理部は、航海で取得したデータを適切に公開するためのオープン／クローズ対応や、増加するデータ公開数、複雑化する公開要件に対応するためのマネジメント効率化に取り組んでいる。

本稿では、航海に関するデータ公開について紹介するとともに、航海データマネジメント効率化に向けた一環として実施した、Google Forms を用いたデータの取扱いに関する情報管理改善の試みについて報告する。

## 2 航海に関するデータ公開

### 2.1 航海で取得可能なデータ・サンプル情報

航海では、船体に固定されている観測機器を用いて、研究者及び観測技術員の労力をあまり必要とせず、どの航海でも取得できる定常観測データ[7]をはじめ、研究者自身が観測のために持ち込む観測機器によって取得されるデータや、生物サンプル、岩石サンプル、堆積物コアサンプルのような実サンプルまで、様々なデータ及びサンプルを取得することが可能である。これらのデータ・サンプル情報は、データ・サンプルの取得者が優先的に利用できるよう公開猶予期間を設定することが可能である。また、個別の契約による公開猶予期間の設定や映像・画像に対する著作権表示等も指定することが可能である。

表 1 データ・サンプル情報の公開猶予期間

|       | データ・サンプルの種類                                         | 公開猶予期間                                               |
|-------|-----------------------------------------------------|------------------------------------------------------|
| メタデータ | 航海情報・潜航情報<br>クルーズサマリー<br>クルーズレポート                   | 2ヶ月                                                  |
| データ   | 海洋物理<br>海洋化学<br>固体地球<br>海上気象<br>生態系<br>・・・<br>映像・画像 | 定常観測項目<br>任意観測項目<br>潜水船・探査機データ<br>待込機器・サンプル分析<br>データ |
| サンプル  | 岩石サンプル<br>生物サンプル<br>堆積物コアサンプル                       | 2年あるいは5年 <sup>*2</sup>                               |

\*1：公開猶予期間の設定が可能    \*2：適用されるデータポリシーによって異なる

### 2.2 DOI の付与

利用者に対する検索性及び引用の利便性向上のため、JAMSTEC の研究航海等に関するデータとして、航海単位にまとめ

たデータ及びクルーズレポートに対して DOI の付与を行っている[8].

航海単位にまとめたデータは、航海のメタデータ、航跡図、定常観測データを含む観測データ、関連する潜航で構成され、上述の DARWIN[2]で公開される。

クルーズレポートとは、研究調査航海に関する成果の速報でありデータ・サンプル利用時や、新規研究航海の検討にあたって参考となる文書であることから、「クルーズレポート・データブックカタログ」[9]より公開されている。2025 年 6 月末時点で 1982 年以降の航海データ 2,600 件以上、1992 年以降のクルーズレポート 1,300 件以上に DOI を付与している。

### 3 航海で取得されたデータの取扱いに関する情報管理の改善

#### 3.1 これまでの情報管理

従来、航海で取得した定常観測データの公開猶予期間や取得したデータ・サンプルに関する取扱いの要件については、情報管理部署が中心となり航海に関係する研究者や契約を担当する各部署への問い合わせを行い、情報の収集を行ってきた。しかしながら、JAMSTEC では年間約 70～80 件の航海を実施しているだけでな

く、一つの航海に対して複数の契約が含まれることもあり、都度公開猶予期間やデータの取扱い要件を確認することは非常に多くのマンパワーと時間を要する業務となっていた。

また、1 章で述べたように、データ・サンプルのオープン/クローズを含むデータマネジメントについては以前にも増して研究者自身の意識醸成が不可欠な状況となっていた。

#### 3.2 Google Forms を用いた情報管理

3.1 の状況を踏まえ、研究者がデータ公開のトリガーを担うデータ管理の方策についての検討を行った。検討にあたっては、研究者が入力しやすい形式・項目とするため、2022 年から海洋物理・海洋化学・地質学・海洋工学・生物学等複数の研究分野の研究者及び関係部署に対するヒアリングも実施した。

また、定常観測データを論文等で使用する場合や、個別の取決めがある場合の公開猶予期間について、航海の責任者である首席研究者が乗船前に Google Forms へ入力・送信することとした新たなデータ・サンプル情報管理の運用を 2024 年 9 月から開始した。手順の概要を図 1 に示す。

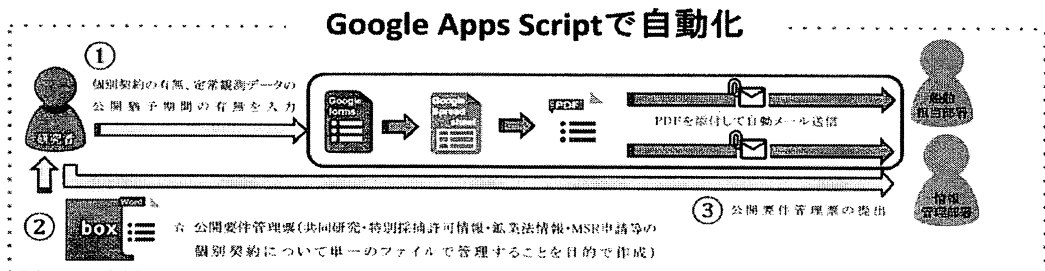


図 1 Google Forms を用いたデータ・サンプル情報管理。赤枠で囲った部分は自動処理している。

まず、首席研究者はGoogle Formsに定常観測データの公開猶予希望の有無と公開猶予を希望する場合の対象海域、並びに個別契約の有無について入力する。Google Forms送信後、Google Apps Scriptを用いてデータを視認性の高い形式[図 2]に整え、PDF化し、各部署担当者へメール送信する処理を自動化した。これにより、情報管理部署の担当者は自動送信されるメールに添付されたPDFを確認することで、当該航海における定常観測データの取扱い及び個別契約等の有無を容易に把握することが可能となった。個別契約が存在する場合は、別途MS Wordファイル「航海で取得されたデータ・サンプル利用／公開要件管理票」[10]を用いた管理を行うことで、研究者の情報入力及び情報管理部署の情報管理について効率化並びに省力化を実現した。管理票の具体的な入力項目については参考文献を参照されたい。

管理票は航海実施にあたり想定される契約内容を記載することが可能なフォーマットになっており、研究者には自身が首席として乗船する航海の契約について記入して提出することを求めている。これにより、首席研究者は自身が首席となる航海において取得されたデータをどのように取り扱うべきか再認識することができ、データ管理に係る意識醸成にも大きく寄与していると考えている。Google Formsを用いた効率化並びに管理表を用いた契約事項の整理を行うことで、効率化・省力化後も適切なデータ管理の継続が可能となっている。

図2 視認性向上のためGoogle Formsから入力された内容を自動で整形

### 3.3 まとめ

これまで述べたようなデータを取り巻く環境の変化に対して、限られた人員・予算の中で適切なデータ管理を行うため、本稿で述べたようなツールを用いたデータ管理の効率化・省力化は効果的である。

今後も適切なデータ公開を継続するため、運用を進める中でさらなるデータマネジメントの効率化を目指すとともに、引き続きツール及び管理票のブラッシュアップを行う予定である。

