

総合科学技術・イノベーション会議

有識者議員及び健康・医療戦略推進に係る有識者による合同意見交換会 [公開議題]

議事概要

- 日 時 令和5年3月9日(木) 9:00~9:53
- 場 所 中央合同庁舎第8号館8階特別中会議室
- 出席者 上山議員、梶田議員(W e b)、梶原議員、佐藤議員(W e b)、菅議員、篠原議員、波多野議員、藤井議員(W e b)
(事務局)
森総理補佐官、松尾事務局長、奈須野統括官、坂本事務局長補、井上審議官、
覺道審議官(W e b)、高原審議官、次田参事官、赤池参事官、
(健康・医療戦略推進事務局)
西辻局長、長野次長、神田参事官
(健康・医療戦略推進に係る有識者)
永井自治医科大学長、宮園理化学研究所理事・東京大学卓越教授
(文部科学省研究振興局)
工藤参事官(情報担当)
(国立情報学研究所)
喜連川所長
(科学技術・学術政策研究所 データ解析政策研究室)
林室長
(大学改革支援・学位授与機構)
山口特任教授
(東北大学)
大隅副学長・附属図書館長
(京都大学)
引原理事・副学長
(情報通信研究機構)
村山研究統括・ナレッジハブ長
(オブザーバ)

橋本内閣官房科学技術顧問、（文部科学省）井上諭一総括審議官

- 議題
 - 1. 総合科学技術・イノベーション会議有識者議員及び健康・医療戦略推進に係る有識者による合同意見交換会の開催について
 - 2. オープンサイエンスについて（日本学術会議からのヒアリング）
 - 3. その他

○ 議事概要

午前9時00分 開会

○上山議員 皆様、おはようございます。

定刻になりましたので、ただいまより第1回の総合科学技術・イノベーション会議有識者議員及び健康・医療戦略推進に係る有識者による合同意見交換会を開催いたします。

本意見交換会は、科学技術・イノベーション推進事務局と健康・医療分野の司令塔組織との連携を深めるために開催するものです。

健康・医療分野は我が国のアカデミアで行われている研究開発において重要な位置を占めており、この二つの連携は重要なものだと考えております。

本日は、健康・医療戦略推進に知見を有する先生方として、永井良三先生と宮園浩平先生にご参加いただいています。

それでは永井先生と宮園先生から一言ずつ最初にお言葉を頂きたいと思います。

まずは、永井先生の方からお願いできますか。

○永井有識者 健康・医療戦略推進専門調査会の座長もしております自治医科大学の永井です。

今までCSTIと健康・医療戦略室のプロジェクトを切り分けてきているようですが、これからは一体化して進めないといけないと思います。こうした機会に情報交換やこれからの方針について御意見を頂ければと思います。

どうぞよろしく願いいたします。

○上山議員 ありがとうございます。

続きまして、宮園先生から一言お願いいたします。

○宮園有識者 私は健康・医療戦略推進専門調査会の副座長も務めております、宮園と申します。理化学研究所と東京大学をクロスアポイントメントで両方に所属させていただいております。どうぞよろしく願いします。

AMEDの方でも疾患基礎プロジェクト等を担当しております、それから理化学研究所の方では生命科学系の各センターと一緒に推進しているということですので、C S T Iの皆様と生命科学のいろいろな事柄について議論させていただければと思います。

どうぞよろしくお願いいたします。

○上山議員 両先生、どうもありがとうございました。

今後、少しずつではありますが、C S T Iと健康・医療戦略室との間でコミュニケーションができれば大変有り難いと思っております。

なお、本意見交換会を公開で行う部分については、資料及び議事概要を公表予定となっております。

続きまして、オープンサイエンスについての議論に入りたいと思います。

これは永井先生、宮園先生、両方の方々も何らかの形で御発言していただければ有り難いと思っております。

これは、質疑応答の部分については非公開という形にさせていただきます。

本日は、日本学術会議から国立情報学研究所の喜連川優所長にご参加いただきまして、内閣府からの議事依頼、研究DXの推進、特にオープンサイエンス、データ利用推進の視点に関する審議について、御回答を頂きたいと考えております。

