

# 情報学委員会国際サイエンスデータ分科会

## CODATA 小委員会(第 25 期・第 5 回)

### 議事要旨

日時: 2023 年 6 月 16 日 15:00-17:00

場所: オンライン開催

議題:

1. 10 月総会に向けて(芦野委員長)
  - 役員候補の選出、総会代表の選出
2. 活動報告
  - EC 報告(芦野-CODATA EC member)
  - 小谷次期 ISC 会長との懇談会について(岩田委員)
  - 他
3. その他

資料: 議事次第

DRAFT CODATA Strategic Plan 2023-2027: Making Data Work to Improve our World

出席者(敬称略): 井上 純哉、中西 友子、宮崎 久美子、村山 泰啓、芦野 俊宏、伊藤 聡、岩田 修一、大武 美保子、五條堀 孝、谷藤 幹子、長島 昭、原田 幸明

出席 12 名(定足数 7 名)

議事:

議題 1 10 月総会に向けて

- 第 25 期は 2023 年 9 月までなので、この構成での CODATA 小委員会は今回が最後ということになる可能性がある
- 2 月の国際サイエンスデータ分科会において、10 月の CODATA 総会に向けて日本学術会議からの代表者及び役員・EC メンバー候補選びを今期の CODATA 小委員会で行うについてご了解を頂いた
- 総会は 10 月であり、日本学術会議 25 期は終了しているが、26 期の分科会・小委員会の開催は 10 月には間に合わないと思われること、候補者のエントリーが今回は 6 月末に締め切られるため、25 期において選んでおく必要がある

- これまで代表・候補については特に幹事会にお諮りするということではなく、分科会・小委員会において決定しており、今回も分科会の了解のもとに小委員会が決定ということで差し支えないであろう

- 前回は EC メンバー候補として芦野委員長、日本学術会議の代表は井上副委員長であった。前回委員会では候補として大武委員というお名前も挙げたが、研究プロジェクトのリーダーなどとして多忙とのことで辞退された

- 日本がこのところ国際的に影が薄いというようなこともあるが、AI など、データの活用が次の段階に入りつつある中で国際的なデータ活用などについて議論する CODATA のような場で発言できる立場を確保することは日本にとって重要であり、役員・EC メンバー候補は出すべきであろう

- 次の選挙では、これまでの活動を見てきたということからも芦野委員長が引き続き EC メンバー候補ではないか

- 国際的なバランスなどを考えると日本からも候補は出るべきで、分野的には COVID-19 のような話題もあるのでライフサイエンス系はどうか。例えば日本学術会議の委員会活動などでも積極的に発言しておられる国際サイエンスデータ分科会の有田委員などは候補となり得るのではないか

- 有田委員はこれまで CODATA の活動に関与してこれなかったので経緯の説明などが必要であろう。また、今後の環境問題、例えば EV へのシフトなどでは世界的な規模でのデータの勝負になるような流れがあり、このようなところに CODATA がどのようにして関わって行くかは分からないがこれまでの経緯などを知り、人脈も作ってきている芦野委員長が継続すべきではないか

- CODATA の EC メンバーには再任制限などはない

- 世代交代が必要であり、次の期の最大の仕事は世代交代になるであろうと思われるが、その際トップは代えずに他は大胆に入れ替えることが重要であり、今回芦野委員長は継続すべきではないか

- 特に異論なく、芦野委員長が次回の EC メンバー候補となることが認められたが、次期に向けて委員の皆様には若手のご紹介をお願いしたい

- 10 月 CODATA 総会に向けて日本学術会議からの代表について、前回は井上委員にお願いしたが、今回大武委員が総会と併設される SciDataCon2023 にセッション提案を出していただいております、もしこちらに対面で参加されるのであれば代表をお願いすることも考えられる

- しかし、今回の総会はハイブリッドで開催されることが予想されるのでオンラインでの出席・投票も可能であろう

- セッション提案については現在 CODATA の EC にてレビュー中であり、採択についてはまだ決定していない

- 総会に代表として出席した場合には、前回総会以降の CODATA 活動報告・財

務状況の報告に対する承認、役員・ECメンバーの選挙、タスクグループの採択などに投票を行う

- 井上委員は学内の業務のためオンラインであれば総会出席が可能であるが対面では出席できない。ザルツブルグ開催なので時差を考えるとオンラインであれば出席が可能

- 総会への旅費については日本学術会議の代表派遣の旅費を申請して承認されているので大武委員が対面で出席可能であればご自身の提案されたセッションへの出席も含めてご検討をお願いしたい。代表派遣については井上委員で申請してあるため必要であれば大武委員への派遣者変更を依頼する