## データ利用可能性宣言

本稿の結果を裏付けるソースコードは一般には公開されていない。

しかし、合理的な理由があれば著者から入手することが可能である。

## 参考文献

[1] JAMSTEC データ・サンプルの取り扱いに関する基本方針（データポリシー）：

[https://www.jamstec.go.jp/j/database/data\\_policy.html](https://www.jamstec.go.jp/j/database/data_policy.html)（2025 年 7 月 15 日参照）。

[2] JAMSTEC 航海・潜航データ・サンプル探索システム（Data and Sample Research System for Whole Cruise Information in JAMSTEC）：

<https://www.godac.jamstec.go.jp/darwin/ja/>（2025 年 7 月 15 日参照）。

[3] 福田 和代：「JAMSTEC のデータポリシー、データマネジメントとベストプラクティス」，Japan Open Science Summit 2019, <https://jir.repo.nii.ac.jp/records/3617>

[4] 公的資金による研究データの管理・利活用に関する基本的な考え方：

[https://www.mext.go.jp/content/20210608-mxt\\_jyohoka01-000015787\\_06.pdf](https://www.mext.go.jp/content/20210608-mxt_jyohoka01-000015787_06.pdf)（2025 年 7 月 23 日参照）。

[5] Wilkinson, M. D. *et al.*: “The FAIR

Guiding Principles for scientific data management and stewardship.” *Sci Data* 3, 160018 (2016) 10.1038/sdata.2016.18.

[6] 学術論文等の即時オープンアクセスの実現に向けた基本方針：

[https://www8.cao.go.jp/cstp/oa\\_240216.pdf](https://www8.cao.go.jp/cstp/oa_240216.pdf)（2025 年 7 月 23 日参照）。

[7] 定常観測データの取得について：

[https://www.jamstec.go.jp/ceist/j/datasample/files/Routine\\_Observation\\_jp.pdf](https://www.jamstec.go.jp/ceist/j/datasample/files/Routine_Observation_jp.pdf)（2025 年 7 月 15 日参照）。

[8] JAMSTEC データ DOI について：

<https://www.jamstec.go.jp/datadoi/doi/10.17596/j/about.html>（2025 年 7 月 15 日参照）。

[9] クルーズレポート・データブックカタログ：

[https://www.godac.jamstec.go.jp/cr\\_catalog/view/base?lang=ja](https://www.godac.jamstec.go.jp/cr_catalog/view/base?lang=ja)（2025 年 7 月 15 日参照）。

[10] 航海で取得されたデータ・サンプル利用／公開要件管理票：

[https://www.jamstec.go.jp/ceist/j/datasample/files/JAMSTEC\\_Publish\\_Regulations\\_jp.docx](https://www.jamstec.go.jp/ceist/j/datasample/files/JAMSTEC_Publish_Regulations_jp.docx)（2025 年 7 月 17 日参照）。

エラータ

## コミュニケーションプロセスとしてのデジタルアーカイブの理論的基礎 付け: 基礎情報学の情報概念を中心に

### Laying the Theoretical Groundwork about “Digital Archive as a Communication Process”: Based on the Concept of “Information” in the Fundamental Informatics

棕本輔<sup>1 2\*</sup>, 上松大輝<sup>3</sup>, 山中司<sup>2</sup>

Tasuku MUKUMOTO<sup>1 2\*</sup>, Hiroki UEMATSU<sup>3</sup>, Tsukasa YAMANAKA<sup>2</sup>

1 デザイン及び学術研究教育業

Independent Design and Academic Research/Education Office

〒248-0013 神奈川県鎌倉市材木座5-15-12

E-mail: [tasuku.mukumoto@pie-ce.org](mailto:tasuku.mukumoto@pie-ce.org)

2 立命館大学

Ritsumeikan University

〒603-8577 京都府京都市北区等持院北町56-1

3 国立情報学研究所

National Institute of Informatics

〒101-8430 東京都千代田区一ツ橋2-1-2

\*連絡先著者 Corresponding Author

情報知識学会誌, 第35巻2号, p. 244

上記, 第33回年次大会予稿本文中に誤りがありましたので, 以下のとおり訂正いたします。

箇所: 左段 12行目

訂正前: 「伝搬」 訂正後: 「伝播」

箇所: 左段 17行目

訂正前: 「伝搬メディア」 訂正後: 「伝播メディア」

情報知識学会 2025 年度総会 議事録

1. 開催日

2025 年 5 月 31 日（土）11:30～12:30
2. 会 場

筑波大学(筑波キャンパス中地区)（対面開催）
3. 議 長

芦野会長
4. 議 事

1) 総会有効成立確認

長塚副会長・事務局

[資料 1]

出席者 22 名、委任状 31 通、計 53 名。依って定足数(正会員の 10 分の 1)を満たし、総会成立

2) 2024 年度事業報告：

原田副会長

[資料 2]

3) 2024 年度決算報告：

原田副会長

[資料 3]

4) 2024 年度監査結果報告：

根岸・田窪監事

[資料 4]

5) 2025 年度事業計画説明：

孫副会長

[資料 5]

6) 2025 年度予算説明：

孫副会長

[資料 6]

7) その他の報告

副会長／常務理事

[資料 7]

①2025 年度永年会員表彰：芦野会長

②第 22 回（2025）論文賞：江草選考委員

③名誉会員への推挙 芦野会長

上記の審議事項、報告および事業計画・予算案はいずれも原案通り承認された。

2025 年度総会資料

[資料 1] 会員数

| 会員種別    | 2024年3/31現在 |      | 2024年度入会 |     | 2024年度退会 |     | 2025年3/31現在 |      |
|---------|-------------|------|----------|-----|----------|-----|-------------|------|
| 正会員     | 192         |      | 12       | (2) | 12       | (2) | 192         |      |
| 学生会員    | 23          |      | 9        |     | 9        | (3) | 23          |      |
| ユース会員   | 7           |      | 1        | (1) | 1        |     | 7           |      |
| シニア会員   | 18          |      | 2        | (2) | 3        |     | 17          |      |
| 名誉会員    | 0           |      | 0        |     | 0        |     | 0           |      |
| 賛助会員(*) | 4           | (13) | 0        |     | 0        |     | 4           | (13) |
| 合計      | 244         |      | 24       |     | 25       |     | 243         |      |

入会者、退会者の()は会員種別変更者数で内数。正会員→シニア 2、学生→正会員 2、学生→ユース 1。\* 賛助会員は全て団体。()内数字は口数。賛助会員 1 口の年会費¥30,000

## [資料2] 2024 年度事業報告

### (1) 会議

- ・総会 2024 年 5 月 25 日(土)12:00~13:00 ハイブリッド開催
- ・理事会 (年 2 回以上開催)
- 第 1 回理事会: 2024 年 4 月 22 日(月) 18:30~19:30 オンライン開催
- 第 2 回理事会: 2024 年 6 月 10 日(月) 18:30~19:30 オンライン開催
- ・常務理事会
- 第 1 回常務理事会: 2024 年 7 月 5 日(金) 19:00~19:50 オンライン開催
- 第 2 回常務理事会: 2024 年 11 月 25 日(月) 18:30~19:15 オンライン開催
- 第 3 回常務理事会: 2025 年 3 月 5 日(月) 19:00~20:00 オンライン開催
- ・論文賞推薦委員会 2025 年 2 月-4 月 メール開催
- ・監査 2025 年 4 月 14 日(月) アドスリー社で監査実施