また、同じく日本学術会議から大学改革学位授与機構の山口周特任教授にもご参加、内閣府からの審議依頼、同じ問題についての回答の後、オープンアクセスに関連する部分について質疑の御対応を頂きます。

さらに、本日は外部有識者として東北大学の大隅典子副学長・附属図書館長、京都大学の引原隆士理事・副学長、そしてG7オープンサイエンスワーキンググループの共同議長を務めている情報通信研究機構の村山泰啓研究総括・ナレッジハブ長にもオンラインでご参加いただいております。

さらに、文部科学省より工藤参事官、科学技術・学術政策研究所より林データ解析政策研究室長にも御同席していただいております。

それではまず内閣府から説明をお願いします。

○赤池参事官 私から、本日の趣旨の説明をさせていただきます。

資料1を御覧いただければと思います。

2ページです。

以前から申し上げているとおり、オープンサイエンスはオープンアクセス、それからオーブ

ン研究データの両面がございます。今日はこの両方から学会議との議論をさせていただきたいと考えております。

ページ、めくっていただきまして、3ページです。

全体の俯瞰ということで、時間軸で少し整理しましたが、将来のデータ駆動型研究からのバックキャスト、それからこれまでオープンサイエンスの流れの中で、成果のオープン化の潮流、そうした中でいわゆるオープンアクセス問題とか、ジャーナル問題が残っているところですが、こうしたものを一体的に俯瞰的にとらえるべきだということで考えております。

そして、4ページ、5ページ、6ページ、この辺りは前回は御説明していますので、少し省略をさせていただきます。

7ページです。

本日の議論、関係機関からのヒアリングの一環としてやっておりまして、これをG7科学技術大臣会合や統合イノベーション戦略に活かしていくと考えております。

8ページです。

これは評価専門調査会における見解で3月1日の委員会で出たものですが、ここにも書かれているとおり、3の下の方でございますが、アカデミアも研究の自律性の観点から主体的に検討する必要がある、としておりまして、CSTI、それから学会議と適切な役割分担をしながら取り組んでいくということです。

参考資料ですが、10ページ、喜連川先生の方からお話をされます。

11ページ、こちらは学会議の研究力強化の方の御回答なのですが、この中にオープンアクセスに関する御提言が書いてあります。

少しビジーな文章になっておりますが、真ん中以下には提言が書かれておりまして、例えばナショナルライセンスへ移行すべきであるとか、あとは交渉力をしっかり強化していくこと。大学法人がまとまって契約するなどの中間的段階の姿も模索すべきだとか、それから、学協会の発信機能の強化、そういったことについても御提言をされております。かなりの部分は、私どもの議論の方向性と合いたしているものと認識しております。この点については、質疑対応ということで山口先生にも御同席を頂いております。

私からは以上です。

○上山議員 ありがとうございます。

それでは、喜連川先生、よろしく申し上げます。

○喜連川所長 おはようございます。喜連川と申します。

あと3週間ほど情報研の所長をして、最後までぎりぎりでごにお伺いさせていただいて幸いです。

冒頭、上山議員から御紹介いただきましたように、これは内閣府さんから学会に対して審議依頼がまいりまして、それに対して、議論を重ねて回答を提出したということです。

こうした書類が届きまして、お時間がないのでぱっぱと見ますが、審議依頼という長い文章があるのですが、1、2、3と三つ質問をしますというのが来ております。

見ていただきますと、原則、データ、データ、データ駆動ということで、先ほど赤池参事官からのお話のオープンアクセスではなくて、むしろデータの方に注目して、答えてくださいということです。

当然のことながらもう論文ではなくて、論文の基になるデータが主役になるという時代ですので、こうしたことについて議論を重ねまして、とてもたくさんの先生方、これはもう少し御紹介できませんが、いろいろな先生方に御支援していただきながら答申をまとめていったというふうにご理解いただきたいと思います。

最終的には、これは赤池参事官からのまとめですが、こうしたふうに提案を1から6までまとめているということです。それだけだと無味乾燥としていますので、少し私の方で論点を申し上げたいと思いますが、もはや財はデータだけではないということで、これから考えるときはもっともっと広がりを持つべきだろうと。

