

平成 21 年 10 月 16 日

統合検索技術の開発に伴う予算増額に関する審議のお願い

1. 審議事項

・ライフサイエンス統合データベースセンター（以下、DBCLS）で開催を予定している第 3 回 BioHackathon（平成 22 年 1 月予定）の場を活用し、統合データベースプロジェクトの課題である統合検索技術開発のための課題を検討し、必要な開発を実行いただけるような国内外の専門家を招待することにより、セマンティックウェブ技術に関する取り組みを行いたい。

なお、年度開始までに、高度なデータベース統合化のためのセマンティックウェブ技術開発に関する業務実施内容を具体化するに至らなかったため、委託契約締結時には、本取り組みに関する予算見積りを計上しなかった。

・予算増額分は、本実施項目に関わる外国人招へい旅費、国内旅費、会議開催費等、総額 5,072,983 円となる。

・本承認後は、すみやかに文部科学省に委託業務変更承認申請の出願を行うこととする。

2. 概要

DBCLS では、これまで統合データベースプロジェクトの中核機関として、生命科学系データベースカタログの作成、生命科学系横断検索システムの構築、生命科学系データベースの受け入れなどを当初設定した年次計画に従って進めてきた。

年次計画で残された最後の課題は、統合されたデータから知識発見につながるような検索を実現するための統合検索技術の開発である。

このような先進技術を取り込んだ検索を実現するに当たっては、国内のリソースだけを統合しても必ずしも十分ではなく、すでに広く使われている海外のリソースや技術的な規格を参考にしつつ、国際的にも相互運用性の高い統合のあり方を検討する必要がある。

そこで、本プロジェクトにおいても、これまで DBCLS が主催した BioHackathon の場などを利用して、ウェブサービスにおけるデータとプロトコルの標準化、およびそれを用いた統合的なデータベース利用環境構築のための基盤となるソフトウェアの開発を実現してきた。

過去 2 回開催した BioHackathon の議論から、データ量が爆発的に増大している現在の生命科学においては、単なる大規模データの統合と横断的な検索を超えて統合されたデータから意味のある推論を行うための、セマンティックウェブによる知識化が有望、ということが明らかになってきた。

そのため、本プロジェクトの残された重要課題である生命科学系データベース統合検索技術を開発するに当たっては、国内外の分散するデータベースリソースをセマンティックウェブ技術の枠組で扱うための基盤となるソフトウェアの開発に着手し、そのフィージビリティの評価を行うことが急務であることが分かった。

3. 実施内容（第3回 BioHackathon にて行う業務実施内容）

<p>(1) 統合検索技術開発の課題検討</p> <ul style="list-style-type: none"> ・技術課題の洗い出し <ul style="list-style-type: none"> ①国内外のデータベースリソースをセマンティックウェブの枠組で扱うための基盤技術整備 ②統合と推論を促進するためのオントロジーの整備 ③コンテキスト依存の可視化およびユーザーインターフェースの整備 <p>(2) セマンティックウェブ技術による統合検索技術開発</p> <ul style="list-style-type: none"> ・基本仕様、方式検討 ・基本部分の試作開発 <p>【実施時期・期間】平成22年1月中旬（期間は6日間）</p> <p>【実施場所】インテック（株）大山研修センター（富山県富山市）</p> <p>【業務実施に協力予定者】別紙参照</p> <p>※会議期間中の守秘すべき情報の漏洩防止に関しては、DBCLS 職員が責任をもって対応する。</p>

4. 予算

<ul style="list-style-type: none"> ・外国人招へい旅費（¥5,952,000）：招へい者（20人）の渡航費用 ・国内旅費（¥934,100）：国内移動旅費 ・会議開催費（¥83,500）：実施場所の賃借料（6日間） <p>※本実施に伴い、予算増額を圧縮するため、海外出張に伴う外国旅費の見直し（滞在日数等）や、業務用図書等の消耗品購入の削減を行う。</p>
--

(内訳)

費目	種別及び内訳	当初計画額	増減額	改計画額
業務実施費		170,047,044	4,611,802	174,658,846
	消耗品費	14,156,237	△ 2,290,238	11,865,999
	国内旅費	5,280,140	934,100	6,214,240
	外国旅費	3,181,600	△ 272,400	2,909,200
	外国人招へい旅費	0	5,952,000	5,952,000
	会議開催費	110,200	83,500	193,700
	消費税相当額	7,933,867	204,840	8,138,707
一般管理費		38,815,090	461,181	39,276,271
合計		426,966,000	5,072,983	432,038,983