### (2) 第 32 回 (2024 年度) 年次大会 (江草由佳 実行委員長)

本年度の年次大会は 5 月 25 日(土)と 26 日(日)の 2 日間に渡り、慶應義塾大学 (三田キャンパス) にて、研究発表会、総会が対面形式で開催されました。

研究発表は、一般セッション・学生セッション合わせて 14 件でした。例年同様に多様な分野からの報告がありました。そして、42 名(会員 38 名、非会員 4 名)の参加者が会場にお越しいただき、非常に活発な議論が行われました。

大会 1 日目は、研究発表と総会が行われました。また、夜には懇親会も行われ会員間の交流を深めました。

大会 2 日目は、研究発表と論文賞授賞式・記念講演、および、永年会員表彰が行われました。第 21 回論文賞(2024)は、坂本さん、中村さん、山本さん、田中さんの「マイクロブログユーザの類語に着目した地域属性の推定に関する研究」が受賞し、授賞式および坂本さんによる記念講演が行われました。その後、永年会員表彰式が行われました。また、学生奨励賞は、学生奨励賞審査委員会にて審査を行った結果、石引花明さん(筑波大学)の「動物園・水族館の学校向けプログラムに関する情報公開状況の調査」が選出されました。

オンライン発表・視聴がない完全に対面のみの開催としては久々の大会となりました。発表件数としては例年にくらべやや少なめとはなりましたが、その分、すべての研究発表セッションをシングルセッションとして開催することができ、大会参加者が同じ発表を聞きディスカッションできる良いの機会となったかと存じます。開催にあたっては、学会役員の先生方、実行委員ならび、学生奨励賞選考委員の皆様、学会事務局および編集委員会事務局、そしてご参加・ご発表および座長の皆様方から、多大なご協力をいただきました。誠にありがとうございました。

ぜひ来年度もご発表、ご参加くださいますよう、どうぞよろしくお願い申し上げます。

### (3) 学会誌編集委員会 (常川真央編集委員長)

#### 1. 学会誌発行

- ・情報知識学会誌 34 巻 2 号(研究報告会論文集) 2024 年 5 月 26 日発行
- ・情報知識学会誌 34 巻 3 号(通常論文) 2024 年 9 月 30 日発行
- ・情報知識学会誌 34 巻 4 号(フォーラム特別号) 2024 年 11 月 30 日発行
- ・情報知識学会誌 35 巻 1 号(通常論文) 2024 年 2 月 28 日発行

#### 2. 2024 年度情報知識学会論文賞授与(孫媛論文賞推薦委員長)

3. 研究論文 6 件、事例/調査報告 0 件、解説/展望 0 件、部会報告 2 件、予稿 31 件を掲載

#### (4) 広報関連

##### 1. メールマガジン(メールマガジン編集室編集長村井源、副編集長岡本由起子)

・2024 年 4 月から 2025 年 3 月まで、8 月休刊を除き、第 190 号から第 200 号まで、11 回(4 月号中川、5 月号常川、6 月号宮本、7 月号村川、9 月号田良島、10 月号高久、11 月号梶川、12 月号孫、1 月号編集部、2 月号長田、3 月号阪口 各理事の方々の編集担当により)発刊されました。その他に、臨時号を 5 回(5 月 15 日、8 月 13 日、9 月 4 日、11 月 15 日、2 月 17 日)配信しました。

・メールマガジン・アーカイブ(2024 年 1 月号-2024 年 12 月号)は、情報知識学会誌 Vol. 35 No. 1(Feb. 2025)に掲載されています。

##### ・記事の内容

例年に倣って、各回の常務理事会・理事会の議事概要、年次大会、情報知識学会フォーラム、論文賞、など、本学会定例行事のお知らせと報告、学会誌、各部会の研究会等の活動後援・共催行事、その他関連団体の行事、事務局からのお知らせなどでした。24 年度は、この他、臨時号として、年次大会・フォーラムの開催や発表募集など急ぎの会員連絡が行われました。

##### 2. ホームページ(江草由佳常務理事)

2024 年 4 月~2025 年 3 月に随時更新を実施。

3. ジャパン・オープンサイエンス・サミット 2024 (JOSS2024) は、主催機関は本学会に加え、国立情報学研究所、科学技術振興機構、物質・材料研究機構、科学技術・学術政策研究所、情報通信研究機構、学術資源リポジトリ協議会、合同会社 AMANE で、2024 年 6 月 17 日(月)~ 6 月 21 日(金)の会期でオンライン開催された。

#### (5) 第 29 回情報知識学フォーラム(高田良宏 実行委員長)

第 29 回情報知識学フォーラムは、2024 年 11 月 30 日(土)に石川県金沢市の金沢未来のまち創造館にて対面形式で開催された。テーマに沿ったシンポジウムとポスター発表が行われ、参加者 33 名(会員 15 名、非会員 17 名)が会場にお越しいただき、非常に活発な議論が行われた。

前半のシンポジウム「研究データエコシステム × 地域資料の保存・継承」~災害を乗り越え地域資料継承に貢献する研究データエコシステムの未来~ は、研究データエコシステムを研究データとしての地域資料の共有・利活用基盤としてのみではなく、減災、縮災に資する情報共有基盤としてどのような可能性と展開が期待できるだろうかという問いに対して、研究データエコシステム事業、ユースケース創出課題、被災資料レスキュー、および、政策に関わる方々にご講演頂き、取り組みの具体例や現状での課題を紹介していただいた。講演の後の総合討論では、講演者と参加者との間で活発な意見交換もでき、参加者全員で非常に有意義な時間を共有することができた。

後半の研究発表(ポスター発表)では、12 名の発表者が、それぞれ 90 秒間ポスター概要を発表した後、約 45 分間のポスターコアタイムに臨んだ。講演、総合討論からの熱気が引き継がれ、終了時刻に続く情報交換会のために皆さんの会話をさえぎるのが心苦しいほどであった。

#### (6) 部会活動

##### ・CODATA 部会(岩田修一部会長)

2024 年度は投稿内容を基に、データに関する記号、言語に関する基礎的な課題についての意見交換のための座談会を 2 回開催し、論点整理を開始した。

・人文・社会科学系部会（田良島哲部会長）

2024 年度は、独自の企画を実施するには至らなかった。

・専門用語研究部会（長田孝治部会長）

1. 専門用語シンポジウムの開催

2024 年度はシンポジウムの開催は出来ませんでした。

2. 国際協力

TC37 (Language and terminology) の SC2 および SC5 の総会が 2024 年 6 月 3 日から 1 週間の日程で 2023 年に引き続きベルギー（ブリュッセル）の EU 議会を借りてハイブリッド形式で開催されました。用語関連では、ISO 10241-2:2012「規格における用語エントリ 第 2 部：標準化された用語エントリの採用」が改定について検討され、オーストリアと日本で共同のプロジェクトリーダーとして進めることが再確認されました。この関連で、2025 年 2 月にマニサレス（コロンビア）で担当者会議が開催されました。この改訂に対して専門用語研究部会や有識者で会合を行っています。

