



国際
オープンアクセスウィーク

2010.10.18-24

openaccessweek.org

第 6 回 SPARC Japan セミナー2010

日本発オープンアクセス

新たな日本語 Web コンテンツ 「ライフサイエンス 新着論文レビュー」

飯田 啓介

(ライフサイエンス統合データベースセンター 特任技術専門員、
ライフサイエンス 新着論文レビュー 編集人)

講演要旨

「ライフサイエンス 新着論文レビュー」は、ライフサイエンス統合データベースセンターが発信する日本語コンテンツのひとつとして、2010年9月1日にスタートしたばかりのサービスです。Nature、Science、Cellなどに代表されるトップジャーナルに掲載された日本人を著者とする生命科学分野の論文について、論文の著者自身の執筆による、専門分野の異なる生命科学研究者にむけた日本語によるレビューを、だれでも自由に閲覧・利用できるよう Web 上にていち早く無料で公開します。論文の出版から1か月以内の公開を目標とし、また、毎週平均2本以上、年間で100本以上の公開を予定していて、この1か月ですでに20本以上が公開されています。最新の研究成果を、日本語で、その背景からわかりやすく紹介・解説すること、そして、それらコンテンツの自由な引用・転載・再利用を可能とすることで、生命科学分野のサイエンスコミュニティ全体に寄与することをめざします。



飯田 啓介

1990年 東京農工大学大学院農学研究科 環境保護学専攻 修了（農学修士）、同年 東京化学同人、1997年 シュプリングー・フェアラー東京（現 シュプリングー・ジャパン）、2004年 共立出版『蛋白質 核酸 酵素』編集部と出版社を渡り歩き、生命科学分野の書籍および雑誌の編集者として活動する。2005年4月『蛋白質 核酸 酵素』編集長、2009年9月『蛋白質 核酸 酵素』の休刊決定に伴い共立出版を退社。2010年4月より現職。

本日はライフサイエンス統合データベースセンターがこの9月にはじめたばかりの「ライフサイエンス 新着論文レビュー」について紹介します。

自己紹介

まず簡単に自己紹介させていただきます。私は1990年に東京農工大学大学院農学研究科修士課程を修了し

ました。そののち、東京化学同人という、大学生を対象とした化学の教科書を主力としている出版社に入社しました。私はそれからずっと生命科学分野において編集の仕事をしてきたのですが、最初のこの会社で編集者として基本的なことをたたき込まれました。

1997年にシュプリングー・フェアラーク東京(現 シュプリングー・ジャパン)に移りました。ここではライフサイエンス分野の日本語書籍の編集をすべて担当することになり、年間10冊程度の書籍の企画・編集・制作を行いました。この会社で編集者として独り立ちして仕事ができるようになったと思っています。また、このころは海外の大手出版社で英文ジャーナルの電子ジャーナル化のはじまった時期で、その中心となったこの会社でその過程を身近に見聞きできたこともとてもよかったと思っています。

2004年に共立出版の『蛋白質 核酸 酵素』という雑誌の編集部に移り、2005年からは編集長を務めました。『蛋白質 核酸 酵素』は生命科学分野ではもっとも古くからある雑誌で、1956年の創刊以来50年以上の歴史があったのですが、残念ながら今年初めに休刊になってしまいました。

私自身は『蛋白質 核酸 酵素』の休刊が決まった時点で共立出版を退社し、2010年4月より現在のライフサイエンス統合データベースセンターに移り特任技術専門員を務めています。

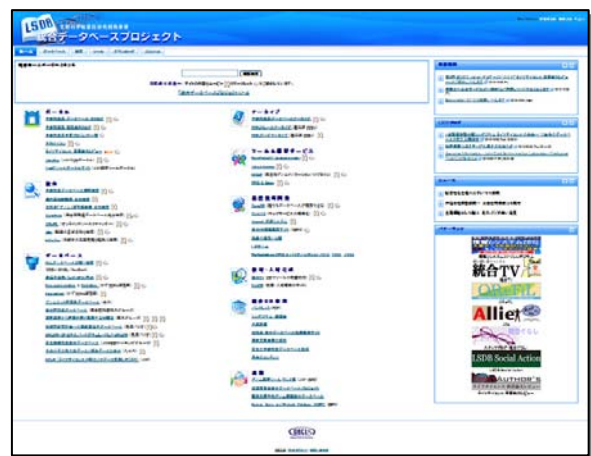
ライフサイエンス統合データベースセンター (DBCLS)

