

データサイエンス共同利用基盤施設
令和２年度 第５回運営会議 議事概要

日 時：令和２年１１月２０日（金）１５：００～１７：００

場 所：ウェブ会議（zoom を利用）

出席者：

【外部委員】近藤委員、諏訪委員、高橋委員、照井委員、中井委員、堀委員、三好委員、
矢嶋委員、横山詔一委員、横山雅之委員

【内部委員】藤山委員、小原委員、門倉委員、前田委員、北本委員、野口委員、上野委員、
野水委員

【陪席】 荒木特任教授、平野企画連携課長、橋本企画連携課副課長、昨間 DS 推進係長、
遠藤 DS 推進係員

議 事：

議事に先立ち、昨間 DS 推進係長から、定足数を満たしており本会議が成立することが報告された。

審議事項

１．前回議事要旨確認

修正意見等がある場合は会議終了までに申し出ることとし、特段の意見はなかったため、案のとおり確定した。

２．令和２年度人事について

前田委員、藤山委員から、資料２－１・２－２（回収資料）に基づき、社会データ構造化センター特任教員の採用及び、データサイエンス推進室の兼務（２件）について説明があり、審議の結果、承認された。

３．第４期中期目標期間に向けての検討状況について

藤山委員から、資料３－１に基づき、第４期に向けての検討会やWGの開催予定及び概算要求に関するスケジュールについて説明があった。次いで３－２・３－３に基づき、外部評価委員からの評価結果及び対応案について説明があり、現在、外部評価委員からの意見を元に、各研究所、DS 施設内では各センターにヒアリングを行っている旨報告があった。

上記のヒアリング内容について、各センターより活動内容、今後の構想等の説明があり、以下の意見交換を行った。

○ライフサイエンス統合データベースセンター

・開発したソフトウェアは民間にライセンス等しているのか。それによる収入はあるのか。

→基本はオープンになっている。ただし、現在進めている SIP では産業化に繋げるプロジェクトを行っているため、今後クローズドのデータを扱い課金制度のようなものを採る可能性はある。

- ・データのオープン&クローズド制について、将来的な考えはあるか。

→オープンの情報をクローズドの情報と組み合わせて使用したいという要望がある。オープンからクローズドへの一方向のアクセスを可能にする取組が SIP で実装段階にあり、これによって企業は自社のデータを外部へ開かずオープンなデータを組み合わせて使用することができる。

- ・資料中に「生物資源」とあるが、その観点では自然科学研究機構の基礎生物学研究所のバイオリソース等との連携は考えているのか

→バイオリソースプロジェクト（NBRP）のデータをほかのデータと組み合わせて使う形で、基生研だけでなく理研、遺伝研とも協力・連携している。

○極域環境データセンター

- ・データを統合するにあたり、日本だけでなく海外のデータも含めて行う必要があると思う。そのために、海外とデータとの標準化の方法論等はシンポジウムの中で検討しているのか。

→例えば学術データベースについては、南極データ管理委員会の議論に沿ってシステムが構築されている。登録されたメタデータは NASA の南極マスタディレクトリに自動登録される仕組みとなっており、メタデータの総合運用ができるようなシステムになっている。また、北極域のデータベース（ADS）は、DIAS や IASC 等の国際的なデータ活動と連携したシステムとなっている。

- ・専門家ではないユーザーが増えることも想定されるが、使いやすくするために、例えばバイオハッカソンのように互いのデータの設計の仕方を話し合う機会はあるのか。

→ライフサイエンス分野とは性質は異なるが、極域分野では研究者や大学院生を対象に、研究コミュニティの中でいかにデータを扱いやすくできるかという観点で統合化を進めている。

○社会データ構造化センター

- ・ソーシャルビッグデータ事業の中に「地方のソーシャルビッグデータの専門家養成」とあるが、具体的な活動内容、今後の計画等があれば教えてほしい。国のオープンデータの活用面で、地方自治体ではデータの活用苦労しているため、成功事例等を公開していただけると、他地方への展開も望める。共同研究成果の還元という点でもお願いしたい。

→DS 施設全体で取り組むデータサイエンティスト人材を育成する教員の養成プロジェクトの機能として何らかの貢献をして活動を進めたいと考えている。

○人文学オープンデータ共同利用センター診断

- ・将来構想として、生命科学分野の「PubAnnotation」を歴史的ビッグデータに活用する試み、とあるが具体的にはどのような活動か。

→PubAnnotation は、文献アノテーションのレポジトリであり、論文等の中である単語をア

ノテーションするツールである。これは歴史文書の中で地名や人物名をアノテーションするのと同じ仕組みであり、活用できる可能性がある。現在試験的に使用を始めており、講演も予定している。

○ゲノムデータ解析支援センター

- ・支援数について、5人で約10件/年となっているが、1人当たり2件/年の支援を行っているということか。

→ほぼ全ての課題はセンター長が担当している。比較的簡単な課題（1,2件程度）をポストドクが担当し、センター長が指導している状況である。多くの課題を受け入れるために自動化可能なところは自動化を心掛けている。

○データ同化支援センター

- ・逆引きデータ解析支援システムの「データを持っているが〇〇をしたい」、という視点は、様々な分野にいる研究者の視点であり、他分野からデータサイエンス分野へアクセスするきっかけになり、とても良いと思う。

○全体

- ・素晴らしい取り組みや成果が出ている一方、研究者の任期の問題含め人材面で危機感がある。第4期の計画が重要となるが、DS施設として大きくなれるような予算、体制を実現して欲しいと思う。
- ・データサイエンスは継続して行うべき事業だと感じた。人員の問題が大きいですが、任期なしのポストを増やすことを含め検討が必要である。
- ・現在のセンター名には分野名が含まれているものが多いが、センター名に記載された以外の分野からアクセスしづらくなるのではないか。分野名ではなく手法等をキーワードとして入れることも検討してはどうか。手法等は分野が異なっても活用できるため、他分野への接点が作れたらよいと思う。
- ・DS施設は手法を他分野へ横展開することが想定されて作られた組織である。ただし、各分野の実問題に当たるにはドメインも必要であり、手法とドメインのバランスが重要である。

以上

資 料：

0-1	データサイエンス共同利用基盤施設運営会議委員名簿	P.2
0-2	情報・システム研究機構データサイエンス共同利用基盤施設運営会議規程	P.3
1-1	令和2年度第4回運営会議議事概要（案）	P.5
2-1	令和2年度人事案件 審議対象者リスト	P.8
2-2	令和2年度人事案件 審議資料（回収）	画面共有
3-1	第4期中期目標期間に向けたスケジュール	P.9
3-2	外部評価委員会による評価報告／外部評価結果に対する取組	P.10