3. 国内協力

ISO/TC37 の国内事務局が規格協会から情報科学技術協会に移動し、日本からも新規&改定 ISO 提案を行っているので、専門用語研究部会として協力を行っています。

・関西部会（村川猛彦 部会長）

第 14 回知識・芸術・文化情報学研究会（2025 年 3 月 22 日、ハイブリッド開催）を、アート・ドキュメンテーション学会関西地区部会と共催した。

・シニア情報知識学研究部会（長塚隆 代表世話人）

2024 年度は、4 月に世話人会を対面で開催し、2024 年度第 1 回卓話会（通算第 22 回）話題提供者：植村 八潮氏（専修大学文学部ジャーナリズム学科 教授）を年 6 月 5 日（水）15：00-16：30 に、紀伊國屋書店 新宿本店 3F アカデミックラウンジで、「電子出版 40 年個人史—CD-ROM から電子図書館まで」と、第 2 回卓話会（通算第 23 回）話題提供者：田良島 哲氏（東京文化財研究所 客員研究員）「インターネット以前のネットワークコミュニティ」を 11 月 14 日（木）15：00-16：30 に、紀伊國屋書店 新宿本店 3F アカデミックラウンジでハイブリッド開催した。

・オープンサイエンス・オープンデータ (OS-OD) 研究部会（高田良宏 世話人代表）

・当学会が主催団体の一つである Japan Open Science Summit 2024 (JOSS2024)（オンライン、6/17~6/21）において、同研究部会世話人が中心となり企画・運営に貢献した。また、JOSS2024 のセッション報告および個人発表予稿（希望者）をとりまとめ、JOSS 特集として情報知識学会誌第 34 巻第 3 号で出版した。

・第 29 回情報知識学フォーラム（金沢市、11/30）を同研究部会世話人が中心となり、「研究データエコシステム × 地域資料の保存・継承」～災害を乗り越え地域資料継承に貢献する研究データエコシステムの未来～というテーマで実施した。

・その他、研究部会世話人は OS・OD を推進すべく、各々が研究データ利活用協議会 (RDUF)、大学 ICT 推進協議会 (AXIES) RDM 部会、NII 研究データ基盤運営委員会システム作業部会等に参加するなどの活動を継続して行った。

（7）論文賞

会員および推薦委員による推薦に基づき、

1. 海外学術情報基盤が提供する機関向けダッシュボードの比較分析

池谷 瑠絵, 大波 純一, 金沢 輝一, 西岡 千文, 高久 雅生, 山地 一禎  
情報知識学会誌, Vol.34, No.1

2. 物語展開の共通因子とジャンルの関係性

村井 源, 青山 美月, 大田 翔貴, 大場 有紗, 福元 隆希, 齊藤 悠仁, 佐藤 央一  
情報知識学会誌, Vol. 34, No. 3

を論文賞候補として、会員投票を行うこととした。

2025 年度年次大会において、受賞者の発表と記念講演を行う

(8) 後援行事等

1. TEI (Text Encoding Initiative) 東アジア／日本語 SIG 勉強会 オンライン開催
2. 2024 年度アート・ドキュメンテーション学会年次大会 6 月 15-16 日
3. 第 13 回国際シンポジウム JADH2024 9 月 18-20 日 3 日間
4. Code4Lib JAPAN カンファレンス 2024 9 月 7-8 日 オンライン開催
5. 第 33 回 整理技術・情報管理等研究集会 「TP&D フォーラム 2024」9 月 14 日 オンライン開催
6. デジタルアーカイブ学会第 9 回研究大会 11 月 1~2 日 東京大学本郷キャンパス
7. 「人文科学とコンピュータシンポジウム (じんもんこん 2024)」12 月 7~8 日  
東北大学川内南キャンパス

## [資料3] 2024年度決算

令和 7年3月31日現在

(単位:円)

## 収入の部

| 科目       | 細目      | 2024年度予算  | 予算細目      | 2024年度実績  | 実績細目      | 備 考                    |
|----------|---------|-----------|-----------|-----------|-----------|------------------------|
| 会費       | 正会員     | 2,160,000 | 1,470,000 | 2,062,000 | 1,470,000 | 年会費¥8,000              |
|          | 学生会員    |           | 150,000   |           | 86,000    | 年会費¥4,000              |
|          | ユース会員   |           | 20,000    |           | 44,000    | 年会費¥4,000              |
|          | シニア会員   |           | 100,000   |           | 72,000    | 年会費¥4,000              |
|          | 賛助会員    |           | 420,000   |           | 390,000   | 一口¥30,000              |
|          |         |           |           |           |           |                        |
| 参加費      | 年次大会    | 160,000   | 90,000    | 154,000   | 136,000   | 32回 非会員参加費/懇親会費/原稿超過買代 |
|          | フォーラム   |           | 70,000    |           | 18,000    | 第29回 超過買代              |
|          | セミナー    |           | 0         |           | 0         |                        |
|          |         |           |           |           |           |                        |
| 売上金      | 学会誌     | 320,000   | 130,000   | 262,252   | 130,803   | 定期購読・バックナンバー           |
|          | 別刷り     |           | 60,000    |           | 45,000    |                        |
|          | 著作権料    |           | 130,000   |           | 86,449    | 学術著作権協会:著作権料(国内・海外)    |
| その他の収入   | 預金利息    | 60        | 60        | 523       | 523       |                        |
|          | 寄付金・その他 |           |           |           | 0         |                        |
| 小計       |         | 2,640,060 | 2,640,060 | 2,478,775 | 2,478,775 |                        |
| 前年度繰越金   |         | 2,583,517 | 2,583,517 | 2,583,517 | 2,583,517 |                        |
| 合計       |         | 5,223,577 | 5,223,577 | 5,062,292 | 5,062,292 |                        |
| 特別会計     |         | 0         | 0         | 0         | 0         |                        |
| 特別会計 積立金 |         | 2,401,928 | 2,401,928 | 2,401,969 | 2,401,969 |                        |
| 総計       |         | 7,625,505 | 7,625,505 | 7,464,261 | 7,464,261 |                        |

