

- 【日時】 平成21年1月30日（金）10:00～11:35
【場所】 ライフサイエンス統合データベースセンター大会議室
【出席者】 森下真一(作業部会委員)、中谷洋一郎(東大)、池村淑道(長浜バイオ大)、瀬々潤(お茶の水女子大)、高木利久(作業部会主査)(文科省ご担当、急病のため欠席)
永井啓一、西川哲夫、川本祥子、坊農秀雅、箕輪真理(以上、DBCLS)

(敬称略・順不同)

【議事】

平成21年度業務計画・平成20年度プロジェクト進捗状況について

議事に入る前に文科省報告書(追加資料)について高木主査から説明があった。「ライフサイエンス委員会の下の情報基盤整備作業部会からプロジェクト終了後の体制について報告があり、内容はJSTのBIRDとの一体運営によって、このプロジェクトを永続化に向けて体制を整える、というもの。ROISの中のDBCLSは残り、JSTの中に新たにできる組織と連携して運用していくものになると思うが、具体策については今後検討される。」

➤ 予算配分額について

中間評価での踏み込んだ評価内容として次年度予算案(各組織に一部増減がある内容。資料1-2)が提出され、これを反映して提案(資料1-1)が作成されている。人材教育については変更なし。PJ全体の総額は今年度の額と変更がない予定だが、2年後以降の一体運営のための移行措置として2.5億円分がJSTから予算執行がなされることになり、DBCLS分のうち2.2億、遺伝研(補完課題)分についてはBIRD課題の分に3000万円がそれに当たる。すでに業務計画書を各機関から提出いただいております、資料1-3のスケジュールに従って、予算の配分を調整し関係機関の承認を経て、最終的に決定する予定。

資料1-4は中間評価の指摘内容とそれに対する対応、また資料1-5は今年度の目標と昨年12月時点での進捗をまとめたもの。これらを参照しながら、来年度の目標を交えてご議論いただきたい。

➤ 東京大学

◇資料説明◇

演習の目的はバイオDBの構築できる人材育成で、1年目はスキル取得、2年目は独創的サーバ構築を行う。19~22年次における計画のうち、プログラミング演習に関しては、理学部の演習(今年度まで)をそれに充てている。21年度からはプログラミング演習が無くなってしまいますので、経験のある人を募る。WSが15台のみなので、各段階の演習受講者の合計は最大15名。大学院生と社会人では継続への意欲が異なる(大学院生<社会人)ようなので、次年度は経験者として社会人を募りたい。

演習の具体的な目標(10ステップ)、実施内容について説明。20年度受講者についてはすでに2名がステップ8(Ensemblのミラー)まで完了、6名はEnsemblの示す手順でやっているが同ステップでエラーが出ており、解決作業中。受講生の傾向として、指示された手順でやることは可能だが、問題解決はなかなか難しいようだ。2年目の受講者が行う独創的サーバ構築の内容説明。

◆質疑応答◆

○外部受講者が少ない状況については、宣伝不足なのか？外部の人が来にくい状況か(曜日等)？

→木曜日の午前10-12時。授業との重なりもあるかもしれない。

○企業の方が研修に来る、といったケースは？

→企業からの問い合わせはなかった。

○プロジェクトでも2回くらい宣伝し、特定研究のメンバーにも通知を出したが…。

→Solexaをお持ちの先生からは1件問い合わせがあった。

○来年度に関してはまた3月位にアナウンスしよう。センターHPなども活用して。

○教材は蓄積されていく予定か？それらを使った自習ができるか？

→Webに公開している。つまりいたところの情報も具体的に出している。

○受講生のバックグラウンド(BG)はどういうものか？

→本来の勉強もしながらこちらもやってみようという大学院生が主。サーバ管理経験者は一人くらいであまり多くない。

○来年度プログラミング経験者募集ということだが、その要求レベルは？

→生物のBGの人はできないこともあるが、自分でミニプログラムを書いたことがあるくらいでできる。

○そんなに高度でなくても、1-2言語で書いたことがあれば大丈夫なのか？

→そうだと思う。出した課題をこなせれば。

▶ 長浜バイオ大学

◇資料説明◇(追加資料あり)