5. 変更が業務計画に及ぼす影響及び効果

- ・業務実施費の増額により、本業務の重要課題であるライフサイエンスデータベース統合検索の高度化を目指した、セマンティックウェブ技術の調査および開発を実施することで、費用対生産性・品質の高い成果の実現が可能となる。
- ・この成果により、たとえば、本委託業務において提供している TogoWS, TogoDB をはじめとしたサービスを、セマンティックウェブを通じて他のサービスと連携させつつ、よりユーザがストレスなしで、高度な利用ができるようになる。
- ・上記業務の追加に伴い外国旅費の回数及び消耗品費の単価等を見直し、経費を減額することとしたが、業務計画そのものに影響を及ぼすことはない。

6. BioHackathon について

Hackathon と呼ばれる会議形式は、開発者が一定期間一堂に会して開発を行うものとして、近年 IT 分野でその有用性が認識されつつある。

ライフサイエンスにおいても、過去の BioHackathon が示してきたように、合宿形式で共通の課題について議論し、同時に集中的なソフトウェア開発を行う機会を提供することで、時差によるギャップを越えた効率的な共同作業が可能となり、プロジェクトの垣根を取り払って様々な新しい成果が生み出されてきた。

このような国際開発会議は非常に高い創造性や生産性を生み出す会議のスタイルであり、今後新しい国際会議のパラダイムとしてもっと普及していくものと考ええる。

また、日本がイニシアチブをとってこのような場を提供していくことは、日本側の希望に沿った開発成果を上げることができるだけでなく、世界の中での、この分野の日本の存在感を高める上でも重要なことと考える。

(過去の実績)

回	実施期間	主な成果
第1回	平成20年2月11日 ～15日	TogoWS、TogoDBの基本部分の仕様決定と試作
第2回	平成21年3月17日 ～20日	統合的な利用環境を構築するために必要な技術開発 (DBCLS Galaxy)

(参考) 業務実施協力予定者一覧

氏名	所属	
	所属機関 部門	役職
Arek Kasprzyk	Ontario Institute for Cancer Research (カナダ)	Director of Bioinformatics Operation and Principal Investigator
Greg von Kuster	Penn State University (アメリカ)	Galaxy Project Manager
Raoul Bonnal	Fondazione Istituto Nazionale Genetica Molecolare (イタリア)	Research Scientist
James Taylor	Emory University (アメリカ)	Assistant Professor
Jan Aerts	Wellcome Trust Sanger Institute (イギリス)	Senior Bioinformatician
Jose M. Fernandez	Centro Nacional de Investigaciones Oncologicas (スペイン)	Bioinformatician
大野圭一郎	University of California, San Diego (アメリカ)	Research Associate / Programmer Analyst
Mark Wilkinson	University of British Columbia (カナダ)	Assistant Professor
Martin Senger	International Rice Research Institute (フィリピン)	Researcher, Bioinformatician
Oswaldo Trelles	Universidad de Malaga (スペイン)	Professor
Pierre Lindenbaum	The Center for the Study of Human Polymorphisms (フランス)	Bioinformatician
Rutger Vos	University of British Columbia (カナダ)	Research Assistant
Young Kim	Korea Research Institute of Bioscience and Biotechnology (韓国)	Professor, Principal Resercher
Alberto Labarga	Univeristy of Granada (スペイン)	Bioinformatics Consultant
Hilmar Lapp	National Evolutionary Synthesis Center (アメリカ)	Assistant Director
Bruno Aranda	European Bioinformatics Institute (イギリス)	Assistant Director
Christian Zmasek	Burham Institute for Medical Research (アメリカ)	Post Doctoral Associate
Pjotr Prins	Wageningen University (オランダ)	Researcher
Alex Gutteridge	Cambridge University (イギリス)	Post-doctoral Researcher
Alan Ruttenberg	Science Commons (アメリカ)	Principle Scientist
片山 俊明	東京大学 医科学研究所ヒトゲノム解析センター	助教
川島 秀一	東京大学 医科学研究所ヒトゲノム解析センター	助教
豊田 哲郎	(独)理化学研究所横浜研究所生命情報基盤研究部門	部門長
中村 保一	国立遺伝学研究所生命情報・DDBJ 研究センター	教授
野口 保	(独)産業技術総合研究所生命情報工学研究センター	研究副センター長

瀬々 潤	お茶の水女子大学大学院人間文化創成科学研究科	准教授
荒川 和晴	慶應義塾大学大学院政策・メディア研究科	特別研究講師
荻島 創一	東京医科歯科大学難治疾患研究所	助教
川路 英哉	(独)理化学研究所オミックス基盤研究領域	研究員
川島 武士	(独) 沖縄科学技術研究基盤整備機構	研究員
木下 聖子	創価大学工学部生命情報工学専攻	准教授
後藤 直久	大阪大学微生物病研究所	助教
西村 邦裕	東京大学情報理工学系研究科知能機械情報学専攻	助教
藤澤 貴智	(財)かずさDNA研究所植物ゲノム研究部植物ゲノム情報研究室	プロジェクト研究員
山崎 千里	(独)産業技術総合研究所バイオメディシナル情報研究センター	テクニカルスタッフ