例えば、ロボットのプログラム、ロボットがどんどん入りますので、そのときのプログラム構造というのは著作権法で守られるとか、あるいは学習済みモデルをどうするか、あるいは最近ではいわゆるAPI、ここが非常に活性化してきました。こうしたものを丸ごと考えましようということと、それからデータ基盤を構築するところに対しては、やはりアプリと基盤、セパレートしたデカップリングがあるだろうということ。生態系がどういうふうに動いていっているのか。生態系というのは生きている生態系ではなくてそのデータの生態系ということです。最終的にはプラットフォームを育成することが重要だろうというふうに思っております。

それから、オープンと知財、このバランスをどう取るのかというのは誰もまだ分かっていない。いろいろな難しい論点があるというところですが、最初にまずお聞きいただいたのは、オープンサイエンスは日本学会はどう思うのだという話なのですが、こんなものはもう答えようもなく、推進すべきだということで、御案内のように、材料の科学ということになっている訳ですが、少し面白い例といたしまして、スーパーコンピュータとデータ駆動科学というのがコンピートしているようにお考えになるかもしれませんが、そんなことは全然なくて、富岳

で作ったものをどうやって料理するかというところがデータ駆動科学であるというふうに御理解いただければと思います。

これはDS2022という、エネルギー基本計画等でもある種温暖化がどうなっていくかという基礎データというのは文部科学省環境エネルギー課が作っている訳ですが、一番左、文部科学省で、それが環境省にあって影響評価を行い、それが更に金融庁に、こうしたふうにデータをいじる場が必要になってくる訳です。

ところがこのデータ、大体20ペタバイト弱ございますので、20ペタバイトというのははや研究者が自分のところでダウンロードできないです。したがって、いじる場そのものを作ってあげることが非常に重要になってきて、今日、坂本審議官もおいでになりましたし、長野さんもおいでになりまして、いろいろな局面でDAISを御支援していただいた訳ですが、こうしたシステムはもうなくてはならないというマストな方向になっているということで御理解していただければと思います。イベントアトリビューションというのがキーワードになっているということです。

これはほんのまだひと月もたっていないのですが、こんなもんができましたよと言ったら、3,000人のユーザーがもう登録されておられまして、結構大きなインパクトが出てきているということです。もう少しビジネス寄りですと、これは知財の審議会の中で出てきて面白いなと思ったのですが、日本からの特許は総量は変わらないです。左側のブルーのライン、この辺がそれに相当します。

右側はAIに関する特許なのですが、やはり中国とアメリカが大きいということで、なんでやねんということなのです。なんで日本のAIは特許が出てこないのという話になる訳ですが、これはData Fuels AIということで、データがないとAIが動かないということから、日本はやはりデータをもっと整備していくところの領域にどんどん出てくるのではないかというのが喜連川の勝手な意見で、これはこの間、波多野議員と一緒に審議会に出させていただいた次第です。

2番目に、この文章は中々読解が難しいですが、現世、何をすればいいかと、データ管理、データ利活用のために具体的に何をやるのかということ。それから、もう1個はこれからどないするのかということ、現世と未来とこの二つだと思っています。

最初に申し上げますのは、まずきっちりデータプラットフォームを作りましょうということで、これは赤池さんがとても真剣に考えられて、素晴らしいのですが、基本的な考え方というのをR3の4月にお出しいただいています、これは非常に精緻にまとめられています。

しかしながらこの研究者の責務というところに行きますと、1、2、3、4、5、6と、これもしろ、あれもしろとずっと書いてありますと、賢い研究者だったらどう考えるかと言うと、例えば僕はあほなので、データの研究をするかもしれませんが、こんなたくさんやるのだったらペイしないなと思って別の研究をする可能性があるということになります。要するに、守らなければいけないルールがあまりにも多過ぎると。

したがって、ポイントはこれをどうやってデジタルで全部ミニマイズするかという、これは全ての勝負になるというのが我々の見方ということで、デジタルが研究者に寄り添うということとです。

N I Iはこのプラットフォームというのを実は5年前から作ってまいりました。2016年にSINET5というものをオープンにしました。その前は4だったのですが、これはぼろぼろでして、どこを一体どうするのかということで、完全にアーキテクチャをダークファイバーに変えまして世界一になっています。

ようやくデータプラットフォームを作らせてくださいということを文部科学省に申し上げまして、5年間やってきましたので、世界中に対して全然ビハインドしていないということです。

