

総合科学技術会議

第5回

ライフイノベーション戦略協議会

平成24年8月30日

内閣府 政策統括官（科学技術政策担当）付
ライフイノベーショングループ

午後 1時00分 開会

○北窓参事官 それでは定刻となりましたので、ただいまから第5回ライフイノベーション戦略協議会を開催させていただきます。大変お忙しい中ご出席いただきましてまことにありがとうございます。事務局を務めさせていただきます内閣府参事官の北窓でございます。どうぞよろしくお願いたします。

それでは、まず本日の協議会の配布資料の確認をさせていただきます。お手元の資料、上から2枚目に本日の配布資料の一覧がございます。資料1-1から参考資料までを資料としてセットしています。なお、別途机上配布資料を準備させていただいていますが、これにつきましては後ほど中野審議官から説明をさせていただきます。資料の過不足等ございましたら、お知らせお願いたします。よろしいでしょうか。

それでは、本日は大久保委員、原澤委員、小原委員、柳田委員、成宮委員、埴岡委員、菊地委員よりご欠席の連絡を受けております。本日は14名に参加いただいております、過半数のご出席をいただいておりますので、協議会開催の定足数を満たしていることをご報告いたします。

それでは、以降の進行を福井座長にお願いたします。

○福井座長 それでは、最初に前回の戦略協議会の議事録の確認をお願したいと思います。お手元の資料1-1でございますが、この件につきましては各委員、各省庁のそれぞれのご発言の部分について既にご確認いただいているということですので、特にここで何かご発言がなければ全体としてご承認いただいたということで進めたいと思いますが、よろしいでしょうか。それでは、そのようにさせていただきたいと思います。

また、議事録をもとに、委員の先生方のご意見を概要として資料1-2の表に整理いたしました。ご確認いただければと思います。

それでは、本日の議題の2に入りたいと思います。本日は医療、健康等に関する情報の利活用についてのご説明と、委員の先生方のご検討をお願したいと思います。前回の第4回ライフイノベーション戦略協議会からは、社会システムの変革・実用化に向けた隘路、基礎研究等について議論してまいりました。前回は「イノベーションの成果の実用化について」をテーマとして議論をお願いいたしましたが、時間が足りずに積み残しがございましたので、次回の第6回で引き続き議論をしたいと思っております。本日のこの第5回の会議では、ご案内のように医療・健康等に関する情報の利活用についての議論ということになります。

それでは、事務局から説明をお願いたします。

○北窓参事官 本日のテーマ、医療・健康等に関する情報の利活用につきましては、第2回戦

略協議会において委員の先生方からご提出いただいた意見から、情報に関連する話題提供を設定いたしました。前回の協議会でご案内いたしましたとおり、1つ目はデータベース統合利用センターの整備、2つ目は全ゲノム遺伝子等のビッグデータを取り扱うゲノムコホート研究です。データベース統合利用センターの整備につきましては、先月、JSTバイオサイエンスデータベースセンターで取りまとめられた報告書、ライフサイエンス分野の統合データベースの第2段階のあり方についてを、本日ご出席いただいております高木副センター長よりご説明いただきます。

また、ゲノムコホート研究につきましては、今月、日本学術会議ゲノムコホート研究体制検討分科会で取りまとめられた提言、ヒト生命情報統合研究の拠点構築、国民の健康の礎となる大規模コホート研究を、本日ご出席いただいております本庶副委員長よりご説明をいただきます。また、竹内委員より意見書を、庄田副座長から資料の提供をいただきましたので、それぞれご紹介いただくこととしております。

以上でございます。

○福井座長 ありがとうございます。それでは早速ですけれども、高木副センター長よりご説明をお願いしたいと思います。どうぞよろしく申し上げます。

○高木副センター長 今ご紹介いただきました高木でございます。

それでは、資料2-1、パワーポイントを印刷したものでございますけれども、それを使いましてライフサイエンス分野の統合データベース整備の第二段階のあり方についてご報告申し上げます。

第二段階のあり方ということがメインでございますけれども、まず最初に統合データベースとは何なのか、それを整備する意義はどういうことなのかということ、その背景を簡単にご説明させていただきます。資料をめくっていただきまして、まず1枚目でございます。左上と右下のページ番号が振られておりますが、1枚目でございます。

皆様よくご存じのように、ヒトゲノムは大体15年ぐらいかけて、数千億円のお金を費やして決定されたわけでございますけれども、その後その計測装置の伸びが非常にパフォーマンスが上がって、今ですと大体それが1日で決めることができると。それから、今どんどん価格が下がっておりますけれども、せいぜい数十万円のオーダーぐらまで下がってきていると。間もなく1人のゲノムを全部決めるのは10万円ぐらいのオーダーでできるようになるだろうと言われております。

この計測装置の伸びは、いわゆるコンピューターの性能の伸びを凌駕しておりまして、コン

コンピューターの場合は、ムーアの法則というのはご存じだと思いますけれども、18カ月で大体パフォーマンスが倍になるということがございますけれども、これを圧倒的に超えるスピードで性能が上がっているというわけがございます。そういうものを背景にいたしまして、ヒトゲノムに限りません。いろんな生物の、1万を超える生物のゲノムプロジェクトが進行していると。

それからもちろん、これは今ゲノムの話をししましたけれども、それ以外のいろんな細胞の画像でありますとか、生体の画像でありますとか、あるいは動画でありますとか、そういうものも急増しているという状況でございます。ライフサイエンス、まさに情報爆発、データ爆発が起こっているという状況でございます。

右下のほうに少し量的なことを書いておりますけれども、単位としまして大体データを今考えるに、petaのオーダー、10の15乗オーダーのデータ量が出ていると。これはもうすぐこの10の3けた上がりまして、10の18乗、exaの単位で数える時代もそう遠くないだろうというふうに言われております。

これに伴いまして、論文数、それからデータベース数も爆発的にふえています。論文に関しましては、年間100万までいきませんが、それに近い数が生産されているという状況でございます。もちろんデータベースもそれに伴いふえているということで、これらのデータ、あるいは論文、データベースを使いながら、どうやって研究を進めていくかということが非常に重要になってきているというわけがございます。

次めくっていただきまして、2枚目でございますけれども、じゃこのような傾向がいつごろから始まったのかといいますと、もちろんヒトゲノムプロジェクトが始まったのが1990年ごろでございますが、その当時はまだデータの伸びはそれほどでございませでしたけれども、その後1990年の終わりぐらいから、今申しましたいろんな計測装置の伸びがございまして、爆発的にデータがふえると。ちょうど我が国でも、これは我が国の例でございますけれども、ミレニアムプロジェクトということで、2000年ごろからここにありますような4つの省を中心にライフ関係のデータが生産され、それがいわゆるデータベースをつくるというプロジェクトが多数走ったというような状況になっておるわけがございます。これが今から七、八年前に調査されたデータで、少し古いものでございますけれども、歴史的なことをごらんいただくためにこの古いデータを持ってまいりました。

ちょっとめくっていただきまして3枚目でございます。今幾つか申しましたけれども、別の観点でまとめますと、ライフサイエンスはいわゆるデータを生産する、あるいはそのデータをみんなが使えるようにするということが、データベースをつくるということをも目的とした大型

プロジェクトが先ほど申しましたミレニアムプロジェクト以降活発になってまいりました。これを別の言い方をしますと、多額の研究費がデータになっていると。我が国のライフサイエンスに投入されているお金は3,000億とか4,000億というオーダーと言われております。科研費などを入れますともう少し、また1,000億以上上がると思えますけれども、そういうものが実際には計測装置を使ってデータになっているというわけでございます。

こういう状況の中で、その研究スタイルが従来の仮説駆動型からデータ駆動型、あるいは別の言い方をしますと、データセントリックとかデータインテンシブという言い方もありますけれども、データ中心科学。最近の言葉で言いますとビッグデータというような方向で来ているわけでございます。ですから、先に仮説を立てるというより、先にデータを全部とってしまつて、ヒトの遺伝子を全部決めてしまつて、その中からいろんなものを探していくというようなデータ駆動型になっておるというわけでございます。

こういうふうな研究の流れとは、ある種逆行するといひましようか、大学の法人化、それからバイドール法がありまして、データが大事であればあるほどデータを囲い込むという傾向になるわけでございます。

そういう中でデータの、あるいは知識の分断化・断片化が起きていると。私が申し上げるのも恐縮でございますけれども、イノベーションにはやはりいろんなデータを自由自在に使いこなしていくということが重要でございますして、いろんなものを再利用したり、転用したり、あるいはいろんなデータとデータを組み合わせて、違う種類のものを組み合わせて、マッシュアップと言ったりしますけれども、ということが重要だと。

ここでよく言われておりますのは、データをつくった人ではなくて、そうではなくて不特定多数のユーザーがイノベーターになり得ると。これも考えてみれば当たり前のことでございますして、ヒトゲノムを決めた人が1人で独占していても、それは余り価値を生みませんけれども、いろんな人が使うことによっていろんな病気の研究もできれば、進化の研究もできれば、発生の研究もできる、こういうことでございます。こういうようなことは、右のほうに資料がございますけれども、ビッグデータとか、あるいはデータ共有ということで、数年前から欧米でも盛んにこういうことの重要性が言われるようになってきたわけでございます。

次めくっていただきまして、4ページ目でございます。では、こういうような状況の中でデータ共有、あるいはデータを使いやすくすることがイノベーションに、特にオープンイノベーションと呼ばれるものが非常に重要なわけでございますけれども、そういう中で、データベースの状況を見ますと、我が国にも1,000を超えるデータベースがこれまでつくられてきており

ますけれども、似たようなものがありまして、どれを使っていいかわからないと。特に所在情報とか利用法がわからない。それから、そこに入っているデータがどの程度信頼できるものなのかということもよくわからない。

先ほど申しました大型プロジェクトがここ10年以上走っておりますけれども、それらがなかなか共有されるように使いやすくなっていないと。あるいはプロジェクトが終わるとそのままどこかに死蔵されてしまうという状況がございました。あるいはばらばらで、なかなかいろんなデータを簡単に組み合わせすることができないと。こういう背景には、我が国にはデータベースの戦略とか中核センターがないということが言われてきたわけがございます。これが今から七、八年前の状況でございます。

1枚めくっていただきまして、5ページ目でございます。こういうような背景がございまして、平成17年ごろより総合科学技術会議でもこの問題を取り上げていただきまして、調査研究がスタートいたしました。その中で各省もそれを踏まえて統合データベースプロジェクトというのは平成18年ごろからスタートさせました。文科省、農水省、あるいは経産省は少し遅く、20年ぐらいから始まったように書かれておりますけれども、類似したプロジェクトがやはり17年、18年ごろから始まっていると。厚労省さんにおいても類似のプロジェクトが走っていると。こういうような状況でございまして、これらをどういうふうに進めていくかということが議論され、それに基づいて各省でプロジェクトがつくられてきたと。またそれを総合科学技術会議で取りまとめ、連携をさせていただいて今日に至っているというわけでございます。