ライフサイエンス統合データベースセンター(DBCLS)は情報・システム研究機構(ROIS)の組織のひとつです(図1)。情報・システム研究機構にはほかに、国立情報学研究所、国立極地研究所、統計数理研究所、国立遺伝学研究所という4つの大きな研究所と、新領域融合研究センターがあり、DBCLSはそのなかでもっとも小さな組織です。

ライフサイエンス分野におけるデータベース統合化の拠点を形成することを目的として2007年4月に設



(図1) ライフサイエンス統合データベースセンター



(図2) 統合データベースプロジェクト

立され、国内のデータベースを中心に、データベースの統合化と保全に努め、利便性を高めるための情報技術の開発やポータルを整備を行うことを目標としています。また、統合データベースプロジェクトの中核機関として機能しています。

この統合データベースプロジェクトは文部科学省の委託研究開発事業で、2006年度から2010年度まで行われます。散在する種々の生命科学系データベースや文献などを、ユーザーにとってわかりやすく、使いやすくしていこうというもので、さまざまなサービスを無料で提供しています。アドレスは<http://lifesciencedb.jp/>です(図2)。

オープンアクセスということでは、DBCLSは文献ではなく、データベースあるいは実験データのオープンアクセス化を行っている組織だといえます。たとえ

ば、ライフサイエンス分野では大規模なプロジェクトがたくさんあり、新型シーケンサーでゲノムを読むと膨大な量のデータが発生します。それを広くみんなが使えるようにしようとか、プロジェクトが終わって管理する人がいなくなり散逸してしまうデータベースをひき受けてあとの面倒をみようとか、散在するデータベースをうまく整理しポータルサイトをつくりそれをつないで使いやすいようにしていこうとかいうことをしています。

新たな日本語 Web コンテンツ

そのなかで、新しい日本語 Web コンテンツとして「ライフサイエンス 新着論文レビュー」というサービスをはじめました。「First Author's」という愛称でよんでいます。http://first.lifesciencedb.jp/がアドレスです(図3)。

私が編集人という肩書きで、事務的なことから編集・制作・公開までひととおりすべてを行っています。ただし、私は Web サイトの構築や管理はまったくできませんので、その部分は DBCLS の中尾光輝特任研究員にお願いしています。トップページのタイトルは DBCLS の岡本忍特任准教授がデザインしてくれました。サイトは WordPress でつくっていて、ブログ的な感覚のページにしています。

「新着論文レビュー」とは

「新着論文レビュー」というのは、生命科学分野の「Nature」「Science」「Cell」などトップジャーナルに掲載された日本人を第一著者とする論文について、論文の著者自身の手により日本語のレビュー(解説記事)をご執筆いただき、それを Web 上でいち早く無料で公開するというサービスです。

最新の研究成果を、日本語で、その背景からわかりやすく紹介・解説することを目標としており、それらコンテンツの自由な引用・転載・再利用を可能としています。そうすることで生命科学分野のサイエンスコミュニティ全体に寄与することができればと考えてい



(図3) ライフサイエンス 新着論文レビュー

ます。

対象とするジャーナルとして、生命科学分野で総説誌を除いたもののインパクトファクターの上位 20 誌を選び出しています。そこに掲載された原著論文のうち、筆頭著者が日本人(筆頭著者と"contributed equally"の場合も含む)である論文を対象とします。

対象とする読者は、広く生命科学全般にかかわる教員・研究者および大学院生・学生です。新聞記事あるいはプレスリリースのように一般市民を対象としているわけではなく、同じ生命科学分野の研究者で専門分野が異なっている人を強く意識しています。そのため、結果や結論ばかりでなく、前提となる研究の経緯やバックグラウンド、将来の展望などもあわせて示しています。

ライフサイエンス統合データベースセンターより日本語コンテンツのひとつとして公開していて、論文の出版から1カ月以内の公開を目標とし、遅くとも2カ月以内には公開することになっています。いまのところ出版から公開までの平均は40日ですが、あと10日短縮して1カ月以内に公開できるようにしたいと思っています。

毎週平均2本以上、年間100本以上の公開を目標としています。対象となる論文は機械的に発生するのでこれについて執筆を依頼し、ひき受けてもらって原稿がいただけたら公開することになります。当初は執筆を断られてしまうことが多いのではないかと考えてい

たのですが、いまのところ 6 割ぐらひはひき受けていただいています。そのため、現在は年間の公開本数を 100 としていますが、最近では上方修正が必要になっていて、年間で 120 とか 130 くらいの数字になりそうです。