## 支出の部

(単位:円)

| 科目       | 細目        | 2024年度予算  | 予算細目      | 2024年度実績  | 実績細目      | 備 考                                 |
|----------|-----------|-----------|-----------|-----------|-----------|-------------------------------------|
| 印刷費      | 学会誌       | 840,000   | 800,000   | 882,920   | 794,200   | 34巻2,3,4号 35巻1号 2024年論文賞受賞記念別刷      |
|          | 選挙        |           | 0         |           | 0         |                                     |
|          | その他       |           | 40,000    |           | 88,720    | 総会案内/長3角4封筒印刷                       |
| 人件費      | 事務局       | 340,000   | 0         | 330,990   | 0         |                                     |
|          | 編集事務局     |           | 300,000   |           | 297,550   | J-STAGE登載作業・Webデータ作成作業              |
|          | HP管理      |           | 40,000    |           | 33,440    | Webページ更新作業                          |
| 事業費      | 第32回年次大会  | 320,000   | 150,000   | 348,531   | 254,645   | 第32回年次大会 会場費、懇親会費、雑費                |
|          | 第29回フォーラム |           | 150,000   |           | 82,750    | 発表原稿等の受付・管理/会場費                     |
|          | 部会補助金     |           | 20,000    |           | 0         |                                     |
|          | その他       |           | 0         |           | 11,136    | シニア部会卓話会講演謝金                        |
| 会議費      | 理事会       | 20,000    | 10,000    | 0         | 0         |                                     |
|          | 常務理事会     |           | 10,000    |           | 0         |                                     |
|          | その他       |           | 0         |           | 0         |                                     |
| 業務委託費    |           | 821,700   | 821,700   | 821,700   | 821,700   | (株)アドスリー / @68,475 * 12ヶ月           |
| 賃借料      |           |           | 0         |           | 0         |                                     |
| 通信費      | 学会誌発送費    | 205,000   | 110,000   | 180,524   | 87,142    | 網ワコー                                |
|          | 編集通信費     |           | 0         |           |           |                                     |
|          | 総会        |           | 40,000    |           | 36,897    | 総会案内発送費                             |
|          | 選挙        |           | 0         |           | 0         |                                     |
|          | 電話代       |           | 0         |           | 0         |                                     |
|          | インターネット   |           | 25,000    |           | 24,838    | さくらネット、MOVFAX                       |
|          | その他       |           | 30,000    |           | 31,647    | 郵便切手代、宅急便代                          |
| 交通費      | 役員旅費      | 10,000    | 0         | 0         | 0         |                                     |
|          | 事務局通勤費    |           | 0         |           | 0         |                                     |
|          | その他交通費    |           | 10,000    |           | 0         |                                     |
| 消耗品費     |           | 10,000    | 10,000    | 8,865     | 8,865     | ラベルシール、賞状額                          |
| 維持修繕費    |           | 0         | 0         | 0         | 0         |                                     |
| 雑費       | 手数料       | 60,000    | 30,000    | 67,593    | 41,413    | 振込手数料: 年会費郵便振替料の学会負担含む              |
|          | 法定調書作成    |           | 30,000    |           | 26,180    | 合計表作成提出1式20,000 調書作成個人送付3名@1,800+送料 |
|          | その他       |           |           |           |           |                                     |
| 小計       |           | 2,626,700 | 2,626,700 | 2,641,123 | 2,641,123 |                                     |
| 予備費      |           | 2,583,517 | 2,583,517 | 0         | 0         |                                     |
| 次年度繰越金   |           | 0         | 0         | 2,421,169 | 2,421,169 | 2024年度収支差額 -162,348円                |
| 合計       |           | 5,210,217 | 5,210,217 | 5,062,292 | 5,062,292 |                                     |
| 特別会計     |           | 0         | 0         | 0         | 0         |                                     |
| 特別会計 積立金 |           | 2,401,928 | 2,401,928 | 2,401,969 | 2,401,969 |                                     |
| 総計       |           | 7,612,145 | 7,612,145 | 7,464,261 | 7,464,261 |                                     |

貸借対照表  
2025年3月31日現在

(単位:円)

| 科目     | 借方        | 貸方        | 備考                                                                                        |
|--------|-----------|-----------|-------------------------------------------------------------------------------------------|
| 1.資産の部 |           |           |                                                                                           |
| 現金     | 0         |           |                                                                                           |
| 普通預金   | 1,159,537 |           | 三菱東京UFJ銀行                                                                                 |
| 定期預金   | 2,401,969 |           | 三菱東京UFJ銀行                                                                                 |
| 振替貯金   | 983,499   |           | 振替00150-8-706543                                                                          |
| 振替貯金   | 343,133   |           | 振替00130-9-706558                                                                          |
| 2.負債の部 |           |           |                                                                                           |
| 前受金    |           | 65,000    | 2025年度の年会費10件(正会員2名.シニア会員8件)<br>2026年度の年会費(ユース会員3件)<br>過入金分正会員1件1000円<br>2027年度(ユース会員各1件) |
| 3.特別会計 |           |           |                                                                                           |
| 積立金    |           | 2,401,969 |                                                                                           |
| 4.正味財産 |           | 2,421,169 |                                                                                           |
| 合計     | 4,888,138 | 4,888,138 |                                                                                           |

[資料4] 監査報告書

監査報告書

情報知識学会  
会長 芦野 俊宏 殿

作成日 2025 年 4 月 14 日

情報知識学会 監事 根岸正光

情報知識学会 監事 田窪直規

当監事は、情報知識学会の2024年4月1日から2025年3月31日までの財務諸表について、監査を行いました。この監査にあたって、当監事は一般に公正妥当と認められる監査基準に準拠し、通常実施すべき監査手続きを実施致しました。

監査の結果、当監事は上記財務諸表が2025年3月31日現在の情報知識学会の財務状況を適正に表示しているものと認めます。

## [資料5] 2025年度事業計画

### (1) 会議

- ・ 総会 2025年5月31日(土) 11:30-12:30  
会場：対面開催
- ・ 理事会 年2回以上開催：2025年4月22日、第2回 6月初旬（オンライン開催）
- ・ 常務理事会 随時開催
- ・ 論文賞推薦委員会 年2回開催
- ・ 監査会 2025年4月14日（月）（学会事務委託先アドスリー社にて開催）

### (2) 第33回(2025年度)年次大会（宮川創実行委員長）

本年度は情報知識学会第33回（2025年度）年次大会は、5月31日（土）および、6月1日（日）の両日にわたって、筑波大学（筑波キャンパス）において、対面形式で総会、研究報告会、第22回（2025）論文賞発表・記念講演、永年会員発表、学生奨励賞発表を実施する。研究報告会は、一般セッション14件、学生セッション10件の計24件を予定しており、学生セッションでは、審査を行い学生奨励賞受賞者を決定する。なお、懇親会を実施する。（URL:<http://www.jsik.jp/?2025program>）

### (3) 学会誌編集委員会（常川真央編集委員長）

1. 学会誌発行
  - ・ 情報知識学会誌35巻2号（研究報告会論文集）2025年5月発行予定
  - ・ 情報知識学会誌35巻3号（通常論文）2025年9月発行予定
  - ・ 情報知識学会誌35巻4号（フォーラム特別号）2025年11月発行予定
  - ・ 情報知識学会誌36巻1号（通常論文）2026年2月発行予定
2. 2025年度情報知識学会論文賞授与