めくっていただきまして、6ページ目でございます。その当時、今から四、五年前でございまして、文部科学省のプロジェクトが一番規模的にも大きいし、扱うデータベースも大きかったものですが、文科省の統合プロジェクトが5年の時限で終わってしまう、18年から始まったものが5年間で終わるということで、その後どうしようかということでご議論いただきました。それで、そのときには内閣府の総合科学技術会議の統合データベースタスクフォースでご議論いただいて、こちらにいらっしゃる本庶先生が中心になって、こういうロードマップを書いていただいたわけです。平成22年の終わりに文科省の統合プロジェクトと、先ほどご紹介ありました科学技術振興機構 J S T にバイオインフォマティクス推進センター (BIRD) というのがございましたので、それを一体化すると。それから23年度から統合データベースセンターというものを走らせた。これは文科省だけではなくて、ほかの省とも連携しながら進めていると。まずはこの3年間連携を図って、25年度まで進めると。ちょうど今、3年プロジェクトの真ん中に来ておるわけでございます。このピンクの絵のちょうど真ん中ぐらいにいるという

状況でございます。

ただ、これはここにありますように、3年間まずやってみるということで、その後、26年4月からは第2段階に入ると。それはまた今後議論いたしましょうということになっておったわけですが、それが内閣府の体制等もございまして、議論をする場がなかなかなくて、きょうこの場でご紹介申し上げているという次第でございます。

こういうふうな活動をしておりまして、中身は後でまたもう少しご紹介申し上げますけれども、第2段階をどういう形でつくっていくかというのが今の課題になっておるといってございまして。

めくっていただきまして、こういうようなことを受けまして、先ほどちょっと時間軸戻りますけれども、文科省の統合プロジェクトとJSTのバイオインフォマティクス推進プロジェクトを合わせる形で、昨年4月にバイオサイエンスデータベースセンターというのが発足いたしました。これが7ページでございます。4つの柱がございまして、我が国でデータベースをどう整備していくか。戦略を立て、それから実際にデータベースのサービスをし、それからそのための将来に向けた技術開発と、それぞれの分野での統合化を図ると。この4つの柱で進めてまいりました。

8ページをごらんください。8ページが今申しましたバイオサイエンスデータベースセンターの機構図でございます。総合科学技術会議のご指導のもとにバイオサイエンスデータベースセンターを運営していくということでございまして、その中では2つのファンディングで、基盤技術プログラムと統合化推進ということで走らせてきたというわけでございます。

9ページにまいります。今申し上げましたことの少し繰り返しになりますけれども、9ページにありますように、この4つの柱でバイオサイエンスデータベースセンターを進めてきております。これは平成18年から始まりました文科省の統合プロジェクト、あるいはその他の省のプロジェクトとも連携しながら、それを発展させる形でこの4つの柱を進めておるといってございまして。この4つの中身は先ほどより少し詳しく書いておりますけれども、3番目が技術開発ということで、将来に向けたいろんな技術を開発すると。4番目が分野ごと、あるいはヒト、植物、微生物、その他のところで統合化を図るといって進めてまいっております。

10ページにまいりまして、この2つのファンディングプログラムは、10ページはその中の一つの基盤技術開発プログラムということでございまして。従来の文科省の統合プロジェクトになっていた情報システム研究機構のライフサイエンス統合データベースセンターがこれを受託したと、進めておるといってございまして。

その次めくっていただきまして、11ページでございます。分野ごとのデータベースの整備ということでございまして、23年度に10課題選ばれております。微生物、植物からヒト、あるいは医薬品関係ということで、非常に幅広いデータベースが選ばれ、それぞれの分野の統合化を図るというわけでございます。これによって、例えば植物なら植物をかずさディー・エヌ・エーが取りまとめをして、我が国の植物のデータベースを全部まとめていくと。微生物は東工大がまとめると。こういうような形で進めてきております。

24年度はもう一つプログラムが追加になりまして、これは理化学研究所でございますけれども、生命動態システム科学の1分子計測とか、そういうような分野のデータの統合化を図り始めているというところでございます。

12ページをごらんください。これまでずっと統合化といっても、なかなかその中身をご紹介できませんでしたが、非常に簡単に申しますと、12ページにありますように、それぞれさまざまなところで行われているバンク事業、それから各プロジェクトのデータベース、いろいろ個人、あるいは科研費等で作られているデータベース、それからさまざまな解析ツール、こういうものが非常に我が国で散在して、1,000とか2,000のオーダーあるわけでございますけれども、それらをまず全部カタログ化して一律に使えるようにすると。ここから入れば全部のデータベースなりツールが使えるようになります。

それから、2番目がデータベースやツールの使い方とか使い分けの情報、これは動画とかいろいろなマニュアルとか整備いたしまして、データベースを使えるようにすると。

3番目からいわゆる統合化ということと関係してまいりますけれども、データベースやツールを統一的、シームレスに使えるようにするというわけでございます。現在大体この3番目までがほぼ終わっておりまして、今はこの4番目、5番目の、先ほど申しましたようにヒトの病気の観点でデータを集める。あるいは先ほど申しましたアルツハイマーなんかございます。それから植物というような観点。そういうものでデータベースを整備していくと。これをイノベーションにつなげていくということで、今4番目、5番目というあたりに取り組んでいるという次第でございます。

じゃ、今どれぐらいのユーザーがいるかと、どういうデータベースが集まっているかということで、それが13ページから14ページでございまして、13ページは今これ2つのホームページを挙げておりますけれども、従来の文科省の統合プロジェクトで行っていたもの、それから昨年4月からできましたバイオサイエンスデータベースセンター、これは将来的には1つにしないといけませんけれども、まずは2つで今やっておりますけれども、2つでいろんなもののサ

ービスを行っておるといふわけでございます。

その中のメジャーなものが14ページに載っております、先ほど申しました段階を踏んでカタログをつくり、それから一括横断的に検索できるもの、それからさまざまなデータの権利関係等を整理したものであること、そこにありますような数のデータベースの整備が終わっております。

データベースに関しましては1,000を超えておりますので、大体我が国のデータベースを網羅しつつあるかなと思っておりますが、ただどんどん新しいデータベースもできておりますので、この活動もまた続けていかないといけない。それから、我が国にあるメジャーなデータベースの300ぐらいは一括して検索できるようになっています。

左下のデータベースアーカイブと呼んでおりますのは、これはやはり企業の方などからよく言われるんですけども、データベースを使いたいけれども、その権利関係とかいうのがよくわからなくなって使いづらいというお話ございましたので、そこはきれいに整理をいたしまして、権利関係、フォーマット、すべての標準化を行って、簡単にダウンロードして自由に使っていていいというのが今53データベースできている。これを今ふやしている段階でございます。ユーザー数が今約10万人弱ぐらい月におりまして、月間に230万ぐらいのアクセスがあるようなところまで何とか来ている次第でございます。

それから、15ページにまいりまして、今、文科省の活動を中心に申し上げましたけれども、最初のほうにご紹介しましたように、各省でいろんなデータベースの統合化が進められております。今、15ページに4つの省の、文科省、それから厚労省、それから農水省、経産省の4つの省の、いわゆる統合データベースプロジェクトのホームページをそこに挙げております。こういうふうになんかそれぞれが頑張っておりますけれども、これだけではやはり不十分だということで、この4つをまとめていこうと。別にどの省のデータベースを使いたいというわけじゃないでしょうから、我が国にあるものは1カ所で使えるようにしようということで、16ページをござらんください。

16ページ、これは昨年12月でございますけれども、この4省のポータルサイトを合わせた1つのポータルサイトをつくりまして、ここからすべてに入れるようなことで、4省の連携も十分進んできつつあるという状況でございます。

さて、ここまでこれまでの活動の背景、それから取組をご紹介してまいりましたけれども、先ほど申しましたように、今のプロジェクト、あと1年半で終了といたしましうか、一応第1段階を終えるということでございます、その後どうしていくかということが非常に難しい問

題でございます。そういう意味で、今後の課題といたしまして17ページに書いておりますけれども、今申しましたような活動、4省連携の活動も含めまして、どういうふうにこれを恒久的なセンター、組織につなげていくのか、拠点につなげていくのかということ、いろんなところで議論が進んでおります。きょうご報告いたしますバイオサイエンスデータベースセンターの運営委員会の報告もその一つでございますが、ここでもそういうことをご議論いただければというふうに思っています。

これが組織的な課題でございます、2点目はよくご存じのように、私、冒頭にも申し上げましたけれども、最近次世代型シーケンサーということで、大量のデータがゲノムデータを中心に出てまいりますけれども、そういうものをどう処理するのか。それから、イメージングデータも非常に大量に出てまいりますので、こういうものをどううまく扱うようにするのかということがございます。

それから3点目は、いわゆるゲノムコホート、この後、本席先生もご紹介されると思っておりますけれども、そういうものの動きも東北メディカル・メガバンクを含め、非常に活発になってきておりますので、そういうものと基礎生物学のデータをうまく組み合わせて、それを国民、あるいは研究者全体が使いやすくしていくにはどうすればいいのかということが問題になってきております。

その中で一つ大きな問題といたしまして、ヒトのデータ、これはプライバシーの問題もございますので、そういうものをどううまく、共有を図るとのことと保護を図るとのこと、これをどう両立させるかということも一つ課題として今取り組んでいる次第でございます。

こういうようなことを背景にいたしまして、18ページでございますけれども、そこにありますような活動をこれまでやってまいりました。この中には今申しましたヒトのデータをどう扱うか、あるいは4省の合同ポータルサイトをどう運営するか、あるいは国際的な、これはハッカソンと呼んでいますけれども、国際的ないろんなデータの標準化をどう進めていく。こういうような活動に取り組んできております。

こういう中で、第2期の計画ということで、JSTの中でも議論いたしまして、報告書を取りまとめた次第でございます。それがお手元にあると思っておりますが、こういう冊子である報告書でございます。これはとてもご紹介できませんので、この抜粋、中身だけを簡単にこの後、少しお時間をいただいでご説明いたします。

19ページをごらんください。19ページ、今、私どものバイオサイエンスデータベースセンターの運営委員会でございます。こういうようなメンバーで、23年10月からことしの6月まで、

計5回議論をしてみいました。その中で統合データベースのあり方はどうあるべきなのか、我が国の拠点形成はどうあるべきなのかということも議論をしてみました。その骨子をまとめたのが20ページでございます。

20ページにありますように、まず第1段階は23年度から25年度ということで進めてきたと。26年度以降どうするかということで、これは先ほど申しました内閣府のほうでのロードマップに従いますと、この3つの案があると。引き続きJST内の組織として拡充を図る、あるいは分野的に近い機関に移行する、あるいは新たな機関を設置する。こういうような3つの選択肢があるということで挙げられておりますけれども、私どもとしてはなかなか2番目、3番目というのも現実的に難しいということでございまして、今は3番目、私どもがこのままセンターを発展させていくとすればどういう体制が望ましいかということで、この報告書をつくっております。

その左下のほうに書いてございますけれども、第一段階としては今、大体予算16億円でやっておりましたけれども、第二段階としましては先ほどデータ量の爆発、それからいろんなゲノムコホートとの対応、イメージデータの対応、そういうものがございまして、大体予算としては倍で、人数としても100名規模のセンターにしないとなかなかうまくいかないのではないかと。ちょうどこの30億円といいますのは、先ほど申しました、我が国で投入されておるライフサイエンスの予算が3,000億とか4,000億でございますから、大体その1%ぐらいに相当するということも言えるかと思えます。