著作権はそれぞれの著者が保持しています。ただし、クレジットの明記を条件として、営利目的での二次利用も含め、転載・改変・再利用すべて自由に行えることにしています。

クリエイティブ・コモンズ (CC) 表示のもとでライセンスしており、とくに図については、再利用を考えて解像度の高いものを JPEG ファイルとしてダウンロードできるようにしています。

分量は、5000 字+図が 2 点+参考文献 10 報を目安としてお願いしています。紙媒体と違って分量に制約はないので、これより長くてもそのまま受領・掲載しますが、このくらいの分量が書く方としては書きやすいかと考えています。これは一般的な和文総説誌として組版すると 3~4 ページに相当します。そこそこの読み応えのある量で、たとえば論文のアブストラクトを翻訳しただけのものより充実したものになっていると思っています。

図については、専門外の読者が容易に概略を把握できるように、原則として生データは載せず、できるだけわかりやすいポンチ絵を添えるようにしています。引用・転載・再利用を自由として公開するため、著作権の関係から原著論文の図そのものを掲載するわけにはいきません。オリジナルの図を作製して掲載するようにしています。

デモ

ここで実際に Web ページをご覧ください。

Google 検索で「新着論文レビュー」を検索すると最上位に表示されますので、ここからトップページに行くことができます。あるいは、「統合データベース」を検索すると統合データベースプロジェクトのホームページが最上位に表示されますので、そこにあるバナー

から「新着論文レビュー」のトップページに行くこともできます。

トップページには、公開されたレビューのタイトルが並んでいます。上部には「新着論文レビューとは」という紹介のページがあります。ページの右手には「このサイトについて」という説明や、「最近の記事」のタイトル、あるいは、「ジャーナル別」というカテゴリー分けもあります。「アーカイブ」をみると、9 月に 23 本、10 月に 16 本で、現在では 39 本が公開されていることがわかります。

原稿を受領しなくても掲載される段階、および、実際にレビューが公開されたときには、そのことを Twitter で情報発信しています。現在はまだまだ少ないのですが、300 人弱の方が登録しています。

それぞれのレビューには、日本語タイトルがあり、執筆者名、執筆者の所属があって、e-mail アドレスへのリンクがあります。対象となる論文のタイトルなどかありここから PubMed のページにもリンクを張っていますので、必要ならここから原著論文をみられるようになっています。

「続きを読む」をクリックすると全体がみられます。500 字程度の要約があり、そのあと「はじめに」から本文となります。

図はクリックすると大きくなり、図説明中の「Download」というボタンをクリックすると、別ウインドウで大きくてきれいな図が表示され、それを保存して自由に使えます。

あとは、たとえば文献や図なども本文からクリックすればジャンプするようになっています。末尾に「文献」があり、これも PubMed にリンクが張ってありますので、そこで要約がすぐにみられ、必要なら原著論文をすぐにたどれるようになっています。

最後に「著者プロフィール」があり、研究室のホームページがあるところにはリンクが張ってあります。その下にはクリエイティブ・コモンズへのリンクがあります。

公開までのプロセス

対象とするジャーナルの出版状況はメールアラートやRSSなどで常時チェックしており、対象となる論文が出版されたら即座にメールで執筆をお願いします。できれば2週間で執筆してくださるようお願いしています。2週間というのは短いのですが、多くの方が期限内に書いてくれます。その代わりに、私のほうでも原稿受領から2週間で制作・公開までをします。そのあいだ、著者による校正が1回、Webページ上での最終確認を1回、行います。

「新着論文レビュー」の売り

ひとつは「早い」こと。論文の出版から1カ月以内の公開を目標としています。2つ目は「多い」こと。毎週平均2本以上、年間で100本以上を公開の予定です。3つ目は「わかりやすい」こと。原稿は一文一文推敲し、徹底的に書き直しています。

誤解を受けやすい点

公開以降、多くの方がかなり好意的に受け止めてくださっていますが、いくつか誤解を受けやすい点もあるようです。

まず、内容はすべて書き下ろしです。原著論文をそのまま翻訳しているのではないかと、あるいは、論文のアブストラクトを翻訳しただけではないかと勘違いされている方もいるようですが、すべて新たに書き下ろしたものです。

また、原著論文の執筆者自身がこのレビューを執筆しています。それ以外の第三者、ライターや私どもが書いたものではなく、著者自身が書いたものです。

そして、掲載論文は機械的に選んでおりとくに選別はしていません。おもしろそうだから、これはいい悪いかといった判断はしません。ただし、執筆を依頼した人すべてがひき受けてくれているわけではないので、条件を満たした論文がすべて取り上げられているわけではありません。