### (4) 広報関連

1. メールマガジン（メールマガジン編集室編集長 村井源、副編集長 岡本由起子）
  - ・ 各号の発刊
    - \*8月を除く、毎月の刊行を予定。
    - \*今年度も各号の編集には、理事各位のご協力を仰ぎ、定期的配信を期す。
    - \*配信はアドスリーに依頼（毎月26日、土日を除く）。
  - ・ 臨時号については、編集室とは別の体制で臨むこととする。
2. ホームページ（江草由佳常務理事）  
2024年4月～2025年3月随時更新
3. 第30回情報知識学フォーラムやジャパン・オープンサイエンス・サミット2025（JOSS2025）では、積極的に関連学協会などに後援依頼などを行い広報に務める。

### (5) 部会活動

- ・ CODATA部会（岩田修一部会長）  
2025年度は、いくつかの課題を選定してデータ活動に関する座談会を開催し、座談会報告として投稿する予定である。
- ・ 人文・社会科学系部会（田良島哲部会長）  
関連する他学会、団体と連携し、特に人文学系の企画の実現を図る。
- ・ 専門用語研究部会（長田孝治部会長）
  1. 専門用語関連会議の開催  
EAFTerm（東アジア専門用語フォーラム）については本年度も各国の動きを見ながら対応を行います。

## 2. 国際協力 ISO/TC37（用語）総会への参加

2025年度はTC37全体での総会は、日本の高松で6月22日から一週間の日程で開催予定です。日本からISO 10241-2:2012の改訂提案が行われているので部会として協力を行います。また、各国から100名程度の参加が見込まれるので、サポートを行う予定でいます。

## 3. 国内協力

ISO/TC37の国内事務局である情報科学技術協会に対して専門用語研究部会として協力を行います。また、言語コードに関するISO規格が改定となりましたが、JISで扱うための幾つかの課題が発生したため、その解決に向けてのサポートを行います。

### ・ 関西部会（村川猛彦部会長）

知識・芸術・文化情報学研究会を今年度も実施し、関西部会の大会的なものと位置づけたい。その他、関連学協会と連携しながら、学会員が興味を持ちそうな行事の共催・後援を行う予定である。

### ・ シニア情報知識学研究部会（長塚隆代表世話人）

2025年4月に世話人会を開催し、年2回の卓話会を実施の予定である。第1回は6月25日（水）に理事の中川 修氏を話題提供者に迎えて開催予定である。

### ・ オープンサイエンス・オープンデータ (OS-OD) 研究部会（高田良宏 世話人代表）

・ 部会企画のシンポジウム／研究会を実施する。関連学協会との共同企画にも積極的に考える。

・ 当学会が主催団体の一つであるJapan Open Science Summit 2025 (JOSS2025)の企画・運営に貢献する。また、JOSS2025のセッション報告や個人発表予稿などをOS-OD研究部会のJOSS特集として情報知識学会誌第35巻第3号で出版する。

・ その他、OS・ODを推進すべく、学会内外で活動を行う。また、世話人を増やす。

## （6）第30回情報知識学フォーラム開催

同志社大学で開催する予定で準備が進められている。

## （7）論文賞

前年度に引き続き、候補の推薦と会員投票を行い、論文賞を選定する。

## （8）新規会員の獲得

対面での年次大会などの開催も可能となり、学会紹介パンフレットや学会クリアファイルなども活用し学会行事や関連学会などの機会を利用し学会の紹介に努める。

## [資料6] 2025年度予算

| 収入の部     |         |           |           |           |           | (単位:円)                 |
|----------|---------|-----------|-----------|-----------|-----------|------------------------|
| 科目       | 細目      | 2025年度予算  | 予算細目      | 2024年度実績  | 実績細目      | 備 考                    |
| 会費       | 正会員     | 2,170,000 | 1,500,000 | 2,062,000 | 1,470,000 | 年会費¥8,000              |
|          | 学生会員    |           | 150,000   |           | 86,000    | 年会費¥4,000              |
|          | ユース会員   |           | 50,000    |           | 44,000    | 年会費¥4,000              |
|          | シニア会員   |           | 80,000    |           | 72,000    | 年会費¥4,000              |
|          | 賛助会員    |           | 390,000   |           | 390,000   | 一口¥30,000              |
| 参加費      | 年次大会    | 160,000   | 90,000    | 154,000   | 136,000   | 32回 非会員参加費/懇親会費/原稿超過頁代 |
|          | フォーラム   |           | 70,000    |           | 18,000    | 第29回 超過頁代              |
|          | セミナー    |           | 0         |           | 0         |                        |
| 売上金      | 学会誌     | 320,000   | 130,000   | 262,252   | 130,803   | 定期購読・バックナンバー           |
|          | 別刷り     |           | 60,000    |           | 45,000    |                        |
|          | 著作権料    |           | 130,000   |           | 86,449    | 学術著作権協会:著作権料(国内・海外)    |
| その他の収入   | 預金利息    | 500       | 500       | 523       | 523       |                        |
|          | 寄付金・その他 |           |           |           | 0         |                        |
| 小計       |         | 2,650,500 | 2,650,500 | 2,478,775 | 2,478,775 |                        |
| 前年度繰越金   |         | 2,421,169 | 2,421,169 | 2,583,517 | 2,583,517 |                        |
| 合計       |         | 5,071,669 | 5,071,669 | 5,062,292 | 5,062,292 |                        |
| 特別会計     |         | 0         | 0         | 0         | 0         |                        |
| 特別会計 積立金 |         | 2,401,969 | 2,401,969 | 2,401,969 | 2,401,969 |                        |
| 総計       |         | 7,473,638 | 7,473,638 | 7,464,261 | 7,464,261 |                        |