このことに関しまして、ちょっと1枚めくっていただきまして21ページに海外の状況をまとめております。当然ながら米国、欧州、あるいは最近では中国でもこういうようなデータベースの活動が非常に活発に行われていまして、大体米国、欧州、それぞれ20年ほどの歴史がございますけれども、人員としては五、六百名ぐらいの組織でデータベースの整備、統合化を図っているということでございまして、我が国としてもこういう大規模データの時代を迎えたライフサイエンスとしては、ある程度の規模で進めていかないと、とてもそのデータの中におぼれてしまうというような状況でございまして、それを利活用するためにこういうセンターが必要だろうということで掲げておる次第でございます。

それから、あと最後に22ページをごらんください。その中で先ほど申しましたように、運営委員会では第二段階のセンターのあり方ということで議論をしてみましたけれども、その中でやはりそれを担う人材、いわゆるバイオインフォマティクスの人材というのが我が国で非常に決定的に不足していると。ですから、データベースをつくるにしても、あるいはそれを使っ

て何かいろんな新しい発見をするにしても、そういう人材が欠けていれば幾ら箱を用意してもだめだろうということで、この運営委員の中で強い議論がございまして、そこにバイオインフォマティクスの人材育成にもぜひ取り組むと。あるいは、それをなかなかうまく流通させる、動かす、あるいはキャリアパスを整備する、あるいはいろんな認証、人材にある程度サーティフィケーションを与えるというようなことをやる機関がなかなかないものですから、そういうものを今のデータベースセンター、少しミッション違いますけれども、そういう中心になって進めていくのがどうかということで、いろんな大学や病院や、あるいは先ほど申しましたコホート研究を進めるようなところと連携いたしまして、もちろん企業等もそうですけれども、人材の流通、あるいは教育促進ということを図って行って、それがひいてはデータベースの整備、あるいは活用につながるということで提言がなされております。これに関しても私どもとしては積極的に取り組んでいきたいと思っております。

こういうような報告でございまして、何とぞよろしくご議論いただければと思います。以上でございます。

○福井座長 ありがとうございます。それでは、委員の皆様からご意見・ご質問を伺いたいと思います。30分ほどの時間がとれると思います。いかがでしょうか。はい、どうぞ。

○狩野委員 委員の狩野でございます。よろしくお願いします。

3つの点を教えていただきたいと思っております。1つ目は検索をするときの柔軟性の設計の仕方です。とりわけ新しい視点で学問をつくりたくなったときに、既存のデータを有意義に使えるかというところは、検索の方法の柔軟性にかかっているように思います。テキストマイニングや画像のパターン解析など、いろんな方法があり得ると思っておりますが、その辺はどういうふうにお考えかというのが1点目でございます。

○高木副センター長 おっしゃるとおりでございまして、いかに柔軟といいましょうか、検索できるかが非常に重要だと思っております。

ご存じのように、いわゆる検索にはグーグルがございすけれども、やはりグーグルですとライフサイエンスの専門用語等がわかっておりませんので、私どもとしては、例えば日本語と英語、あるいはいろんな表記のバリエーション、あるいは遺伝子名一つとりましても、大体表記の揺れ、あるいは別名を入れますと10、20のオーダー、同じ遺伝子でも名前が違いますので、そういうものの全部辞書を持っておりまして、検索できるようにしております。そういう意味では漏れがなく探せるんじゃないかというふうに思っておりますし、いろんな略語、その他に関しましても、世界トップクラスの柔軟な検索ができるシステムを私どもっております。

ちなみに私、自分のことを言うのはあれですけども、私はもともとテキストマイニングの専門家でございますので、そういう観点でも取り組んでおります。

それから、画像に関しましては基礎的な研究は大分進めておりますけれども、まだ実装するところには至っておりません。先ほど何度も申し上げておるように、非常に画像が大量に出ておりますので、それをうまく検索できるというのがこれから非常に重要になるだろうと思っております。それにどのような画像にデータを振るか、あるいは画像の中からどうやって意味をとるかということが一つの課題だと思っております。これはまだグーグルなんかもうまくいってないところで、私どもはこれライフサイエンスに特化していろんな技術を開発しておりますので、それも世界をリードできるものが実現できると思っております。

○狩野委員 続きまして二点目として教えていただきたいのは、今後の貴組織における第二段階の構造を考える上で、今までつくられたものを「守る」という方向性と、「新たなものを試していく」という方向性の2つの役割に対し、それぞれ必要な人材の人数比としては、いかがお考えでしょうか。今までは立ち上げの時期でおられたので、「試す」側に、かなり割合があったと思いますが、いかがでございましょうか。

○高木副センター長 これまでバイオサイエンスデータベースセンターができる前に4年ほどやっております。計五、六年このプロジェクトをやっております。それで先ほど申しました検索の柔軟性に関しましても、さまざまな技術が大分固まってまいりましたので、それに関しましては余り人手をかけずに、データ量の伸びに対応できるだろうというふうに思っております。今後は先ほど申し上げましたように、画像データでありますとか、あるいはいわゆるコホート研究などで出てまいります臨床の情報でありますとか、そういうものをうまく扱う技術というものをつくっていききたいと、そちらに注力をしていききたいと思っております。

またもう一つは、今、分野ごとに植物なら植物、微生物なら微生物、ヒトならヒトでデータベースをつくっておりますけれども、当然そこにそれらをまたさらに一段統合したらいいというところもございまして、それに関しまして私どもはRDFって、ちょっと技術的になりますので、きょうはご紹介いたしませんけれども、そういう技術を世界に先駆けて、それを使ってデータベースの統合化を進めるということをやっております。これは国際的にもイニシアチブをとっておりますけれども、そういうもので進めていきたい。

そういう意味では、既存の今までやっているものに関しましては人をふやさずに、新しいものに関してはこれから人をふやしてやっていきたい、そういう考えでおります。

○狩野委員 3点目です。アメリカのNCBIであれば、言語が英語であることから、どうや

っても国際公開になってしまうわけですが、日本語で進める場合にはそれがよくも悪くも日本語限定となるわけです。このあたりについてはどのようにお考えでしょうか。

○高木副センター長 これも以前からいろいろご指摘いただいておりますのでございまして、私どもとしては基本は英語で国際的に発信をしていきたいと思いますというふうに考えています。ただ当面、このプロジェクトのもともとの発端は、やはり日本人がいろんなデータに割とアクセスしやすいというようなことを目指しておりましたので、日本語に関しましても、例えば日本語の文献でありますとか、さまざまな学会の抄録等も日本語で書かれているものが多数ありますので、そういうものを検索可能にしておりますが、これは大体技術もほぼ確立してまいりましたので、今後いろんなものを、全部のサービスではございませんけれども、多くのものが今、英語でサービスをしておりますので、それは両方やっていくつもりでおります。

○福井座長 ありがとうございます。はい、どうぞ、桃井委員。

○桃井委員 特に高頻度単一疾患の解析等に極めて重要な環境因子で、環境省では子供のコホート調査などもスタートいたしました。この4省統合というのは縦割りでなく大変すばらしい方向だと思いますが、環境省とのデータの統合性、連携についてはどういう方向でしょうか。

○高木副センター長 残念ながら、まだそのあたりの議論は正直に申し上げて進めておりません。ただ、私どもとしてはこの4省だけで閉じてやりたいとか、そういうことは全く思っておりませんので、ある程度、特に今コホート研究を初め、そういうもののプラットフォームといいたいでしょうか、ひな形ができましたらいろんな方に呼びかけて、こちらの枠組みに入っていたければ、私どもとしては一緒にぜひやっていきたいというふうに思っております。

○福井座長 堀江委員、どうぞ。

○堀江委員 ただいまの桃井委員の意見と似ているのですけれども、臨床データとの統合を考える場合に、私ども産業保健という分野では原因と結果の関係というところに非常に注目しております。因果関係をつける上では、何が原因なのかというデータをあわせて持つておくと、爆発的にそれこそ解析、いろんな因果関係が出てくるということになると思いますので、全く同じ意見で環境省のデータ、特に日本というのは四季もありますし、東西南北広がっておりますので、こういったものとの連携というのは多分日本で一番よくできる分野じゃないかなと思っております。

あわせてですけれども、労働現場というのは、前回も申し上げたんですけれども、作業関係というのを測定しております。そういったデータ、まだ体系的に集めるような素地ができておりませんが、今後そういった方向に進めば、我々も含めて働く人間がどこでどんなも

のに暴露されていて、場合によってはその人の生活習慣も入れて、その次に臨床データが来て、最後にこういう遺伝子というところで、きれいに因果関係がつかのじゃないかなと期待しております。ぜひご検討いただければと思います。

○高木副センター長 全くおっしゃるとおりで、私どももそういうところに進めていきたいと思うし、この後、本席先生は多分そういうお話もなさるんだろうと思いますが、植物とか微生物とか、いわゆるヒト以外のデータに関しましては、当然今の環境といいましょうか、どこでとれたとか、さまざまなデータをある種の特定の用語、特定のフォーマットで記載しようという方向に来ておりますので、当然それは今、先生がおっしゃられたような方向に行くだろうと思っておりますし、私どもとしてもそういう方向に、まだ今そこは手をつけておりませんが、進めていければと思っております。

○福井座長 環境省の戸田室長でしょうか、お願いします。

○環境省 環境省でございます。子供の研究と環境に関する全国調査、エコチル調査を担当しておりますけれども、現在エコチル調査につきましては、10万人の参加者を集めるのにほとんどの予算を使っている段階でございます。データというのはこれから出てくるというように考えておりますので、そのデータの蓄積につきましては十分連携を図らせていただきたいというふうに考えております。

○福井座長 ありがとうございます。ほかにはいかがでしょうか。樋口委員、どうぞ。

○樋口委員 私はこういう統合データとか統計とか何も知らない人間なので、素人的な質問が2つですけれども、一つは途中に言及のあった、これが人に関する情報も統合していくという中で、やっぱりプライバシーという話がどうしてもある。それについても検討していますよという話だけだったので、一方でご存じかもしれませんが、厚労省のほうで医療情報等の個別法というのをつくろうとしているんですね。それで来月にはまず第1段階の検討会の報告書が出て、一定の方向性が出てくる。そこでも大きな問題になっているものですから、もう少し補足的に説明をしていただければというのが1点。

2つ目は、ちょっとこれは読み方だということか、あるいはほかの資料を見た上で質問すべきことなのかもしれませんが、本当に素人で、だからこういう乱暴なことを言うんですけども、データベースというのは、つまりある特定の目的でつくってしまうとそれだけにしか使えない。しかし、汎用性のあるデータベースというのは何にでも使えるというのは、実は何にでも使えないという意味になる可能性があって、そういう点が。

例えばこれ、さっきアメリカの情報のところにはアクセス数が何か4,000万件/dayみたいな

話が出ていて、それから第2段階になってこの図表を見ればですけれども、もうそんなことを考えているんだよということで、情報の喪失と利用の双方の意見を。つまり利用する側の意見も反映したようなものでないと、結局は本当は立派なデータベースだけはできているけれども、実際に使おうとしてみると極めて使いにくいというのがあり得るんじゃないかなと思っていて、そんなこと本当に素人の懸念にすぎないよと言ってくださればいいんですが、少し言葉を継いでいただければありがたいと思います。