さらに、さきほどもお話ししたとおり、図はすべて

作り直しています。ほかからの引用や転載ではないオリジナルのものを掲載しています。また、受領した図原稿をそのまま掲載するのではなく、専門の業者に依頼して作り直し、その際には生命科学専門の編集者が編集をくわえてよりよいものとしています。

とにかく「わかりやすい」ものを

「わかりやすい」ということに力を入れていることが最大の売りではないかと思っています。わかりやすくするというプロセスでは、手前みそですが“編集”が大事だと思っています。いただいた原稿そのままでは専門外の人が読んでわかりやすいものとはならないのが実際のところで、相当に手をくわえる必要があります。

わかりやすくするため、まず、一文一文を吟味しています。

小学校の国語の時間のようですが、たとえば、係り受け（主語と述語の一致）がきちんとできていない場合がたくさんあります。とくに、文章が長くなると最初と最後が一致しないことが多くあります。

それから、助詞の使い方、「てにをは」を一個一個みて直していきます。

能動態と受動態ということでは、英語では受動態をよく使いますが、それをそのまま日本語にすると表現としてあまりよくない場合があります。能動態と受動態を入れ替えることを行いますが、そうすると必然的に主語も変わってきてしまうので、そういうところも整えます。

さらには、一語一語の語順、主語の位置、ある言葉とそれを形容する言葉はどういう位置関係にあればよいかということの一つ一つ考えます。

また、人によっては句点が多かったり、あるいは、文章が長かったり短かったりということがあります。それを適切な量の句読点に直し、誤解を生まないよう、間違いのないように整えていきます。

言葉が足りず補うべき言葉が必要な場合や、その反対で重複している場合もよくあります。足りないとい

ろは補い、重複しているところは削って、すっきりとした文章にします。

文章ごとのつながりでは、適切な接続詞を挟んでなめらかにつながるようにしたり、足りない場合には少し言葉や文節を足してうまくつながるようにします。このように、国語の時間のように一文一文考えて直すというプロセスを行っています。

きょうはあまりお話ししませんが、当然、サイエンティフィックな点に関しても手を入れていますし、専門用語を統一するというプロセスも行います。

“編集”というプロセス

編集とは、専門の第三者がコンテンツを整え、その正確性を担保する、あるいはわかりやすくするというプロセスであるといえるかと思います。

これまでずっとおもだったメディアであった書籍や雑誌などの紙媒体では、編集者といわれる人が必ず介在し、“編集”というプロセスが機能してきました。それによって生み出されてくるコンテンツのクオリティがある程度まで保証されていたという面があったかと思えます。

ところが、Web 時代になり誰でも自分の書いたものをそのまま公表できるようになりました。そのような世界では“編集”のプロセスはほとんど機能していないのが実際です。このことは、コンテンツの信頼性、あるいは、わかりやすさという点から考えて、非常に大きな問題だと私自身はずっと考えています。

日本語コンテンツの重要性

もうひとつお話ししたいのは、日本語コンテンツの重要性です。日本人にとって、とくに専門領域から少し離れた領域に関して手っ取り早く情報を得ることは、サイエンスの標準語である英語よりも日本語のほうがはるかにたやすいと思います。そうした日本語コンテンツの価値を認めたいうえで、その流通を促進するようなくみがこれからは必要になってくるのではないかと考えています。

コミュニティでの共有と再利用

これまでの紙媒体は、完全な複製のできない行き止まりのメディアだったといえます。もちろんコピーはとれますが画質はどんどん下がっていってしまうし、たとえば、そこから一部を取り出して再利用したりリミックスしたりすることはできません。

それに対して、デジタルデータは劣化することなく、コストもほとんどなしに複製・加工・再利用が可能です。このことは著作権ビジネスにたずさわっている人にとっては頭の痛い問題なのかもしれませんが、サイエンスにおいてはそれは悪いことではないのかもしれない。サイエンスにおいては、コンテンツを囲い込んで保護し独占的な利益を生み出すというよりも、コミュニティ全体に公開・共有して広く利益を生み出すようにすべきだと考えています。

「新着論文レビュー」のサイエンスにおける意義

いまは学問の細分化がどんどん進んでいて、生命科学分野においても、隣の研究室で進められているのがどのような研究なのかわからないことが実際です。そういうなかで、「新着論文レビュー」は研究者のあいだの橋渡しができるのではないかと考えています。また、いまは学問が加速度的に発展して新しい知見をキャッチアップしていくことがどんどんむずかしくなっています。自分の専門分野については最新の論文にも目を通しているのですが、少しでも専門からはずれると追いつけないところができます。そうしたなかで、「新着論文レビュー」は研究者のコミュニティの“科学コミュニケーション”において大きな役割をはたせるのではないかと考えています。