これが現在、加入機関といいますか、データを利用されているところですが、とても皆さん御興味を持たれてどんどん入ってこられて、今テスト利用されているということです。

何でこんなことをするのだと言いますと、これは機関デポジトリの歴史なのですが、見ていただきますと、当初文部科学省はこのブルーのやつで、機関デポジトリというのは機関が作りなさいということで一生懸命作っていた訳です。そうしますと体力のある大学は作れるのです。しかしながらだんだん体力のないところに行きますと、こんなもん僕らはやれませんか作ってくださいということになりまして、ピンク色のものを作るようになりました。そうすると、そんな便利なものあるのだったらもう我々自分でやるのはやめますというので、どんどん減ってきてまして、今はもう5分の1ぐらいです、自分のところでやっているのは。

だったら、丸ごとN I Iが最初からデータ基盤を作った方がいいだろうと。こんなものは競争基盤ではなくて協調基盤でなくて、共有基盤なのですね。こうしたセンスをお持ちいただけただことは非常に有り難いかなと思っています。

最近、我々のところでどういうことが起こっているのかというのを御紹介すると面白いかと思いますが、ある機関が我々のところに来られまして、10ペタをお持ちなのですが、10ペタ丸ごと、もう我々マネージできないので、全部N I Iのデータ基盤に入れさせくれないかということ、研究所の所長さんが我々のところに、10ペタというのは結構大きいです。こ

れは随分御苦労されているなと思いますが、何でそうしているかという、もう調達などできないと言うのです。

これは病院の情報システムと同じで、旧7帝大がばらばらにやるということは病院がやる負荷というのはとんでもなく大きいです。よっぽどプロフェッショナルでないとうことはできない。しかも、これは少しオープンなので、この後余り申し上げにくいのですが、今日は藤井議員もお入りになっていますが、我々が調達するときも東大の本部の調達係と一緒にやってやります。これぐらいコンプレキシティが高いということが一つです。

それから、運用がそもそもできない。そんなことやるようなIT人材をそれぞれの研究分野にみんなばらまくほどの余裕は我が国家には全くありません。そうした意味でデカップリングをしましょうということが我々は重要だと思っていて、応用をやることはIT基盤はもう別個にしておきまして、どうしてこうしたことを言うかと言うと、GoogleもAmazonも何をやっているかという、いわゆるODMをやっています。

これは調達パワーの最大化でして、あらゆる意味でジャーナル問題とほぼ同じことがここに出ているということです。SINETはどうかという、正にこれをやった訳ですね。

SINETはNTTさんのfiber to the homeという昔のアイデアがものとても価値が高いもので、それを御利用させていただいている訳ですが、SINETは丸々手作りです。ですから、NIIとNTTというのは連携ができて、NTTにとってみると、一番でかいネットワークのプレイヤーというのがNIIになっているというふうに見ることができます。

これによって、生態系が全部観測できますので、次の策をどんどん打ち込むことができるということになります。このデータ基盤に対しましては、NIIもやはりソブリンを作っていくべきだと思っております。

NTTさんがよく出てきて恐縮なのですが、私のドクターをあげた方がNTTにおられました、秘密計算というサービスメニューを入れたのですが、上から四つ目、真ん中のところ、暗号計算というのがあります。これは難しくてよく分からないだろうなと思ったのですが、この間プレスを打ちましたら、もう早速希望者が出てきて、産学連携にとっては、これは非常に重要で、NTTさんも非常に喜んでおられるということです。

問題は、どれだけお金が掛かるのかという話になる訳ですが、これは見ていただくと面白いのですが、左隅、2022年の11月に、EOSCのヨーロッパ・オープンサイエンスクラウドのシンポジウムがあって、コロナでうちのスタッフが行きたくないと言うのですが、

行ってくださいと行って、頑張るって行って、大体の現状はもうつかんできました。

原則、1ビリオンということになっております。1ビリオンというのは結構大きなお金で、向うのアーキテクチャはこんな数字になっておりますが、我々と少し違うベクトルで進んでいますが、根源的に何一つ遅れておりませんで、予算額としては一言でいうと富岳一つ分みたいなことになります。