○高木副センター長 お答えします。まず、1点目のヒトのデータの扱い方に関しまして、きょう時間がなくてご報告いたしませんでしたが、数回議論をいたしまして、ヒトのデータ、完全に統計的な情報はオープンにすると。インターネットでだれでもアクセスと。それから、一部は制限つきアクセスということで、申請していただいて、許可が出たらその人は使えると。あるいは、もう少し綿密な共同研究といいましょうか、そういうものを結んで使うと。その3段階ぐらいに分けておまして、その場合にそれぞれに関してデータのセキュリティーはどういうふうに管理すればいいのかとか、そういうルールはもう私どもではつくっております。それはほぼ完成いたしましたので、間もなくこれに従いまして、先ほど私ご紹介いたしましたけれども、私の今お渡ししました11ページに結構ヒトのデータがありまして、例えば一番上ですとヒトの脳疾患の画像データでありますとか、真ん中のほうにヒトのバリエーションと疾患との関係でありますとか、こういうデータがもう大分たまっておりますので、今、私どもが作り出したルールに従って公開と共有、あるいは制限をつけていくということを考えておまして、その準備はほぼ整ったというふうに考えております。

それから、先ほど先生がお話になったご懸念の点も、一応私もその委員会に入っておりますので、そういうことで発言はして、整合性がとれるようにしたいと思っております。

それから、2点目のデータベース、いわゆる汎用のもものと個別なもの、まさに先生のおっしゃるとおりでございまして、汎用にするとほんの簡単なことにしか使えないと。ある問題を絞ると少数の人にしか使えないと。こういうような状況がございます。そういう意味では、きょう統合というふうなことで申しましたけれども、統合データベースは私ども1つだというふうには考えておりません。ですから、非常に汎用に、みんなが使うような文献みたいなものはだれでも使いやすくすると。それから、ある分野に限ったデータベース、例えばがんならがんの研究者が使うようなデータベースは、それはそれで特化してつくっていくと。その2本柱でやっております。そういう意味ではそれなりに満足いくものができているんじゃないかと思っています。

確かにおっしゃるように、ユーザーの意見を反映しないと、単にデータをつくっているだけではだめでございますので、それに関しましては毎年数百名の方にアンケートをとりまして、これは若い研究者、それからシニアな方、それから製薬等の産業界の方、さまざまな方からアンケートをとりまして、どういうサービスはこういうふうにしたらいいとか、そういうアンケートをとってフィードバックしておりますので、そういう意味でも私どもとしてはある程度自信のあるものをつくっているという自負がございます。

以上でございます。

○福井座長 はい、どうぞ。

○吉岡委員 吉岡ですが、海外との比較というのが21ページの表があって、わかりやすくなっているわけですがけれども、この中に日本のものが並べてあるともっとよかったなと思うんですが、私なんかこういったデータベースが進んできているというのは余り知らなかったもので、どれだけ広報しているのかよくわかりませんが、海外のデータベースを利用するわけですね。

それと比べたときに、どこの部分が非常にいいものができてるのかというのをもうちょっと説明していただけて、ああこれはすばらしいなというのがわかるとありがたいんですが、いかがでしょうか。

○高木副センター長 確かに、日本の研究者の多くも、特に米国のデータベースを使っているというのが現状でございます。大きく言いまして、米国のデータベースは2つございまして、いわゆる論文のデータベースが非常に充実しておるものですから、ここを使っているというのが一つでございます。もう一つは、いわゆるゲノム関係のデータベースは古くから歴史がございまして、これを使っているというのが、そういう意味でこの米国のデータベースセンターは世界最大で、日本人も非常によく使っていると。

ただ、今申しましたように、この中に入っているのはいわゆる論文で、日本にさまざまあるいろんな日本語で書かれたものを含めて、活用できる文献が日本で使えるように必ずしもなっていない面がございますので、そういうものを活用するような点。

それから、ゲノム以外のデータ、さまざまなイメージデータでありますとか、日本の強みの分野の、例えばイネのデータベースでありますとか、ある分野の微生物のデータベースでありますとか、そういうものは日本がすぐれておりますので、今整備を図って使いやすくしているということでございます。あるいは、日本人の疾患に関しましては、日本人のデータを集めないといけないということがございますので、そういうもので集めているということで。

○吉岡委員 そうすると、日本にとってプラスアルファで有利になるというのは非常によくわかったんですけども、我々としては両方のデータベースにアクセスするようになさなきゃいけないのか、あるいは日本の今回のデータベースから海外のデータベースにもリンクして飛んでいくようなことができるのか。そこはいかがなんでしょうか。

○高木副センター長 私どもの先ほど日本のデータベースをまとめるということで一言ご紹介していますけれども、実は欧米のメジャーなデータベースも一緒に検索できるようになっておりまして、そういう意味で私どもを引いていただければ、欧米のデータベースも引けるという状況になっております。

○吉岡委員 わかりました。ありがとうございます。

○福井座長 それでは、庄田副座長、お願いします。

○庄田副座長 産業界もライフサイエンスのデータベース統合の必要性というのは、いろんな機会に申し上げてきております。その中で、例えば2005年のバイオテクノロジー戦略会議での議論の中でもデータベースのナショナルセンターというものをつくるべきだ、必要ではないかと提言しておりますし、産業界と各大臣と年に一、二度開催された官民対話、新しい政権では政策対話と呼称が変わっておりますけれども、その場でもライフサイエンスのデータベース統合というのは大変重要なことだということを申し上げてきております。そういう中で、第一段階から現在第二段階に進んでいるということで、非常に心強く感じております。

一つお伺いしたいのは、昨年、バイオサイエンスデータベースセンター（NBDC）が発足したということですが、このセンターが日本におけるライフサイエンスのナショナルセンターである、例えば国立電子図書館であるところか場ではっきり位置づけされてはいないのでしょうか。

○高木副センター長 私どもとしてはまさに今おっしゃったように、ナショナルセンターとして機能したいと思っておりますし、そのような努力をしておりますけれども、それに関しましては、基本的にはまだオーソライズされていないのではないかと考えています。もちろん4省連携等も進めておりますので、実質的なそういう立場にいつつあるというふうには思っておりますが、多分その議論は先ほどから何度も申しましたように、26年度から本格的に中核拠点を整備するということでございますので、その議論の中で今のこの活動を続けていくのがいいというふうにおっしゃっていただければ、それはそれでオーソライズされたことになるんじゃないかと思っておりますけれども、それはこれからの議論だろうというふうに思っております。

○庄田副座長 そのように申し上げているのは、現在は既存のデータベースの統合プロジェクト

トですね。ただ、これからも公的な資金を使っているような研究プロジェクトがデータを創出していくわけです。それが一つのナショナルセンターに行けば、国民の財産で皆さんがアクセスできるということになる。これが大変大事だと思いますので、今までのデータベースを統合することも大事ですが、これからの研究プロジェクトの創出されたデータというものが、一つ同じ統一された仕組みでどこかにあるということが重要になってくると思います。そういう意味で、私はぜひともこの現在の高木先生やってらっしゃるところが日本のナショナルデータベースセンターになるというのが現在までの経緯からいくと一番適切なのではないかと、これは個人的な意見でございますけれども、思っております。

○福井座長 アメリカでは随分前から、例えば疫学的データを研究者や学生が、最近では300ドルとか、学生は50ドルの場合もあるようですが、少額のお金を払ってデータを使わせてもらって論文を書くということが可能なようです。先生のところではお金を取ってデータを提供しているのでしょうか。ユーザーがどういうアウトカム、アウトプットを出してきているのかについてのデータはございますでしょうか。

○高木副センター長 まず、お金は現在のところ取っておりません。将来はいろんな検討をしないといけないことがあるかもしれませんが、先ほど申しました、例えばNCBIという米国のセンターを始め、どこも今無料でいろんなサービスをしておりますので、私どももそういう観点で無料でさまざまなサービスをしております。

それからもう一つ、データの受け入れという観点で、そこに皆さんがこのセンターにデータをつくった人が出していただかないといけないわけですが、それに関しましては米国ですと、例えば国のプロジェクトだったらNIHにデータをデポジットしなさいというルール、それをナショナルアセットにしなさいというルールがあるんですけども、我が国の場合は残念ながらそのところが不明確でございまして、なかなかこれまでデータが集まらなかったという状況もございますが、数年前より、例えば文科省ライフサイエンス課のプロジェクトは、公募要領のデータが出たらこのセンターに入れなさい。それからことしからですけれども、厚労省の科研費のデータが出たらここに入れなさいと。そういうような公募要領にいろいろ書いていただいておりますので、そういう意味でデータが集まりつつあると。それを無料で私どもが少し整理をしてお出しするという仕組みは徐々に整いつつあるというふうに思っています。

それからご質問があった、これによってどういうふうな論文が出たか。そのところはまだちょっと確認をしておりますが、先ほど申しましたように、大体今10万人ぐらいのオーダーの人が毎月使っているという状況でございます。

○福井座長 ありがとうございます。第二段階のあり方につきましては、最終的には総合科学技術会議で検討されるということです。この点につきまして、委員の先生方から何かご意見がございましたらどうぞ。

○狩野委員 たびたび恐縮です。人材育成のことを最後に書いておられましたけれども、例えば電子工学科に行く希望者がだんだん減っていたりとか、あるいは医療系だと国家資格があるほうにばかり人が行ったりとかいうような状況を見ておりますと、バイオインフォマティクスに行ったらどんなすてきな未来が待っているか、うまく表現しないと若者が行かないような気もいたします。どんなすてきな未来を、高木先生はお描きになりますでしょうか。

○高木副センター長 残念ながらそこがなかなか説得できないものですから、人が集まってないんじゃないかという面はあると思います。もちろん、どういうふうな教育プログラムで何をどう教えるかということは非常に重要な問題でありますけれども、やはり今の学生の行動、考え方を見てみますと、将来、今、先生がおっしゃったように、どのようなキャリアパスが開けるのかということ非常に彼らは見ていると思います。

残念ながら、今、例えば産業界でも、製薬企業でも、皆さん非常に大事だとおっしゃってくださるんですけども、とっていただける方が年に1人とか2人とかいうオーダーでございますので、そうすると学生としてはなかなか安定した将来の就職先というふうに考えられないというところがございます。あるいは、アカデミックでもそういう人のニーズは非常に高いんですけども、なかなかパーマネントなポジションが用意できないということがございますので、私ども今考えておりますのは、少しそのあたり私どもがバッファーになって、人が循環していて、余り二、三年で職がなくなるとかいうようなことを気にせずに、いろんなところで働いて、スキルを身につけていって、将来パーマネントなポジションにつける。それを応援するような仕組みを考えたいというふうに思っております。何とかそれを実現していきたいと。そうしないとやはり、私が言うのも僭越でございますけれども、日本のライフサイエンスはそれがうまくいかないとなかなか発展しないのかなというふうに思っておりますので、頑張っていきたいと思っております。

○福井座長 ほかにはいかがでしょうか。はい、どうぞ。

○成戸委員 今は統合データベースというか、既存のデータベースを集めて、いかに引きやすくするか、検索しやすくするか、あるいは利用できるようにするかという観点でご説明いただきました。本当は後の本席先生のご説明にも関係するかもしれないんですけども、データベースの集め方というか、いろんな病気の用語を含めて、どういう形でデータベースを集めるか、