| 支出の部     |           |           |           |           |           | (単位:円)                                   |
|----------|-----------|-----------|-----------|-----------|-----------|------------------------------------------|
| 科目       | 細目        | 2025年度予算  | 予算細目      | 2024年度実績  | 実績細目      | 備 考                                      |
| 印刷費      | 学会誌       | 850,000   | 800,000   | 882,920   | 794,200   | 35巻2.3.4号 36巻1号 2025年論文賞受賞記念別刷           |
|          | 選挙        |           | 20,000    |           | 0         | 2026-2027選挙公示印刷                          |
|          | その他       |           | 30,000    |           | 88,720    | 総会案内/長3角4封筒印刷                            |
| 人件費      | 事務局       | 340,000   | 0         | 330,990   | 0         |                                          |
|          | 編集事務局     |           | 300,000   |           | 297,550   | J-STAGE登載作業・Webデータ作成作業                   |
|          | HP管理      |           | 40,000    |           | 33,440    | Webページ更新作業                               |
| 事業費      | 第33回年次大会  | 320,000   | 150,000   | 348,531   | 254,645   | 第33回年次大会 会場費、懇親会費、管理サポート業務/G-Links       |
|          | 第30回フォーラム |           | 150,000   |           | 82,750    | 発表原稿等の受付・管理//G-Links、会場費                 |
|          | 部会補助金     |           | 20,000    |           | 11,136    | シニア部会卓話会講演謝金                             |
|          | その他       |           | 0         |           | 0         |                                          |
| 会議費      | 理事会       | 20,000    | 10,000    | 0         | 0         |                                          |
|          | 常務理事会     |           | 10,000    |           | 0         |                                          |
|          | その他       |           | 0         |           | 0         |                                          |
| 業務委託費    |           | 821,700   | 821,700   | 821,700   | 821,700   | (株)アドスリー /@68,475 * 12ヶ月                 |
| 賃借料      |           |           | 0         |           | 0         |                                          |
| 通信費      | 学会誌送達費    | 215,000   | 90,000    | 180,524   | 87,142    | 会誌送達費(糊ワコー)                              |
|          | 編集通信費     |           | 0         |           |           |                                          |
|          | 総会        |           | 40,000    |           | 36,897    | 総会案内送達費                                  |
|          | 選挙        |           | 30,000    |           | 0         | 投票用紙送料等                                  |
|          | 電話代       |           | 0         |           | 0         |                                          |
|          | インターネット   |           | 25,000    |           | 24,838    | さくらネット、MOVFAX                            |
|          | その他       |           | 30,000    |           | 31,647    | 郵便切手代、宅急便代                               |
| 交通費      | 役員旅費      | 10,000    | 0         | 0         | 0         |                                          |
|          | 事務局通勤費    |           | 0         |           | 0         |                                          |
|          | その他交通費    |           | 10,000    |           | 0         |                                          |
| 消耗品費     |           | 10,000    | 10,000    | 8,865     | 8,865     | ラベルシール、賞状額                               |
| 維持修繕費    |           | 0         | 0         | 0         | 0         |                                          |
| 雑費       | 手数料       | 60,000    | 30,000    | 67,593    | 41,413    | 振込手数料:年会費振替替料の学会負担含む                     |
|          | 法定調書作成    |           | 30,000    |           | 26,180    | 合計表作成提出1式20,000 調査作成個人送付3名@1,800+送料      |
|          | その他       |           |           |           |           |                                          |
| 小計       |           | 2,646,700 | 2,646,700 | 2,641,123 | 2,641,123 |                                          |
| 予備費      |           | 2,421,169 | 2,421,169 | 0         | 0         |                                          |
| 次年度繰越金   |           | 0         | 0         | 2,421,169 | 2,421,169 | 2024年度収支差額 -162,348円 2025年度予算案収支差額3,800円 |
| 合計       |           | 5,067,869 | 5,067,869 | 5,062,292 | 5,062,292 |                                          |
| 特別会計     |           | 0         | 0         | 0         | 0         |                                          |
| 特別会計 積立金 |           | 2,401,969 | 2,401,969 | 2,401,969 | 2,401,969 |                                          |
| 総計       |           | 7,469,838 | 7,469,838 | 7,464,261 | 7,464,261 |                                          |

## 〔資料 7〕 その他の報告

### ① 2025年度永年会員表彰

永年会員表彰：対象者 3名 大山 健治（オオヤマケンジ）、深澤 克朗（フカサワカツロウ）、土山 玄（ツチヤマゲン）（敬称略、会員期間10年以上）

情報知識学会第33回（2025年度）年次大会永年会員表彰式

2025年6月1日（日）13:10-14:00 論文賞授賞式・記念講演、永年会員表彰式中で行われます。 当日欠席の表彰対象者には後日表彰状を郵送させて頂いた。

### ② 第 22 回（2025）論文賞

規定による候補の推薦及び投票を行った結果、

「海外学術情報基盤が提供する機関向けダッシュボードの比較分析」池谷 瑠絵，大波 純一，金沢 輝一，西岡 千文，高久 雅生，山地 一禎. 情報知識学会誌, Vol. 34, No. 1, p. 3-17, 2024

を受賞論文とすることに決定した。

大会 2 日目の 6 月 1 日（日）に 13:00 より授賞式および池谷瑠絵氏による記念講演が行われた。

### ③ 2025年度第1回理事会で、石塚英弘元会長の名誉会員への推挙が承認された。

推挙理由： 創立時からの正会員である石塚英弘元会長より、引退・退会のお申し出があり、昨年の年次大会では長年の功績に鑑み感謝状をお送りしていることもあり、名誉会員に推挙された。 定款(7)名誉会員：本会の諸活動において、特別の功績があり、理事会の議決を経て推薦された者

以上

第 30 回情報知識学フォーラム

文化と社会をとらえるデータサイエンスの最前線

データサイエンスは、テキスト・画像・音声・時空間情報といった多様なデータを横断的に結びつけ、文化と社会の「いま」を高解像度でとらえる新たな視点を切り拓いています。人文社会科学の理論と計算的方法論を結びつけることにより、地域文化の可視化、芸術表現の計量分析、メディアと消費行動の変容、文化資本と不平等など、従来の方法論では、見落としされがちだったパターンや因果の手がかりが見えてきました。生成 AI を含む最新の機械学習、自然言語処理、地理情報解析は、文化資源の保存と継承や政策立案にも新たな選択肢をもたらししています。本フォーラム「文化と社会をとらえるデータサイエンスの最前線」では、国内外の実践者による最先端の研究を共有し、方法論（再現可能性、妥当性）、基盤整備（RDM、FAIR、オープンサイエンス）、倫理（ガバナンス、バイアス、著作権）までを横断して議論します。学術・公共・産業の協働を視野に、社会実装へ向けた課題と可能性を具体的に描き出すことを目指します。

また、ポスターセッションの時間を設け、フォーラムテーマに関わる研究・実践の発表を募集・紹介いたします。分野横断の対話を通じて、新たな共同研究やプロジェクトの芽を育む機会となれば幸いです。皆さまのご参加を、心よりお待ちしております。

日 時：2026 年 1 月 31 日（土）  
会 場：同志社大学 大阪サテライト・キャンパス  
（大阪市北区梅田 1-12-17 JRE 梅田スクエアビル 17 階）  
対面開催（オンライン参加はありません）  
主 催：情報知識学会

スケジュール：

2025 年 09 月 19 日（金） 研究（ポスター）発表申込開始  
2025 年 10 月 31 日（金） 研究（ポスター）発表申込期限  
2025 年 11 月 11 日（火） 採択可否連絡（予定）  
2025 年 11 月 11 日（火） 参加申込開始（予定）  
2025 年 12 月 12 日（金） 発表原稿提出期限 12:00（予定）

実行委員長 河瀬 彰宏（同志社大学）

## 事務局からのお知らせ

### 〔1〕 個人会員の皆様へ、2025年度の会費納入のお願い

2025年度会費未納入の方は、下記の学会口座へ早めにお振込みをお願いいたします。1年分の年会費は正会員 8 千円、学生会員・ユース会員・シニア会員は 4 千円です。過去数年分未納の方は合計額を納入くださるようお願いいたします。請求書が必要な方はその旨、情報知識学会事務局にメールでお知らせください。