これに対して、高いと見るか安いと見るかということなのですが、研究者にお聞きいただくと分かるのですが、あなたはお金が欲しいですか、データが欲しいですかという質問をされると分かると思いますが、原則データが欲しいという研究者が結構いると思います。これがデータ駆動時代です。したがって、データ基盤にもう少しお金を投資することは全然悪くないと思います。

EUと日本があったときに、EUに対して日本が勝ち代があるかという話になります。これはEUがどういうふうなデータ基盤の作り方をしているかというのをずっと見て行って、我々も調査を重ねた訳ですが、向うに行って喜ばれることは、日本はやはり一枚岩ですねという言い方で、とてもうらやましがられるとうちのスタッフは言っています。

各国はもうばらばらで、こうしたことをコーディネートすることというのは技術屋、つまりIT屋から見ると著しくしんどいだろうなと思っておりますので、我が国は十分なアドバンテージがあるというのが我々の見方です。

それから、2を少しとおり越して、先に3を御紹介しますと、モニタリングというのがありまして、これはデータがどういうふうに進んでいて成長しているかとオープンサイエンスのオブザーバトリミみたいなものを持っておりまして、普通彼らがやるのはどれぐらいがオープンアクセスなのか、どれぐらいがグリーンなのかゴールドなのか、こんなもんでもいいというところで、我が方は少し面白いことをやりまして、オープンエアは本当に面白いのです。

向うはいろいろなロゴを全部オープンにしてくれています。つまりデータセットがあつて、パブリケーション、論文、ソフトウェアという、こうしたものがどういう参照関係かというデータを全部出してくれていまして、右下ですが、テラバイトオーダーのデータ。これをこっち側に持ってきて可視化しますと、見ていただくと分かるのですが、赤がデータセットで緑が論文、それがどう関連しているかというのをこんなことは向うの人はやってないと思うのですが、我々のところでいろいろ解析しますと、分野、分野ごとによって、この生態系が大分違ってきます。

これはフルディスククロージャーでやったのですが、非常に面白い図柄がたくさんできまし

て、これはまだやっている途中なのですが、こうしたものを見ながら、次の策を打っていくということがものとても重要だということを申し上げたいと思います。

Googleが持っている最大の価値はコンテンツではないのです。コンテンツはユーザーが持っております。Googleの最大のアセットはどうみんなが使っているかという、そっち側です。その情報を、理由は何かしら分からないですが、丸ごとオープンにしてもらえるとというのは大したものだなと思っております。

最後に、人材のことですが、いわゆるデータプロフェッショナルがいるという言葉がどんどん出てきている訳ですが、アメリカのUSアカデミーが何と言っているかと言うと、Team Scienceという言い方をされていて、この言葉は非常に心を打ちます。つまり問題はどんどん複雑になってきますので、パーティシパントがものとてもたくさんになる。

喜連川の言いますと、ノーベル賞はつぶれる。たった3人ではないと。100人くらいの人が一緒にならないと解決できないような問題が山のように出てくるというところです。

これは少し面白い論文なのですが、これは2022年にMIT Technology Reviewで出たのですが、コロナ肺炎のAIツールを山のように作ったが、どいつも役に立たなかったということをレポートしています。

これは実はnature machine intelligenceの論文を引っ張っているのですが、この忙しいときに2,000本の論文がグローバルに出て、全部調べた、でも全部あかんかったという言い方をされています。

日本はそう悪くはないと個人的には思っています、これは順天とNIIと名古屋大学、もちろんAMEDさんにお力添えを頂いた訳ですが、我々4億枚も放射線画像を持っております、医療界の中でこんなクラウド基盤を作ったのはAMEDの初代の理事長の末松先生と御一緒させていただいたからですが、クルーズ船の自衛隊病院が持ってきたデータも全部入れますと、こうしたセグメンテーションで、これは10月の時点でこれができたというので圧倒的に世界で一番ということになります。

ところが、蓋を開けてみるとどうということだったかと言いますと、PMDAがやりましたのはアリババのソリューション1号に入れて、何でなのですか、放射線学会はこれが出せないのです。オプトインができないから、やってないからという言い方で、放射線学会も確実に使えますというぐらいの画像処理を作った訳ですが、これはできなかったということで、少し目詰まりをもう少し広く御覧いただくことが必要になるのではないかと思います。