あるいはデータをつくる段階からどうすれば統合データベースが価値のあるものになるかというのを、データベースセンター側から、積極的に関与していただきたい。お互いにかも知りませんが、そこまで関与しないと本当にいいデータベースにならないかと思うんですが、その辺のお考えについてご説明いただけますか。

○高木副センター長 きょうは少し説明を省略いたしましたけれども、先ほど申しました、例えば11ページにファンディングプログラムということで、幾つかのその分野ごとのデータベースをつくるプロジェクトがあります。ただ、これは既存のデータを集めるだけではございません。この横にデータを出す、実際にデータを出すプロジェクトと連携しております。あるいは学会と連携しております、そういう意味ではこれは過去のものを集めるというより、これから未来志向で出てくるデータを全部取りまとめていくということでございます。

一番顕著なものが11ページ一番下の生命動態システム科学、これはまだそんなにデータが出ていないわけですが、世界的にもまだどこがイニシアチブとるようなデータベースもできておりませんし、データのフォーマットも決まってないと。でも、まだデータが少ない段階で私どもと一緒にやりながらデータの設計をし、集め方も工夫しということで、まさに今ご質問あった方向で私どもは進めております。

以上でございます。

○福井座長 ほかにはいかがでしょうか。はい、竹内委員。

○竹内委員 後ほど私のほうから別途意見書として提出させていただいたものを説明させていただきますけれど、今、成戸委員がおっしゃられたように、私ども産業界の要望としましては、先ほどアンケートでいろいろ伺っていらっしゃるという話をお聞きしていますけれども、今後も産業界からの要望も設計、計画の段階から取り入れて頂き、より使いやすい、あるいは世の中の役に立つようなものを構築して頂くことを切に望みます。宜しく願いいたします。

○高木副センター長 実は製薬協からも大変応援をいただいております、そのワーキンググループと密に連携をとっております。それで、その中でいろんな、こういうデータベースがあったほうがいいのか、こういうツールがあったほうがいいのかはお聞きしておりますので、当然そういうことを視野に入れてといいましょうか、むしろそれを念頭に置いて進めていこうというふうに考えております。そういう態勢は今十分できていると思っております。

○竹内委員 今後とも、より連携をはかって頂きますよう宜しくお願いします。

○福井座長 ほかにはいかがでしょうか。先ほど庄田副座長がおっしゃいましたけれども、国のセンターとなるためには、第二段階のあり方の三つの選択肢の中の独立した新たな機関を目

指すほうがいいのではと個人的には思います。平成26年度までには実現可能な方策はないとさらっと書かれていますけれども、これは国としてそれだけの覚悟が固まっていない状況だと先生方は判断されているのでしょうか。

○高木副センター長 前のロードマップ、本席先生が中心になってお書きになったときも、多分ある種の独立した、ちょっと例えがいいかどうかわかりませんが、例えば国立国会図書館のような、あるいは先ほどデジタルセンターのようなお話もございましたけれども、そういうものをつくってはどうかというご議論がございましたし、今でもそういうことをおっしゃる方は多数いらっしゃいます。ただ、私どもが何のめどもなくつくれというのもちょっとあれなものですから、一応このセンターの運営委員会としてはそういうものがなかなか、私どもとしては今すぐどうこうということは申し上げられないので、私どもの考えられる範囲でこういう案をつくりましたということでございます。あのあたりもぜひこの場でも含めてご議論いただければというふうに思っております。

○福井座長 いかがでしょうか。アメリカでは国立医学図書館がPubMedなどをつくった過程で、国としてもものすごい投資をして、何百人ものスタッフを配置して、長い間実績を積み重ねてきております。日本はわずかなお金で、上澄みのところだけを持っていくようなプロジェクトが多く、もっとお金と人を投資して、世界に冠たるデータベースセンターをつくるぐらいの方向性が示されるといいと個人的には思っています。はい、どうぞ。

○青木議員 すみません。せっかくの話なのでぜひ教えていただきたいんですけども、22ページの人材のところに興味があるんですが、このインフォマティシャンという、インフォマテイクスという学問とインフォマティシャンというのの位置づけがいま一つ理解できないんですけども。私がこれ最初聞いたときには、経済におけるエコノメトリシャンという人たちがいて、その人たちは経済学のPh. D. を取るんですけども、トレーニングの間でちょっと特化するんですね。それで、その人たちは学者で論文を書くんですけども、一つは経済的な、例えば新しい賃金法の影響をデータを分析して論文を書く人たちと、あと統計的な研究というものもあるわけで、新しいデータ分析の手法というので論文を書く人たちがいるんですけども、今お話を伺ったところのバイオインフォマティシャンというのは、むしろ技術者みたいな方々なんでしょうか。

○高木副センター長 これは説明をはしょっていて恐縮でございますけれども、いわゆるバイオインフォマテイクスと呼ばれる人たちは3種類か4種類に分けることができます。一つはデータからいかに、最近の言葉でいいますとビッグデータからいかに新しい知識を見出すかとい

うような、ある種の新しい統計的手法もつくったり、いろんな解析、アルゴリズムをつくったり、あるいは非常に新しいデータベースをつくっていくと。それが研究ベースでそういうことをやっていく人たちが必要でございます。

それから2点目は、そういうような新しい技術はつくらないけれども、既存の技術なり既存のデータを使っているいろんな現場の現実の問題を解いていくというような人たち。それから、3番目がそのためのデータベースの論文を読んで少し整理をしたりとか、そういう人たち。いずれにしても我が国では余りいませんけれども、いわゆるPh.D.、ドクターを取った人が欧米では大部分そういう仕事をしておりますので、その3種類どれもですね。我が国としてもそういうような人たちが働ける場をつくりたい。

特に今足りないと言われていたのが、私が申しあげました2番目の、新しい最先端の研究をするわけでは必ずしもないけれども、今あるデータベースやツールやいろんな解析手法を使って、現実の疾患・病気と遺伝子の関係を調べていくとか、そういう人材が今圧倒的に足りないと言われておまして、大規模なコホート研究と疫学研究とを進めていく上でも、そういう人たちが不可欠なのになかなかそういう人がいないということでございますので、そういう人を育てていきたいと。

ただ、そういう人たちは生物学から入ってくる人もいれば、情報系から入ってくる人も両方います。欧米では割と生物系から入ってくる人が多いんですけども、我が国ではなかなか生物系から入ってくる方は、数学とか統計というのが少し苦手という意識がございまして、どうしてもやはり情報系から入ってくるのが今期待されているという状況ではございます。

○青木議員 確認なんですけど、そうすると、欧米では生物なり医学の研究の大学院教育の課程でそういう技術を身につけるカリキュラムがあるということなんですか。

○高木副センター長 はい、そうでございます。あるいはそれを専門に教える学科とか大学院がございまして、それが例えば情報系から入ってくる人、物理系から入ってくる人、化学から入ってくる人、あるいは生物から入ってくる人がそこで数年間トレーニングを受けてそういう分野に入っていくと、またその就職先があるというようにいい循環ができておりますけれども、我が国の場合は、そのところがなかなか今欠け落ちているというか抜け落ちているところがございまして、それをうまく立ち上げていきたいというのが今回の趣旨でございます。

○福井座長 ありがとうございます。

それでは、時間も過ぎてしまいましたので、長い時間、ご説明と討論ありがとうございます。

続きまして、日本学術会議ゲノムコホート研究体制検討分科会で取りまとめられました提言「ヒト生命情報統合研究の拠点構築（提言）－国民の健康の礎となる大規模コホート研究－」につきまして、本庶副委員長よりご説明いただきたいと思います。よろしくお願ひします。

○本庶副委員長 ありがとうございます。

お手元に「提言」という冊子がありますので、基本的にこの冊子だけ資料といたしておりますので、この内容をかいつまんでご説明させていただきます。

このライフイノベーション戦略協議会の年間スケジュールを拝見しますと、既にアクションプランということに参画されておりますので、アクションプランの中でゲノムコホートというものの位置づけにつきまして十分ご理解をいただいていると思いますが、あえて補足させていただきます。

第4期科学技術基本計画の特色は、分野別の推進方策から課題解決型へという大きな変換を行ったということ、それから、基礎研究の重要性を強調したということが、私個人的には大きな特色であると考えております。

ライフ分野における課題はたくさんございますが、中でも最も大きなものは、この医療費の無限増殖的な増大であります。40兆円、国家予算の半分以上、税収を超えるような医療費が毎年使われるという状況につきまして、この課題をいかに解決するかというところからスタートして、予防医学というものを強化し、その成果を医療に役立て、発症そのものを抑えることができれば、当然医療費の削減につながる、こういうことから、アクションプランの中で先制医療ということ、その手段としてゲノムコホート研究というのが推進されているわけであります。

学術会議では、このような政府の取組が始まっている中で、どのようにしてこのゲノムコホートを国の中心的な課題として立ち上げ、そこにかかわるいろいろな分野の研究者がどのような役割分担をし、そしてそれが最も効率よく社会に還元されるであろうかということを検討するために、この冊子の1枚めくっていただきましたところに委員の名前の一覧表が出ております。主としてライフサイエンスにかかわる方々、医学を含めた方々を中心に議論をいたしました、その結果をまとめたものでございます。この提言の非常な大きなポイントは、スタートはゲノムコホートという観点から出発いたしました、実はこれを追求していくことによってヒトの生命情報が統合的に理解できるようになるであろう、また、そうすべきであるという提言に至ったということでもあります。

ライフサイエンスというものの出口が何かといいますと、それぞれの立場の違いがあるにせよ、「ヒトとは何であるか」ということを理解するということが最も根源的なライフサイエン

スのターゲットであります。また、そこに基づいて疾患の解明、予防、治療ということも可能になってくるわけでありますが、大規模ゲノムコホートをきちんとやることによって、これは必ずそういうふうなヒトの研究ができるようになる。また、そのことが最近の非常に大きな技術的な革新によって可能になってきた。従来の生命科学はほとんどネズミとかモデル動物を使った疾患研究ということでありました。それはなぜかという、ヒトの材料あるいはその情報を非侵襲的に得て、それから何らかの生命現象、疾患現象を理解するということがほとんど不可能なほど技術的に未発達でありました。しかし、今日では、1滴の血液から膨大な量の情報を得ることができます。これは機器分析技術の進歩によるところが大きいわけでありまして、先ほど高木先生がおっしゃいましたように、ヒト1人のゲノムのコストが既に数十万円のレベルに達しておりますし、これは大量に使われることによって恐らく、さらに3分の1くらいに下がることは極めて近い将来に起こるであろうと多くの人が予測いたしております。

こういう背景の中で、まず、このレポートの最初のところ、要旨のところを中心にご説明させていただきますが、現状における問題点、これまでの我が国の歴史等について書いております。

これまでの歴史を眺めてみますと、我が国におけるコホート研究というのは、原爆被爆者の追跡調査が1947年に始まったことから始まり、多くのコホート研究というのが行われておりますが、その規模、目的も大小様々であり、特定の疾患に限定した研究というのが非常に数多く行われております。