- 大武委員が総会へ代表として出席可能であるか SciDataCon、総会両方への出席についてスケジュール調整を行うが、対面出席が難しければ井上委員がオンライン出席とする

- 代表申請については総会だけではなく SciDataCon2023 についても申請して承認されているので両方の出席には問題ない

-

## 議題2 活動報告

- 3月27、28の2日間開催された対面のECにおける議論について芦野委員長より報告があった

- 1日目は次回総会に向けて DRAFT CODATA Strategic Plan 2023-2027: Making Data Work to Improve our World についての協議が中心であり、この内容、またこれに向けて各国のCODATA委員会がどのような寄与ができるかという点について議論された。

- CODATAの活動領域が広がり、ISCなどからのプロジェクト予算も増加傾向であり、CODATA独自の加盟国からの分担金のみで運営されていた時代と比較して Strategic Plan にもISCなどの動向が強く影響するようになっている

- 我が国については、日本学術会議そのものの再編等が議論されているところから、日本学術会議によるCODATAへの寄与を継続できるよう努力している旨を説明した

- 2日目にはISC、WDS役員も交えての会議があり、データが今後ますます重要になる中でこの3組織の協力についての話し合いが行われた

- 29日にはUNESCO本部において、UNESCO、ISC、RDF、WDS、CODATAの共催によりオープンサイエンスに関するシンポジウムが開催された

- Strategic Plan について前回のものと基本構造は変わっていないが、AIにデータが使われることに対して参照・引用の問題など含めいかにすべきかを記述してゆくべきではないかという議論があった。

- AIの出力は他のデータを学習に使っていてどの部分がどのデータソースといっ

た明示は難しいがデータの作成者が AI の学習に使われることを承知しておらず、勝手に使われているところが問題なので、どのデータを学習したのかを明示することが必要。AI モデルからは分からないが、出力結果からネット上のどのデータを参照しているかを探することは可能ではないか

- 日本では学習することと利用することを分けて考えてきたが、生成 AI の技術によりそこがうまく切り分けられなくなっているという見方もある

- 世の中の的には、トランスペアレントで説明可能な AI でなくてはならないという方向であり、Strategic Plan では Data Policy Committee で議論する項目として挙げられている。また、ISC にも AI 問題を議論するグループが作られており、CODATA としてはそことの連携を中心に考える

- ISC 次期 President 小谷先生と国際対応委員会関係者との懇談会について(岩田委員)。

- 国際関連の活動をするメンバー40名程度の参加があった

- ICSU と ISSC が合流して ISC になったが、その中で Cross-disciplinary な議論が本当に出来るのかという質問をしたが、今後の課題であり、なかなかコミュニケーションが難しいといった回答であった

- 既にグローバルに活動を進めているグループもあれば国内的な活動のために ISC を使いたいというグループもあった

- Strategic Plan において、これまでの CODATA の重点領域を見ると項目が増えているが、それに伴ってリソースは増えているのか、Cross-Domain というところで具体例はあるのか

- Cross-Domain については従来から重要であると言われ続けてはいるが実際にはなかなか進んでいない。情報の時代において AI なども含め新しい動きがあるのではないか。様々なトレードオフがあるため、何らかの決定的なデータを学界から示す必要があるのではないか

- CODATA のリソースについては、メンバーの分担金に加え、ISC などからプロジェクト予算を取ってくるようになって期限付きのスタッフなどが増えている。また、Cross-Domain については EU の予算による WorldFAIR プロジェクトで災害データ、農業データなど CODATA のタスクグループにあるようなトピックについてケーススタディを行っている

- 以前の CODATA は各国に確固とした実施主体があってこれらの調整を行うといった組織であったが、時限のプロジェクト予算で活動してゆくという方向性はそれで良いのか

- 現在の EC メンバーはデータアーカイブ方面の専門家が多く、かつてのようにデータを実際に作ってゆくという立場の人があまりいなくなった

- Strategic Plan の重点領域の変遷を見ると、70-80年代の対象はかなり具体的な

分野であるが、2000年あたりからはデータに関わる一般化した問題が対象となっており、Data Science 分野における他の活動と差別化することができるのか。

CODATA の役割が失われているのではないか

- Strategic Plan で AI での利用なども含めた Data の流通、交換についての問題があまり述べられておらず「理想論的な」問題が中心となっているのは如何なものか