それから、これはまた言うと、いろいろな御意見があろうかと思いますが、データプロフェ

ッショナルの人材育成はどんなものなのか少しよく分かりません。

喜連川的に見ますと、データを知っている人というのはデータを作った人なのです。その作った人をアシストするというのは関節技になります。我々がやるのは作った人そのものを持ってくるということをやっています。したがって、最近の朝日のニュースなのですが、そうした人のパワーをオンデマンドのクラウドソースで持ってきた方が圧倒的に早いのではないかとというのが個人的な 이슈で、我が方は東京大学では経済安保の問題をやろうとしています。こんな問題を解こうと思うと、東大の中でも全然人が足りなくて、日本国中全部集めてようやく問題が解けるかどうかという話です。

これは実はコロナが始まったときもそうなのです。感染系に対してはもう終わっているのではないということ、中々人が育たなかった。経済安保をやろうと思えば、グローバルな産業連関表が必要になりますが、こんなことを研究している人はもはやいないのです。そうしたことで常に分かっている人をどう集めるかということが非常に重要だと思っております。

それから、未来という意味で申し上げますと、ARWというのが出てきておまして、大量にデータを生み出すものを連携しながらつなげるというのが非常に重要です。

日本では今、コアファシリティといいまして、電子顕微鏡とかあるいはSPRING-8とか装置がございます。こうしたところから膨大なデータが出てくる訳ですが、今までだと無理だった訳ですけど、現在SINET6というのはもう400ギガですから、世界でもすごい性能を持っておりますので、しかも日本海側も含めてありとあらゆるところを結んでいます。

これがSINET6、このアーキテクチャそのものはSINET5とほとんど変わらなくて、これが大変大きな世界のパワーになっていると思います。

先ほど言いましたARWですが、この審議依頼が来て、5、6月にこの話をしていたら、アメリカのアカデミーオブサイエンスがARWという名前を付けてまして、ロボットをうまく作ろうと。

ここはたしか一杉先生が来られてヒアリングをなされたとお伺いしておりますので、これは理研の高橋さんのバイオ系の話ですが、これは双腕のロボットで作らして、iPS細胞に随分役に立ったというふうにおっしゃっていただいております。

つい最近の日経では、非常に大きなマーケットになるということになっておりますので、これは大学共同事業から見ますと、ユーザーがロボットのプログラムをアップロードする。アップロードしたのからデータを生成する、そうしたまた全然違うエコサイクルが出てくるのではないかとこのように思っております。

データという意味ではスパコンが大きなソースになります。ただ、富岳の横に付いているストレージというのはテンポラリーストレージでありまして、パーシステントなストレージ空間ではないので、そうした意味では全然価値観が違うというのは御記憶いただければと思います。

5番目は、これは大した話ではないですが、データの中で皆さんが覚えておいていただきたいのは、FAIRという、Findability、Accessibility、Interoperability、Reusabilityという言葉で、この中で一番重要なのがFです。

つまり検索エンジンというのは言葉で書いてあるので、だから検索できるのです。データはどうやって検索するのだと、データは腐るほど出てくる訳です、これから。そのときのファインダブルというのが分からなくて、みんな世界中で言っていたのですが、最近FIPという言い方をしまして、これはFair Implementation Profileということで、彼らもこんなもの一般則は無理だということで、ドメインディペンデントにあるという、ドメインのオントロジーを作っていくということになってきていますので、これは地味ですが一歩一歩やるしかないなと思っています。

それから、法制度が一番ややこしくて、最近これは非常にいいペーパーが出ました。

Open Science and Intellectual Property Rightsということで、研究を進めようと思いますとやはりインセンティブをリサーチャーにあげなきゃいけないということで、コロナのワクチンのときにWIPOが、WHOはこんなもんだにせんかい、という話でしたけど、WIPOはそんなことはやっちゃいけないと、どこまでオープンにして、どこまでをインセンティブに回すのかというところをこのレポートはしています、こんなことを研究しているレポートはほとんど世の中にないないという言い方をしています。

したがいまして、非常に新しい研究領域でここはしっかり押さえていくべきだと思っております、2022年にData Actが出まして、これは我々も随分丁寧にフォローしてきた訳ですが、これは促進法です。