現在、我々がヒアリング等々で確認した限りにおきまして、いわゆるゲノムコホートとして現在進行中のものは、2002年に「久山町コホート」として九州大学を中心に行われているものと、2005年から始まった「長浜0次予防事業」という京都大学を中心に行われているものの2つであります。

先ほどご紹介がありましたように、我が国の政府は、東北メディカル・メガバンク事業として新しいゲノムコホートを開始しようといたしておりますが、まだその中身は明確ではありません。むしろこの時期においてきちんとした全国的な体制をつくる必要があるというのがこのレポートの中心的な課題であります。

なぜ、全国の中核拠点が必要かといいますと、ここによって統一的なデータの取り方、これをきちんとしないと、AコホートとBコホートのデータを合わせることができません。すべてのデータが全く相互に解釈できない状況では意味がなくなりますので、なるべく早く中央的な、全国統一的な基準の策定に既にスタートしているコホートもこれからスタートするコホートも、

みんなが参加して納得できるような、そういう仕組みと具体的な基準をつくるべきであるということであります。

それで、さらに現状の大きな問題点は、先ほど樋口先生からご指摘がありました倫理指針の問題であります。これは私、ゲノムコホートの開始について口をすっぱくしてお願いしたのですが、この点は非常に大きな課題として残っております。

3省倫理指針というものの改訂が現在進んでおりますが、その内容は多分、そのままゲノムコホートに適用することは不可能な内容であり、研究を進めていく上では非常に不十分なものであります。

現在先行している2つのゲノムコホートで行われている方式は、地方自治体はその住民との約束を取り、または地方自治体が研究機関との契約を結び、個人情報保護に関しての責任を地方自治体が担保すると、そういう形で初めて包括同意を取得することが可能になっております。包括同意という言葉は初めての先生方もいらっしゃると思いますが、この提言の中に大きく強調いたしております。どういう意味かと言いますと、ゲノムコホートをやるということは、その方のDNA配列を含めたすべての健康情報を取得するということであり、現在健康な人が将来、どのような病気になるかということは予測が付きません。ここが疾患特定のコホートと非常に大きな違いでありますので、自分の健康情報、生命情報をどのような医学研究に利用してもらってもよろしいという同意があること、これが包括同意でありまして、これをとるということは非常に困難であり、また、しかし、それをやらないゲノムコホートというのは意味がないといえますか、目的を達することができません。そういうことをやるためにも、きちんとした全国体制が必要であるということでもあります。

それから、このレポートで提言いたしておりますことを、最初のiiページの3のところ、(1) (2) (3) (4) (5)と順次申し上げておりますが、このゲノムコホートをやることによって、第1、ヒトの生命情報統合研究の創出が可能になるから、これをやるべきである。また、それを可能にするためには、単に医学研究としてこれをとらえるべきではない。これは、新しい情報統合、先ほど高木先生がお示しになりましたように、これは次元が10の6乗ぐらい多分優に上がるほど巨大なデータベースになります。その中からどのような情報を、どのような組み合わせで取ってくるか、そういうふうな情報学自身の変革も多分要請される、それで初めてこの大きなデータが生きたものになってくる。情報科学の変革もやっていく中で要求されるであろう。また、その変革こそが新しいイノベーションにつながるであろう。非浸襲的にデータをとるということは、現在可能になっておりますが、完全ではありません。そこにおいて



新しい機器開発、このデータというのはMRIとかCT、PETによる画像解析のデータも含めておりますし、それから血液成分、あるいは尿中のあらゆる化学成分を微量で定量的に分析するということであり、現在可能なものと不可能なものがあります。そういうふうな新しい機器開発をやっていく。ここにおける中心的なプレーヤーは、生物学の人ではなく、エンジニアリングであり、物理化学の専門家である。そのような非常に広範な人の力を結集してこれをトップレベルのゲノムコホートにしていくと、そのこと自身が必然的にヒトの生命情報を現在この規模として提言しているのは100万人であります。なぜ100万人かということは、ご質問があったときに説明させていただきます。従来なぜヒトのデータが使えなかったということは分析が難しかったと同時に、ヒトのゲノムが全部違いますから、AさんとBさんを直接比べられない。しかし、100万人を集めて、その中で比較していくことによって統計的な処理と遺伝的な相関との中でその問題は解決されてくる。また、そのことこそがヒトの生命科学であると、こういうとらえ方をして、これをぜひやるべきだということが（1）であります。

第2は、既に先ほどからご指摘がありますように、我が国においては膨大な医療機関におけるヒトそれぞれのデータがあります。これをいかに効率よく集めてくるか、これに関してはいろいろな法規制の問題もあり、次のページ、ivページでございますが、①として、医療情報システムの標準化、それから国民保険番号、こういったものの法整備、それから医療情報の追跡基盤というのは、マイカルテのように、1つの番号で、どこの病院に行かれても、その診療情報が集約できる。そういった仕組みをつくるということが非常に重要であろうというふうに申しております。

（3）につきましては、このレポートの15ページを見ていただくのが一番おわかりになりやすいと思いますが、全体の構成として、先ほどご紹介いたしました図1の中央のオレンジ色の中核拠点、これをきちっと立ち上げ、ここの中枢的な機能をしっかりさせることが重要である。ここでやるべきことは、先ほど申し上げた統一プロトコール、それから倫理指針、産学連携の知財、こういったものを一体的にやるべきであろう。それから、当然ながら、その真下に書いてありますデータ解析センター、これは2種類ありまして、実際に分析をするところと、そのデータを集約していく、先ほどの高木先生がご提言になったところは、まさにここに入る、ここの部分を受け持っていただくことを想定しているものであります。それから、左側に地域研究拠点、この中に先ほどご紹介した東北メディカル・メガバンクとか九州とか長浜とか、それぞれ地域に根差した現場のコホート、その健常人をそれぞれ何万人かずつ登録していく。そのところで試料の収集、地域住民の健康管理、健診、健康づくり等の住民へのリターンも含めた

地域拠点の構築ということが求められる。そして、そこから出てくる生体試料をバンクとしてどのように扱うか、これは基本的な考え方として、その地域の拠点に1つ、中央に1つという考え方と、全部を中央化し、中央を複数個にして災害や危機に備えるという考え方もあります。これは今後中核拠点を中心にすべてのステークホルダーと合意をすべきものであるというふうに提言いたしております。

それから、もとへ戻っていただきまして（4）この包括同意がとれるかどうかということがゲノムコホートの成否を決めるものであります。これをやるためには、国民のこの予防医学に対する理解というものがなくして不可能であります。現在の地域自治体による個人情報の秘密保護という担保のやり方は一つのやり方であり、現実的に機能いたしておりますが、本来であるならば、国全体としてこのようなことがもっとスムーズにできるような倫理指針を目指していただくということが望ましい。

（5）は、これがきょうここへ私が出てまいった最大の目的でございますが、このような仕組みをスタートするために、まず推進協議会を立ち上げ、そしてその中で先ほどの15ページにあるような全体像を具体的にだれが、どこで、どのようなことをするのかを決めていくべきであろうという提言をいたしております。これにつきましては、引き続き学術会議としてはこの推進協議会を設置し、前へ進めていきたいと思っておりますが、総合科学技術会議としても、特にこのライフイノベーション戦略協議会としてここに一緒に参画いただき、そしてナショナルプロジェクトとしてゲノムコホートを推進していくことにご支援をいただければ大変幸いですというふうに考えております。

それから、1つ飛ばしたところがありますが、実は、15ページの絵のところ、「産学民コンソーシアムの構築と運営」がこの中核拠点の一番下を書いてありますが、これはどういうふうなことを考えているかということが19ページに絵がございます。これは、そこに「情報統合研究中核拠点」と中央に書いてあるのが先ほどの15ページの図の中核拠点と同じ概念のものであります。そして、ここの右側に「企業」という形で紫色で書いてありますが、ここにはさまざまな業種の企業が参画され、そしてコンソーシアムをつくり、相互にwin-winの関係で新たな課題にチャレンジする企業を包含した共同研究方式をとるのが望ましいという提言をいたしております。当然、先ほど申し上げました情報の分野、機器解析の分析、それから医療機器の分野、もちろん情報をもとに新薬を開発される医薬品産業の方々、多くの企業の参画が望まれるのではないかと、そういうものをうまくマネジメントして利益を還元しシェアしていく体制が国として望ましいのではないかと、そういうことも含めた提言でございます。

以上、簡単ではございますが。

○福井座長 ありがとうございます。

それでは、15分ほど質疑応答の時間をとらせていただきたいと思います。いかがでしょうか。
樋口委員。

○樋口委員 私、早退をさせてもらうので初めに意見というか、昨日の検討会というのが厚労省でありまして、今の先生のご報告、お話と非常にオーバーラップする、まだこれは報告書の案なので、成案にはなっていないのですけれども、2カ所だけご紹介をします。

医学研究、つまり何の検討会かという、もう一回繰り返し言うと、医療等の情報に関する特別法を個人情報保護法の中でというか、枠外で、とにかく特別につくろうという、そういう動きの一つですけれども、とにかく医学研究、学術研究については追跡調査を行える医療共通番号を活用することは大きなメリットがあると、こういう話と、それから本人同意のあり方についてというところですが、今の案では、この医療等の分野では、公衆衛生や医学研究等においては、同意原則を硬直的に適用すれば、サーベイランスや研究そのものが成り立たず、結果として、患者が享受する医療の向上等に支障が生じることになりかねない。それで、ただなかなか同意は要らないよというのは簡単には言えない話ですから、本人同意のあり方については、具体的ケースを想定して検討を深めるべきである。個別の明示の同意が必要か。院内に例えば掲示をする、病院に掲示をするような、その掲示をもって黙示の同意があるということでもいいのか。今でもそういう部分はいっぱいあるんですが、安全に匿名化するという方法で対処するか、あるいは公益目的で、将来活用することも含めた包括的同意を得るかなどの組み合わせをケースごとに考えていくべきであるというような、今のところは報告書案になっております。ちょっとご紹介だけしました。

○福井座長 ありがとうございます。

いかがでしょうか、今の樋口委員のおっしゃったことに関連してでも結構ですし、それ以外でも結構です。

これは、3省合同の倫理指針が、文科省で止まっているのですが、それとは関係のないところでのお話でしょうか。

○樋口委員 今度の特別法は、まず指針ではなくて法をつくろうという話と、それから医療等の情報の中には、当然ゲノム情報やなんかも一々区別しないで全部適用範囲には入るだろうという前提であります。今のところですが、もちろんだろうなるかはこれからの問題なのでわからないんですけど。

○福井座長 わかりました。

それでは堀江委員、どうぞ。

○堀江委員 今の個人情報の取り扱いをする際に、研究者側が持っているIDと、それがだれであるかという読み替え表とといいますか、これが持っている場所とといいますか、そういう人を個人識別情報管理者とといいますか、そういう方々がおられると思うんですけども、現在の考え方では、研究者の中にそういった人を置くのではなくて、研究者はあくまでもIDしか知らない。それがだれであるかは、場合によってはそれをその方に読み替えていただいて本人へのサービスに生かすとか、あるいは追跡した情報を研究者が取得するといったときに発動する管理者を外に置くというのが多分今の考え方だと思うんですけども、今後もそういう形というのはやはり堅持していく方向になるんでしょうか。ちょっと樋口先生への質問になってしまいますけど。