- オープンアクセス、データポリシーなど新しい世界を作ろうというときにあって CODATA のアドバンテージ、昔の CODATA のあり方、COVID-19 の後にあって対面、オンラインの違いなどを見ると各国の文化や感覚などデータの捉え方が国によって異なると考えられ、単に効率的なデータ収集・管理だけではなくそういったところを取り込んだ議論をする場はあるのか

- データやテキストの行間にある意味を捉えきれていないところがあるのではないかと思われ、若い人に新しいタスクグループ提案などに繋げることはできないか

- 我が国ではデータベース化するにあたってデータ構造をどのように作るかを長く議論しているが、海外ではデータマイニングのようにデータを使う議論が中心でデータ構造、それでは表し切れない部分をいかに表現する、知識の体系としてデータがどのようにまとめられているのかという議論が失われているのではないか

- しかしながら研究のトレンドがデータを使う方向なので予算獲得等を考えると知識体系の表現といった研究は難しい

- CODATA の性格として、1960 年の 7 人の創設者の議論を聞いたことがあるが、分野間・国際間でのデータの活用、そのような活動によって先行的に社会の課題を見つけ出していくという二つのことを考えていたように思われる。このために、データをサイエンスとしてだけではなく、テクノロジカルなデータに重点を扱うということであった

- 課題を設定してプロジェクト予算などをつけるという考え方は既に課題が分かっているということで、当然ながら後追いになるので、この 20 年ほどの日本の科学技術政策の失敗でもあるが、CODATA についてもそのようなプロジェクト予算の獲得に向かうことで後追いになってはいけない

- データ倫理が重要と考えられ、CODATA でもしばらく以前からトピックに挙がっているが、データの使い方が中心でデータをどう生み出すかというところでのデータ倫理について述べられることが少なく、CODATA において先行して議論して頂きたい。ここでデータの生成というところにはデータを作る人間だけではなく、AI や、作られたデータによる社会的な影響も含む

- ヒトゲノムに関しては特にデータ倫理について深い議論があったが、CODATA

としては CODATA の使命と全体としてのデータの産生・利用について分けて考えるべき。各国または複数の国が協力して何らかの活動を行おうとするときに CODATA のステートメントがあることがサポートになる、というのが実務的なところで、ステートメントによって実際に世界を変えていこうといったことには力不足か

- 政策的な部分については、政策の専門家が参画するのが良いとも思われるが、個別の研究分野の専門家が集まって議論するのが CODATA の強みであるという矛盾する部分がある。プロジェクト予算などがメンバーにいかにか還元されるか。デジタル庁のような動きもあるがそのような動きを今後のプロジェクトに活かしていくことが重要
- 海洋のデータなども多くの民間の船舶から google などが大量に収集できる状態であり、経済発展・環境保護・安全保障のキーワードのもとに新しいことをやるような試みを後押しするという役割もある
- CODATA の活動や役割について周知する上では、代表やメンバーとして同じ人が継続的に顔を出しているというのも重要
- 9月に中国にて International Symposium Open Science Cloud が開催される
- やはり9月にはインドにおいて International Conference on Open and FAIR Data Ecosystems, Principles, Policies, and Platforms が開催される。このトピックではインド主催の初の国際会議か
- インドについてはグローバル・サウスのリーダーとして自国の立場を強化するためにこういう国際的なトレンドを活用するので見習ってほしい
- TG 提案の締め切りは EC メンバー候補のノミネーションと同じく6月27日となっている
- アフリカについては南アフリカが突出しているが他はどうか
- ロシアについては副会長が出ているが動きが見られない。材料データについては CODATA とも関連のある方がワークショップを企画している
- エジプトについては CODATA のメンバーではないので動きが分からない
- 大武先生の Task Group 提案は今作成中であり、Working Group での方向性への議論を活かしたものにしたい。認知症予防に向けての対応などについて国際的なところから日本国内での予防活動に活かすようなものにしたい
- 認知症基本法が成立したが、COVID-19 の期間に骨抜き・改悪が成されてしまったため、国際的なフレームも活用して改正のタイミングでなんとか出来るようにしたい。予防よりも発症した人への手厚い対応をするような法律になっている