GDPRが抑制法であるのに対して、Data Actというのはどンドンみんなでシェアしなさいという言い方になっているのですが、これはあまりにもやり過ぎで、これは相当物議をかもししているところです。ということで、この領域は全くもってレギュレーションがどっちに行くか分からない、そうした新しい領域です。

そうした意味で、我々としては非常にエクサイトしております、5年前から十分準備を続

けてきたということはよかったのではないかと思っている次第です。

N I I はまずオープンサイエンスのデータ管理ハンドブックというので、これは東大の宍戸先生と御一緒させていただきまして、個情委も入ってやっておりまして、現在、財産研のハンドブック、これは日本の大学全体に対してディプリケートされていくという構図になっています。

これが全体構図です。土木学会と同じような感じで、データが上から降ってくるということなのですが、いろいろな大きなデータ機器からデータが降ってきて、先ほどお見せしましたようなN I I データプラットフォームの中に入る。こんな無味乾燥したものは無理で、やはりT e a m S c i e n c e が支える。そして横から必ずこれをモニターしましょう。この世界はインターネットが始まったときと同じで、何が起こるのかさっぱり分かりませんので、しっかりと見極めた者が勝ちという、そうしたゲームになってくると思います。

そして、それを見ながらこのルールガバナンスを適正に持っていく。大体全体像としてはこうなっているというのが我々の見方で、先ほど言いましたように、六つの提言をさせていただいたということになります。

冒頭に申し上げましたように、アブストラクトデータタイプというのですが、データだけで存在するのではないのです。データの横に必ずそれをいじくるプログラムと一緒にってきます。そのインターフェースをA P I としてオープンにしたものがどんどん成長しているということで、この1例を少し御紹介いたしますと、これが東京大学がずっとやってきましたD A I S という先ほど申し上げました坂本さんや長野さんに随分御支援を頂いたのですが、見ていただきますと、何とかというプロジェクトが出て、そこから1期と2期、土木の先生にやっていただいております、僕は遠慮がちでファーストプレイヤーになりたくないののでやってくださいということで、3期目くらいになって、やはりI T 出てくださいと言われてやったのですが、僕がやったからではないのですが、ここに来てユーザーが増えまして、今は1万2,000人くらい入っておられます。しかもユーザーの半分より少し大きいのは海外になっています。

データを動かすというのは結局何なのだろうと思いますと、やはりこうした分野バイ分野のある種もうここに来ざるを得ないという大きなマグネットを持った場所を作る。ここで勝負が決まるのではないかというのが喜連川的な感覚で、こうしたものを日本でいろいろお作りいただくということが次の方向感ではないかと思っています。

なんでやねん、というのがよく分かりません。何でこう増えて、ちなみに先ほど3,000と言っていましたが、これの外数ですので、D A I S は今、1万5,000くらいの数にな

っているということです。

それから、最後、赤池参事官からジャーナル問題も何か言ってくださいとおっしゃられたので、これは難しいのですが、学術会議ではこの問題は一切議論しませんでした。これは非常に古典的な、もう出来上がった市場なので、データの方はこれからの市場です。2番目に書いてありますように、我々そうした会社さんが来られるときにお伺いしますと、データの方はどうなされているのかというのを聞く訳ですが、ノンライブラリービジネス、こっちからぐっと上がっています。逆に言うと、通常のライブラリーのところは、そこはグロースレーションが高くないということで、明らかにこちらが主力になってきます。

やっていることは何なのかということになりますと、やはりバイイングパワーを大きくするということが全てになってきますので、先ほどのデータ基盤もそうですし、SINETもそうですし、そうしたものをきっちり作っていくということで、このデータの空間だけは世界に取られないといえますか、記者の方がたくさんおられる中、出版者に余りペネトレイトされないようにする必要があって、限定的適用データの空間というのは、日本はフルで持っているのですけど、これは日本だけですので、これをどうするかということを考えていく必要がある。

NIIはもちろんお手伝いするのは全然やぶさかではなくて、我が方のオリジンは学術情報センター、要するに文献をどうオーガナイズするかということです。

どうも御静聴ありがとうございました。

○上山議員 ありがとうございました。

公開の議題はここまでとなりますので、記者の方はどうぞ御退室をお願いします。