○樋口委員 私が適切に答えられるかどうかはあれですけども、今のお話は、いわゆる連結可能匿名化で、それを匿名化していると言えらるためには、そういうようなシステムをつくっておく必要があるよという話ですよ。極めてある意味では形式的な取り扱いをしていて、その匿名化のあり方についてもいろいろ問題があるから、これまでのやり方、その形式的なやり方を、つまり同じ研究機関じゃなくて別の大学でも別のところに置いておくと統合する、府庁と言ったらいいか、そうすると、もうそれで匿名化ですよというようなのはすごく形式的な話なので、やはりそういうようなことで済ませているのではなくて、もっと実質的な意味での匿名化ということをやるのであれば考える、あるいはここまでやればもう十分匿名化されていますよという話を明確にしてあげるとか、いろいろなことが課題として残っているというような内容の報告書です、まだその段階は。

○福井座長 ありがとうございます。本庶先生に何かご質問ございませんでしょうか。

桃井委員、どうぞ。

○桃井委員 桃井でございます。

この15年くらいでしょうか、医学系の先生方にはご承知のとおり、精神疾患の developmental origins of adult diseaseということが多くの疾患で言われておまして、精神疾患のみならず生活習慣病も含めてさまざまな臓器の成人発症疾患が、小児期というよりは、むしろ胎生期も含めた発達期オリジンであると、そういうエビデンスが大変蓄積されてまいりまして、例えば、生涯のストレス耐性を決める根幹の一つであるステロイド受容体の発現は出生早期に決まって、それは生涯持続する。ですから、そこに介入できるとすると、出生の早期

であるというようなデータも蓄積されてまいりました。ですから、11ページに書いてあるように、小児期発症の小児期の疾患のための小児期からのゲノムコホートではなくて、成人疾患も含めた全年代の疾患の発症も含め、特に胎児疾患であります、そういう視点での小児期のさまざまなサンプル収集も含めたコホートでないと、特に他国が成人を中心のコホートでありますので、日本のコホートが特徴を持って予防、あるいは根本的な介入薬、介入法をそこから創出するためにも、やはり次世代期から、出生後ではなくて、私が申し上げるのは、体内期も含めてですけれども、そういうコホートのプランニングがぜひ必要であるというふうに思います。そういうデータが蓄積されてきたのはこの15年ほどでありますので、まだなかなか多くの領域の先生方にはご承知されていないと思いますが、しかし、かなりのエピゲノムプロファイルは小児期早期に決まる。日本でいう三つ子の魂百までというようなものがゲノム、あるいはモレキュラーのデータで蓄積されているということを反映したゲノムコホートについては、この中でご議論されたでしょうか、あるいはもし余りご議論されていなければ、ぜひ日本特有の、よりよいコホートプランとしてお考えいただきたいと思います。

○本庶副委員長 先生にこれを申し上げるのは文字どおり釈迦に説法ではありますが、既にエコチルドレンという環境省が年間40億を投入して数年前からコホート研究をされておられます。その開始に当たり、当時私は、これをゲノムコホートにしてくださいということを強く、現在の担当官にも申し上げました。しかしながら、現状はゲノムは入っておりません。なおかつ、先ほどご紹介があったように、まだ解析は道半ばであります。私はそういう点では、国の財政が逼迫しておる中で、この40億をゲノムなしにこのまま継続することが適切であるのかどうか、むしろこの場できちんとしたご議論をいただきたい。

○桃井委員 私も全くそのように思います。

○福井座長 ほかにはいかがでしょうか。

はい、どうぞ。

○庄田副座長 先ほど本庶先生から、なぜ100万人規模のバイオバンクかということについて質問があればということでしたが、たまたまきょう参考資料でつけております製薬協の医薬産業政策研究所の「バイオバンクの現状と展望」の中で、世界的にも100万人以上のバイオバンクというのは英国のMillion Women Studyと米国のMillion Veterans Programの2件と報告されておりますので、この100万人という規模の理由について教えていただけたらと思います。

○本庶副委員長 簡単に言いますと、これは多いほどいいに決まっております。ですから、アウトカムがどれだけ期待できるか。結局コストパフォーマンスの問題になりますが、このお手

元の提言の最後、23ページのところに参考資料というものを示しております。それで、これはその代表的な疾患が50万人にした場合と100万人にした場合に10年後、あるいは20年後に現在の大きな発症率から予測してどのくらいの人数が集まるかということを示しております。私の理解では、やはり数千人規模の患者が発症することによって、これは福井先生のご専門ですけれども、一定の統計的に有意なデータになり得る。それはケース・バイ・ケースですけれども。

そういうことを考えますと、この表を見ていただきますと、100万人した場合、例えば心筋梗塞の場合、100万人にした場合に、10年後に7,000人、そして50万ですと3,000人。つまりどうということかといいますと、100万人にして10年でデータが十分意味があるものがとれるのか、あるいは50万人で20年待つのか。そして、その間の投下資金は一体どれくらいかということを考えますと、我々はむしろ100万人で10年で、もしデータ解析が可能であれば、そのほうがコストパフォーマンスはいいのではないかと。なおかつ、現在これとの競合相手としてはUKバイオバンクというのがございます。UKバイオバンクは50万人であり、その登録は既に完了しております。ところが、人口はUKは日本の半分でございます。それから考えましても、日本で国家プロジェクトとしてやるならば、100万人を目指すべきであろうという議論でございます。

○福井座長 ありがとうございます。

ほかにはいかがでしょうか。

中核拠点というのは、大体先生のイメージされる場所では、どのくらいのスタッフでできるようなものなのでしょうか。

○本庶副委員長 この中核拠点というイメージは、かなり幅があります。私の個人的なイメージは、ここはそんなに大きくする必要はないと思っております。それは、この中に、例えば情報統合データベースを入れ込むとすると、これは先ほどの高木先生のご説明のように大きくなります。しかし、私はそれは必ずしも入れ込む必要はなく、現在既にある情報統合データベースに参加していただいて、その部分を受け持ってもらって。ほかのお仕事もたくさんありますけれども、それを活用という言い方は適切かどうか分かりませんが、既にある機能を持っているところを活用していく。中核拠点というのは、かなり意思決定とか全体の統合とか、産学連携とかはばらばらにいろいろな約束事で地方でやられると、これは非常に産業界も迷惑されますし、国としても困る。ですから、そういうことはきちっとセントライズしていかないと。ですから、主な仕事はそういう委員会をきちっと集めて、そして約束事を決めていく

ということを主な仕事とすれば、これはでっかいビルディングをつくる必要はないと思うんです。かなりソフト的なパワーで十分できる、私は個人的にはそういうほうがうまくいくんじゃないかと思います。

○福井座長 ありがとうございます。狩野委員。

○狩野委員 よく練られたアイデアだと思って伺いました。実現に当たり、どのぐらいのお金がかかるかという議論をここではきっとしなければいけないということが一つです。例えば、私、最近、岡山へ移りまして、同地ではこれに使っていただけそうなデータベースの構築が既に進んでいるということを承知しており、同様に他の各地でも存在するかと思います。そのときに、この中にも書いておられますけれども、電子カルテの標準化を含めたような、すなわち全体としてまとめるに当たって必要な項目がふえてきた場合に、それをどのように後から実装できるかという部分に、どのぐらいのお金がかかるかというのも算定しなければいけないかと思いますが、あわせてどのような規模を、どういうふうにお考えでしょうか。

○本庶副委員長 これは先生おっしゃるように、何をどこまで含めるかによって非常に大きな違いがございます。当初、数十万人規模でこのウェット部分のゲノムコホートの概算をしたときでは、エコチルドレンの予算規模と大して変わらないぐらいでできるであろうというふうに考えておりました。しかし、総務省絡みの情報関係をどのようにするのか、このところはちょっと我々では推計が難しく、今のところ入っておりません。したがって、ざっくばらんな話、東北メガバンクのお金の規模で私は十分足りると思っています。

○福井座長 竹内委員、どうぞ。

○竹内委員 15ページの図の中核拠点のところの上から5つ目の中点（・）に、知的財産の管理という項目が記載されています。これは、その下に記載されている産学官コンソーシアムの構築と運営というところにかかわってくるのではないかと思います。知財に関してはある程度コンソーシアムに独占権を与えて運営されるのか、あるいは、公的な施策ですので、使いたい方々にそういう独占権は関係なく幅広く使えるようなシステムになるのか、どちらなのかということをお伺いさせていただきます。

○本庶副委員長 非常に重要な問題だと思って我々も議論いたしました。企業にとって、ここに参加するメリットは何かということになります。それがなかったら、だれもお金を出してここに来たいという企業はないと思いますので、当然、企業に何らかのメリットが還元されるような仕組みがビルトインされなきゃいけない。しかし一方で、企業から来るお金で全部をやるわけではない。19ページの図をごらんいただきたいんですが、公的資金、もちろん個人的な寄

附等々の国民のお金ということは当然入ってきますから、その部分に対しての透明性ということ担保しなきゃいけない。しかし、その中で企業にどのようなメリットが与えられるか、これは非常に難しい問題であります。従来、国のプロジェクトでそのことをうまく整合性をとりながらやっておられる例はありますので、私どもは一般則をここの中核拠点で決めて、きちんとした運用は可能であろうというふうに考えております。

○福井座長 ありがとうございます。先生が最後のところでおっしゃった研究拠点の整備等の実現を図るため、専門家による推進協議会を設置することを提言されるということですので、そのことを念頭に置いて、この戦略協議会でも議論を進めていきたいと思っております。

どうも先生ありがとうございました。

それでは、続きまして、竹内委員より提出いただきました意見書についてご説明いただきたいと思っております。

○竹内委員 時間がそれほどありませんようですので、手短かに説明させていただきます。

資料2-4をごらんいただければと思います。

医療・健康等に関する情報の利活用についての意見ということでまとめさせていただいております。

疾病構造の変化や少子化、高齢化等に伴いまして、質の高い医療サービスを提供していくためには、エビデンスに基づく医療の標準化を進めることは皆さんが述べられておりますように非常に重要なことございまして、ICT技術を活用した医学研究によりこれらが可能になると考えております。

ご存じのように、欧州等におきましても、国の政策として医療情報を蓄積して利活用することが既に行われております。我が国におきましても医療情報の利活用によりイノベーションを実現することが喫緊の課題であると考えており、明確なグランドデザインのもとに実際に利活用できる医療情報データベースを構築することが求められているのではないかと思います。しかし、現有の医療情報データベース等は、先ほどの説明にもありましたように、標準化されていないことから互換性がなく、使い勝手がよくないように感じております。したがって、今後構築されるデータベースに関しましては、標準化を進め互換性のあるものとして確立すべきではないかと考えております。

また、今後このような仕組みを構築するに当たりまして、利活用する側の視点も重要ではないかと考えております。実際に用いられて期待される成果を上げるためにも、企画設計段階から産業界、私たちは利活用する側ですけれども、産業関係者も参画して検討することが必要で