学部学生が育成の主対象である。DBCLSや参画機関等の支援要員の育成、研究へのDB活用やDB作成に参画できる人材の育成、DBの有用性が理解できる人材の育成が目的である。特に、アノテーションができないデータが急増する今後の状況に対応して、大量データに対応できる人材を育成する。

2年生向け実習課題つきテキスト(全39p)には、DBCLSやJSTで作っているDB等の紹介、生物資源に関するDB紹介、Jabionの紹介、などを盛り込んでいる。テキストどおりに何かするのではなく、独自の課題設定を課しているので、学生ごとに異なった独自性のあるレポートの提出を必須とした。3年生(250名)向けには、追加資料の前期用のテキスト(54p)に沿っているが、今年度は「健康への貢献遺伝子DB」の構築を実施したところ、各学生が興味を持つ様々なテーマが提案された(追加資料)。環境由来のメタゲノムデータを対象にして、約3500の「健康へ貢献する可能性のある」新規遺伝子が発見されたが、アノテーション内容についてはシニア世代に精査してもらう予定である(昨年と同様)。結果全体をみると、服部グループ(東大)のヒト腸内由来のメタゲノムデータからは、多くの有用遺伝子の候補が出てきた。ベンターのサルガッソ海由来データからは登録数の割には少ないのは、すでにスクリーニングされて有用なものは出されていないのか？という印象を持った。後期はゲノムアノテーション(テキスト9p)に従い、本年度は各学生別でアノテーションを実施した。4年生についてはシニア世代と一緒に世界に通用するDBを作るという課題を実施し、tRNADBを公開した。学生は複数のプログラムを用いてtRNA遺伝子を推定し、それをシニア世代がキュレートした。キュレートの際にシニア世代から出た要望をtRNADBの多様な検索機能として盛り込んでいる。自分で統合TVを作りたい、という学生もいたので、統合TVへ登録した。

▶ ◆質疑応答◆

○21年度の計画については？

→今年度実施内容の延長に加え、エキスパートについてはタンパク質遺伝子分野のシニア世代専門家を入れたい。既存のtRNADBについては毎年の更新が必要であり、高等動植物ゲノムに関しては、公的機関からの作成依頼があれば「エキスパートが精査したtRNA遺伝子DB」に順次加えたい。イネゲノムに関しては既に開始している。

○テキストは一応完成？来年度新たなものが出てくるのか？

→統合TV的な可視化を進めたものはさらに拡充したい。21年度の4月からは、コンピュータバイオサイエンス学科を新設するので、その一回生用のテキストがひとつは、今年度内後期用に必要になる。ゲノムアノ

テーションは自ら考えてもらうためにシンプルなテキストにしているので、拡充すべきかどうかの見極めが必要。次世代シーケンサーの大量情報用に作成を予定している。

○講義を受けた学生が世の中に出ている？

→東大の森下先生のところ、ならびに奈良先端大にひとりずつが、DB 作成に関係する分野で大学院へ進学している。もうひとり民間企業へ就職し、教育研究機関からの依頼で、文部科学省プロジェクトとして、ライフサイエンス分野 DB の作成を担当している。ライフサイエンス分野 DB の作成に直接に関与しているのはこれら 3 名である。他には、コンピュータグラフィクスや DB に関係した IT 分野への就職者がいる。

○学生の興味を引くうまいテーマとしてはどんなものがあるのか？

→「健康」には興味あるようだ。「環境」については、バイオエタノール生産、リグニン分解といったものが出る。KEGG 中の薬剤に関与する代謝経路などについては、エキスパートの指導のもとにメタゲノムから遺伝子を発掘する作業を、学生と研究員が共同して実施している。