#### (1) 振込先

- 1) 郵便振替口座 00150-8-706543 情報知識学会
- 2) ゆうちょ銀行 ○一九店(ゼロイチキョウ店) 当座 0706543 情報知識学会

#### (2) 会費の納入年月の確認方法

お手元に届いた学会誌の封筒の宛名ラベルには、ご自分の年会費の納入日が年度毎に西暦下 2 桁、月(2 桁)、日(2 桁)の 6 桁の数字で印字されています。会費未納年度には〔未納〕と表示されております。なお、お振り込みの後、事務局に通知が届き、宛名ラベルに印字、発送するまで10日ほどかかりますので、ご了承ください。

#### (3) 課税区分

不課税ですので消費税はかかっておりません。また、適格請求書発行事業者登録番号につきましては、当学会は登録しておりません。必要でしたら、経理処理のご担当者以上に 2 点をお伝えください。

### 〔2〕 学会誌送付先、会員種別、メールアドレスの変更について

会員種別、学会誌送付先、メールアドレスが変わられ、変更手続きがお済でない方はご連絡ください。年会費を納入していただいているのに学会誌やメールマガジンが届かないのでは申し訳ありませんので、変更の情報を事務局：office@jsik.jp まで、メールでお知らせください。新・旧の情報を並べて書きいただけると確認できるので助かります。

### 〔3〕 新規入会申込方法

入会ご希望の方は情報知識学会ホームページ <http://www.jsik.jp/> から、「本会について」→「入会案内」→「入会申込フォーム」に必要事項を入力・送信してください。あるいは申込用紙を pdf 形式、doc 形式でダウンロードし、ご記入のうえ下記の事務局へ電子メール・FAX 送信または郵送などをお願いいたします。

なお、退会のご希望の方は、事務局までメールにてお知らせください。

#### 情報知識学会事務局

〒164-0003 〒162-0814 東京都新宿区新小川町5-20

サンライズビルⅡ 3F (株) アドスリー内

FAX:050-3730-8956

E-Mail:office@jsik.jp

URL:<http://www.jsik.jp/>

---

---

## 情報知識学会誌 編集委員会

編集委員長 常川 真央 中央大学  
副編集委員長 江草 由佳 国立教育政策研究所  
佐藤 翔 同志社大学

### 編集委員

|        |              |        |              |
|--------|--------------|--------|--------------|
| 相田 満   | 国文学研究資料館     | 芦野 俊宏  | 東洋大学         |
| 天野 晃   | 国立情報学研究所     | 石塚 英弘  | 筑波大学名誉教授     |
| 宇陀 則彦  | 筑波大学         | 大槻 明   | 日本大学         |
| 岡 伸人   | 近畿大学         | 岡部 晋典  | (株) 図書館総合研究所 |
| 岡本 由起子 | 欧州情報協会       | 小川 恵司  |              |
| 梶川 裕矢  | 東京大学         | 五島 敏芳  | 京都大学         |
| 阪口 哲男  | 筑波大学         | 孫 媛    | 国立情報学研究所     |
| 高久 雅生  | 筑波大学         | 高田 良宏  | 金沢大学         |
| 田良島 哲  | 国立近現代建築資料館   | 時実 象一  | 東京大学         |
| 中川 修   | 大日本印刷 (株)    | 長田 孝治  | ロゴヴィスタ (株)   |
| 長塚 隆   | 鶴見大学名誉教授     | 中山 堯   | 神奈川大学        |
| 西澤 正己  | 国立情報学研究所     | 西脇 二一  | 奈良大学         |
| 根岸 正光  | 国立情報学研究所名誉教授 | 原 正一郎  | 京都大学名誉教授     |
| 原田 隆史  | 同志社大学        | 藤田 桂英  | 東京農工大学       |
| 細野 公男  | 慶應義塾大学名誉教授   | 村井 源   | 公立はこだて未来大学   |
| 村川 猛彦  | 和歌山大学        | 村田 健史  | 情報通信研究機構     |
| 森 純一郎  | 東京大学         | 山下 雄一郎 | 産業技術総合研究所    |
| 山本 昭   | 愛知大学         |        | (五十音順)       |

---

---

### ■複写をされる方に

本誌に掲載された著作物を複写したい方は、(社)日本複写権センターと包括複写許諾契約を締結されている企業の従業員以外は、著作権者から複写権等の行使の委託を受けている次の団体から許諾を受けて下さい。  
著作物の転載、翻訳のような複写以外の許諾は、直接本会へご連絡ください。

〒107-0052 東京都港区赤坂 9-6-41 乃木坂ビル 学術著作権協会

TEL : 03-3475-5618 FAX : 03-3475-5619 E-mail : naka-atsu@muji.biglobe.ne.jp

アメリカ合衆国における複写については、次に連絡してください。

Copyright Clearance Center, Inc. 222 Rosewood Drive, Danvers, MA. 01923, USA

TEL : 978-750-8400 FAX : 978-750-4744 URL : <http://www.copyright.com/>

情報知識学会誌 Vol. 35, No.3 2025 年 9 月 30 日発行 編集・発行 情報知識学会

頒布価格 3000 円

---

---

日本学術会議協力学術研究団体

情報知識学会 (JSIK : Japan Society of Information and Knowledge)

会長 芦野 俊宏

事務局 〒162-0814 東京都新宿区新小川町 5-20 サンライズビル II 3F (株) アドスリー内

FAX : 050-3730-8956

E-mail : [office@jsik.jp](mailto:office@jsik.jp)

URL : <http://www.jsik.jp/>

---

---



# *Journal of Japan Society of Information and Knowledge*

## Contents

### Research Papers

- Evaluation of summaries using a language model of parliamentary proceedings  
..... Kouichirou NACHI, Shoji NISHIMURA ..... 341
- A Theoretical Consideration on Applicability of Floridi's Information Ethics  
Framework Development ..... Tahee ONUMA ..... 350

### Report

- Report on JAPAN OPEN SCIENCE SUMMIT 2025  
..... JSIK, Special Interest Group on Open Science and Open Data, ed. .... 368
- JAPAN OPEN SCIENCE SUMMIT 2025: Individual Presentation Session Proceedings  
Exploring the Use of AI in Metadata Creation for the Social Science Japan Data  
Archive: Metadata Extraction from Social Surveys with an LLM Application  
..... Yukihiro NISHIMURA, Koichi IRIYAMA ..... 389
- Improving the efficiency of cruise data management using Google Forms  
..... Yoshihiro TAKATSU, Chizuru SAITO, Takako SATO, Kazuyo FUKUDA ..... 395

### Errata

- Laying the Theoretical Groundwork about "Digital Archive as a Communication  
Process" : Based on the Concept of "Information" in the Fundamental Informatics  
..... Tasuku MUKUMOTO, Hiroki UEMATSU, Tsukasa YAMANAKA ..... 400

### Information

- Minutes of General Assembly 2025 ..... 401
- 30th Forum on Information and Knowledge ..... 412
- Others ..... 413

情報知識学会誌 第35巻3号 2025年9月30日発行

編集兼発行人 情報知識学会

〒162-0814 東京都新宿区新小川町5-20 サンライズビルⅡ 3F (株)アドスリー内

E-mail : office@jsik.jp

URL : <http://www.jsik.jp/>

(振替 : 00150-8-706543)

学術刊行物 ISSN 0917-1436