はないかと考えております。

このような認識のもとで医療情報の利活用に関しまして以下の4点を提案させていただきます。

1つ目は、医薬品の安全対策や研究開発に向けた医療情報データベースの利活用の推進というところでございます。

先ほども申しましたように、英国を初めとして海外では同様の施策が既に開始されていることは皆さん十分ご存じかと思っておりますけれども、我が国におきましては、電子化された大規模医療情報データベースを産業界等で利活用する環境には現時点ではありません。我が国の強みであるICTを利用して医薬品開発を推進するためには、医療情報データベースの産業界での利活用は必須と考えております。

産業界が利活用する目的には、安全対策と研究開発、その2つがでございます。

医薬品の安全性、有効性の確保には、現場の医療関係者のみならず企業も責任を有しております。このためにも企業も利活用できる仕組みづくりが必要ではないかと考えております。

また、疫学情報、臨床情報、そしてオミックス情報等を基礎研究とリンクさせることによりまして、疾患メカニズムの解明や新規創薬ターゲットの発見等に結びつくと考えており、期待もしております。

さらに、先ほども申しましたように、利活用を推進するためには、企画・設計段階から利活用者としての産業界の参画が望ましいと考えておりますので、ご検討をよろしくお願いいたします。

次のページになります。2つ目は医療情報DBリンケージの推進についてです。

医療情報データベース間のリンケージが可能になりましたら、医薬品と有害事象の因果関係が解明されやすくなりまして、結果的には、我が国の医薬品の安全監視の飛躍的改善をもたらすことが期待されます。

厚労省で推進されておりますナショナルレセプトデータベースとセンチネルデータベース、あるいはナショナルレセプトデータベースとがん登録のリンケージがその一例ではないかと感じております。

以上が2つ目です。

3つ目は、電子カルテの普及促進です。

安全かつ効率的な情報連携を可能にするためには、医療等ID（仮称）と、それに基づく認証・認可機能、情報提供ネットワークシステムとの接続機能を中心とする医療等情報中継デー

データベース（仮称）の構築と電子カルテの普及が重要であり、電子カルテを普及させるためには国家的な施策が必要であると考えております。

参考までに意見書に電子カルテの普及率を付記させていただいております。

4つ目は人材育成です。

これは、先ほどお2人の先生もおっしゃっていましたが、現在、我が国におきましては、医薬品と有害事象の因果関係や有用性を研究する薬剤疫学専門家がほとんど存在しない状況であり、医療情報を利活用するためには、各種医療情報を利活用しやすいデータとして変換する薬剤疫学専門家を早急に育成することが望まれております。

以上4つの点を提案させていただきます。どうぞよろしくお願いいたします。

○福井座長 ありがとうございます。

ご質問やご意見などございませんでしょうか。どうぞ、狩野委員。

○狩野委員 とりわけ4点目に関して、まことにここにおっしゃっているとおりであります。

ただ、このような人材をアカデミアで配置したいと思った場合の現在の問題点を考えますと、1つ目は、研究評価の指標がインパクトファクターに寄り過ぎているために、このような研究の結果というのは、なかなかインパクトファクターが高いジャーナルというのは存在しておりませんで、その研究結果が認知されにくい、あるいは研究費を取ってくるための基準として評価されにくいという点があるように思っています。この評価方法が国からの財源としての研究費に適応されている限りは、何か別の指標あるいは価値観を持った方々によるファンディングがない限りは、こういう方々がなかなか育ってまいらないということを心配しております。もしそれがご同意いただけますと、例えば、そのような独自の価値観のもとに企業連合体などの形でファンディングソースをつくっていただくことで、こうした育成すべき新分野に対する援助を進めていただくようなことがもし可能であれば、こうしたなかなか大多数の理解をまだ得られにくいけれども、実は重要であるという新たな研究分野を支えていただけるのではないかと、期待をしているところであります。これが1点目。

それから2点目ですけれども、先ほど高木先生にも伺ったように、人材の卒後の活躍先が見えにくいと、なかなか若い人が入りにくいという、これも世情であります。だからどうしてもいいかがなかなか難しいところではありますけれども、例えば、この種の仕事をアカデミアに置いた場合に、学生さんがそこに行きたがるかどうかというのは、その先にどんな道がありますかと必ずみんなに聞かれてしまいます。とりわけ就職先というのはアカデミアにはなかなかたくさんなく、それに比較すると企業により多くあるという状態が現状では続いております。

ですので、こういう方々がもし育ってきた場合に、企業の中でどういう使い方をしていただけるのかという点も、もし何かアイデアがございましたら非常にありがたいというところであります。よろしくお願いいたします。

○竹内委員 非常に重い課題であると思います。即答はしかねますけれども、引き続き産業界、業界団体を含めて検討させていただきます。

○福井座長 ほかにはいかがでしょうか。よろしいでしょうか。

ちなみに私も、本席先生が医学部長だったときの京大医学部教授会で、随分その話をしたことがあります。例えば、バイオインフォマティクを養成するにしても働き場所があるのか、よくわかりませんでした。大学でのニーズと国全体としてのニーズが、微妙にずれてきているようです。大学改革という名前で片づけられる問題かどうかわかりませんが、大きな構造の変革が必要だとは思っています。

ほかによろしいでしょうか。どうぞ、桃井委員。

○桃井委員 大変すばらしいご提言だと思います。3番目の電子カルテですが、今実際に臨床家としても見ていて、問題なのは、普及も大切なんですけれども、あそこにある膨大なすばらしい情報が活用できないような普及のされ方をしているというところが一番大きな問題であると思います。ですから、これは個々の場所ではとても解決できない問題ですので、我々も医療経済的にも電子カルテのために働いているような気がするほど、ほんの小さなバージョンアップに多大な、これは言われたとおりのたん入れたらおしまいという状況でございますので、そういう冗談が飛び交うほど、もちろんメリットも多大でございますが、マイナスも多大である。一番マイナスなのは、あの膨大な登録さえも全国的な活用ができない状況にある。これをどこかでやはり国の施策として、普及ではなくて、その膨大なデータを活用できるシステムをぜひ検討する場所をつくっていただきたいと思います。

○福井座長 ありがとうございます。時間も押し迫ってきましたので、それでは竹内先生、ありがとうございます。

本日最後の議題は、次回の第6回以降のテーマについてです。中野審議官からご説明をお願いいたします。

○中野審議官 「第103回総合科学技術会議における総理指示への対応について」という紙が配付されております。これに基づいて説明させていただきますが、まず、1の背景にありますのは、イノベーションを実現するために必要なシステム改革（規制・制度改革、導入促進策）について国として推進すべき戦略、取組内容案を具体化することは、それぞれの戦略協議会の

役割の一つであるという、この協議会設置の専門調査会における決定でございます。この決定、おさらいのために2枚目の別添1につけておりまして、この協議会が専門調査会によって設置されておりまして、2の役割及び期待される成果の後段、期待される成果というところにありますのは、ここまでご検討いただいたアクションプランの取りまとめということと、2番目にあります、イノベーションを実現するためにシステム改革を取りまとめていただく。これを専門調査会に提案いただくということだったわけでございます。

1枚目に戻りまして、背景の2番目の丸にございますとおり、7月末に開催されました総合科学技術会議の本会議、議長は総理大臣でございますけれども、ここの場で最後に総理大臣から、システム改革等イノベーション実現に必要な施策のあり方について年末までに対応方針を取りまとめるようにという指示があったところでございます。

今後のシステム改革と言われているものの検討の体制が2にございますけれども、イノベーション政策を実効的に運営するための総合的な役割及び運営の改革ということについては、イノベーション専門調査会自体が検討いたします。各戦略協議会におきましては、アクションプランに掲げる政策課題や重点的取組の目的を達成し、社会的課題の解決につなげるための取組改革を整理いただく。あと基礎研究・人材部会においては、基礎研究・人材に関する改革の検討をいただくということになっております。

ご承知のとおり、他の会議、機関でシステム改革あるいは制度改革などについては、過去にも現在も多くの提言が行われておりまして、時間的制約もございますので、戦略協議会でご提言いただくものについては、各省がこれまで行っているものとの重複を排しながら、特に医療イノベーション推進会議のほうで体系的にまとめていただいているものもありますので、これまで出ているものとの重複をなるべく避けながら、総合科学技術会議ならではの、ちょっとこの先、事務局から申し上げるのは僭越なんですけれども、言い方を気をつけなければいけません。官邸あるいは総理に求められているものというのは、総合科学技術会議ならではの、総理大臣、関係大臣が取り組むべき課題、大きな課題を、限られた数整理していただきたい、ご提言いただきたいということでございますので、この後の検討の過程では幅広くご議論いただくことになると思いますが、こうしたことを念頭に置いていただければ幸いです。

以上です。

○福井座長 ありがとうございます。

何かご質問ございませんでしょうか。よろしいでしょうか。

それでは、前回の第4回の会議のときまでにいただきました委員からの意見書のうち、4名

の方についての意見書はまだ議論されておられませんので、桃井委員、吉岡委員、原澤委員、成戸委員からの意見書につきましては、次回の会議で検討していただきたいと思います。その残りの時間と、その後2回の会議でシステム改革について議論をしたいと考えております。そして12月20日の専門調査会で議論の集約化ができるようなスケジュールで進めていきたいと思っておりますので、どうぞよろしくお願ひいたします。限られた時間でかなり濃密な議論になると思っておりますので、効率的に議論を進めるため、できましたら第6回となります次回の開催までに事務局と相談しまして、会議の運営方法や検討の枠組みがわかりやすくなりますように委員の先生方にお諮りしたいと考えております。

それでは、事務局から説明をお願いいたします。

○北窓参事官 ただいま審議官からご説明させていただきましたように、限られた時間で濃密なご議論をいただくこととなりますので、まず1つ目の提案でございますが、大変お忙しい先生方にお集まりいただいているので恐縮ですが、資料3に年間スケジュール予定を示してございますが、次回の第6回、9月26日でございますが、皆様のご理解が得られれば、その開催の時間を、今まで2時間の会議でやってまいりましたが、例えば2時間半とか若干長目に開催させていただけないかということをお諮りしたいと思っております。

それから、システム改革についてでございますが、先生方からいただいたご議論とあわせて、現在、CSTPでは4府省が一体となって進めてきた先端医療開発特区、いわゆるスーパー特区について進めてきているわけですが、これが平成20年度に創設され、今年度は最終年度となります。スーパー特区、それ自体がシステム改革の試みの一環でもございますので、その成果と特区制度の貢献、及びその課題について事務局として分析した上で、今後の展開についてご議論いただくことも考えてございます。

以上について事務局から発言させていただきました。

○福井座長 ありがとうございます。今後の進め方について、何か委員の先生方からご意見ございますでしょうか。はい、どうぞ。

○庄田副座長 今、参事官がご説明されたスーパー特区についてですが、スーパー特区では規制の緩和措置や各種支援措置がとられて、特区でなければできなかったことができたはずで、スーパー特区でなければできなかったことが、どのような措置がとられて実現することができたんだと、そういうことがしっかりわかるような資料、ご説明をお願いしたいというふうに思っています。

○北窓参事官 はい。

○福井座長 ありがとうございます。

それでは、閉会につきまして事務局のほうから何か説明はございますか。

○北窓参事官 それでは、次回協議会は9月26日、13時から、申し訳ございませんが、2時間半にわたって開催させていただきたいと思います。

本日はどうもありがとうございました。

○福井座長 これで終わります。ありがとうございます。

午後 2時02分 閉会