▶ お茶の水女子大学

◇資料説明◇

細かい資料はすべてウェブに掲載済みのため、本日の資料は概要を提示。大学院生向け全学の授業を実施。BG としては情報系：生物系：化学系＝1：1：1 位で延べ 45 人。講義内容は、基礎講義(20 年度新規、BG が異なるため)、ネットワークを通じたライフサイエンスデータの利用(19 年度は NCBI の API、20 年度は TogoWebService を利用、マニュアルのフィードバックも行った)、データ解析の基礎(可視化部分はこちらから提供)、データマイニング技術(20 年度新規、基本プログラムは提供)。資料公開：後半はすでに公開、基礎的講義については著作権の関係で現在図を書きなおして公開準備中。啓蒙活動としてのセミナー開催。21 年度計画については、外部受講生確保(20 年度は当初予定していた 2 名の外部受講生が結局駄目になって 0 名。21 年度については現在 8 名[4 名はテクニシャンとして勤務中、4 名は主婦]予定)、カリキュラムの工夫(主婦の不安、学生と一緒にやるのは落ちこぼれそうで怖いという気持ちに配慮して、学生とは別に TA を多めにつけながら実施する必要あり。社会人の本務との調整への対応。部分完結できるカリキュラムと習熟度判定の組み合わせ)、他を実施予定。

◆質疑応答◆

○テキストのほうは著作権処理をして公開予定？ダウンロード可にする？

→PPT 資料なので話した内容のテキストの追加などするかもしれないが、公開する。

○テキストを横断検索にかけるのは有用ではないか？お茶大に限らず。

○センターとしてはぜひやりたい。

○郷先生がおやめになるについては、PJ への大学のサポートは変わるか？

→サポートはお願いしているが、具体的には不明。文系の方が理解はある。

○教材と人材がアウトプットになると思うが、長浜バイオ大のように DB を作ったら明確になるのでは。

→人材については統合 TV に 2 名投入したのもアウトプットであると思っている。

○文科省の欲しているものはどういったものか？

○(DBCLS)不足している人材のスペックと人数を明確にし、それを供給するための計画を求めている。

○(資料 1-4 を見ながら)普通のバイオインフォの講義とどこが違うのか？ということも言われていたが、東大のほうの方がわかりやすのかもしれないが、お茶大も DB の統合や構築に特化するという点を踏まえて来年度の提案をいただいていると思うがいかがか？

→データを統合するという課題についてはわかりにくいという点もあると思う。

○受講生の卒業後の進路は？

→昨年の例では、製薬の中の IT 系 SE や総研系 SE とか。ただ、バイオではない分野を担当するケースも。

○製薬自体の全体枠が少ないのではないか？

○バイオ系であっても SE になる人が多いのか？

→そのほかには化粧品メーカー。ただ、やはり枠が少ないので、SE 職の人が多くなる。

➤ まとめ（総合討論） 他

○教材・人材育成について、長浜は学生数も多くて問題ないが、東大やお茶大についてはまず来てくれる人がいないことには始まらない。その点で教材に重点を置いてということになってしまうのか？人材の集まりは伸びないのはインセンティブがないから、ということだろうか？バイオに進む人はあまりいないのか？

→(東大)人によってスキルや進捗のスピードがまちまち。IT系の企業への就職に役立つというインセンティブにはなるらしいが、バイオ関連のインフォの職種があまりない。

→(東大)作成したテキストはノウハウの塊、3年生あたりの講義にも使えるのではないかと考えている。外部からこれを受けに来る人はそれなりのスキルが必要ということもハードルか。一方で、これ以上スキルを落とすとDB構築自体ができないということも心配される。

○社会人や主婦のリクルートの方法は？

→(お茶大)主に公開講座に来た人。OGや研究室のつてを使って。

○主婦の方はやはり社会復帰の意欲を持っている方か？具体的にアノテーターになりたいという希望は？

→(お茶大)子供の手が離れたとか、最近の情報を知りたくて(興味)といったきっかけ。授業の中ではアノテーターの不足について訴えているが、具体的に希望があるわけではない。

→(長浜バイオ大)tRNAのキュレーターについては出来高で謝金を支払っている。責任感と同時に継続する上での励みになっているようである。学会にも出席し、ポスター発表も担当している。シニアをもっと発掘し活用すれば、DB作成やその高度化に関して、有意義で特徴のある活動ができると思えるので、ぜひこの活動を広げたい。コンピュータへの苦手意識は、学部学生と共同作業をすることで解消する傾向にある。生物の知識を学部学生へと継承する実践の場ともなっており、そのこともシニア世代の積極的な活動を生んでいると思える。

・シンポジウム案説明 統合DBの成果と合わせて、PJ外のDBに関する活動も紹介できるようなものにした。口頭発表やポスター発表などにご協力願いたい。

(11 : 35 終了)