「新着論文レビュー」の統合データベースにおける意義

「新着論文レビュー」が統合データベースセンターの事業として行われていることについても、それなりの意義あるいは理由があります。

まず、「統合」という意味では、“糊”としてデータベースを支え、つなぐ機能があります。いろいろなデータベースがあり、いろいろなデータがあって、たとえばひとつのタンパク質について、その配列の情報や相互作用の情報、遺伝子の情報など、いろいろな情報があるのだけれども、それらをまとめて使用するとき、そのことについての日本語のわかりやすい解説があれば、それらが全部有機的につながって有効に利用できるのではないのでしょうか。ばらばらの細切れの情報ではなく、それをつなぐ日本語の解説があることは非常に意義があるのではないかと考えています。

もうひとつ、「データベース」という観点から考えると、このサービスはまだ始まったばかりなので、みなさん、新しい論文の情報が得られてうれしいと喜んでくださっているようで、それはもちろんねらいとしているところですが、ほかの価値も考えています。このサービスが1年、2年と続いてレビュー数が100、200、300と積み重なっていくと、それだけでひとつのアーカイブ、データベースとしての意義がでてきて、そこからいろいろなものがみつけられたりするのではないのでしょうか。たとえば、データマイニング的な手法から新たな情報を得る、日本語コーパスとして研究材料に使う、といった役割も担えるのではないかと考えています。

今後の展開

このサービスははじめたばかりなので、タグ付けができていなかったり、検索機能が足りなかったりするところがあります。そういうところはすぐにも充実させていきたいと考えています。それから、専門家と相談してテキストマイニング的な機能を追加し、さまざまな付加的な情報をつけていければと思っています。

さらに少し長期的な話になりますが、新着論文のレビューにかぎらず、さまざまな種類や性質の日本語コンテンツを集めて、同じように編集のプロセスを導入することでわかりやすいものにし、転載・改変・再利用すべて自由に行えるようなかたちで公開していきたい

と思っています。

問いかけ

「新着論文レビュー」にたずさわっていて、いろいろ思うことがあります。

たとえば、原著論文あるいは実験データは誰のものか。いわゆるアウトリーチ活動は誰がするのか。サイエンスの世界に“著作権”はなじむのか。研究活動の出版に営利企業である民間の出版社はどのような役目をはたすのか、ということなどです。これらはどれも非常に重い問題で、簡単に答えがでるようなものではありません。

今回の SPARC セミナーの「背景」にも「研究者はそれぞれの研究について、内容や成果を分かりやすく発信する取り組みを進める必要がある」とあります。研究者はますますアウトリーチ活動を求められつつあるわけですが、研究者は研究に忙しいし、アウトリーチ活動といわれても苦手でできないという方も多いのが実際です。

そのようななかではアウトリーチの専門家が必要で、専門の第三者（サイエンスコミュニケーター、サイエンスインタープリター）がアウトリーチ活動を行う場が必要なのではないのでしょうか。しかし、現実にはそのような場はなかなかないし、人材自体も欠けている、ましてや、それを生かす場も欠けています。私自身はコミュニケーターやインタープリターではないとっていますが、編集者というものはアウトリーチにかかわる立場としても機能していかなければならないと思っています。

コンテンツの信頼性

最後にもういちど、コンテンツの信頼性ということについてお話しします。これまで良質な日本語コンテンツは、おもに出版社が紙媒体というかたちで世に出してきました。しかし、ここに Web 化の波がやってきて、紙媒体以外のコンテンツが非常に大きなウエートを占めつつあるのが現状です。また、これまで出版社

は良質なコンテンツの発信源として機能してきたのですが、その機能自体もやや失われつつあるのかもしれませんが。そうしたなかで、Web 上のコンテンツのクオリティをいかに担保すべきかが大きな問題になっているのではないかと思います。

そこに編集者はどのようにかかわればよいのか。とくにサイエンスの世界について考えると、原著論文と二次文献、あるいは英語と日本語にかかわらず、コンテンツの質を維持するためには、その道のプロ（編集者）の手が必要だと私は確信しています。

では、コンテンツの電子化や Web 化の流れを受けて、どのようなしくみで編集者がそこにかかわっていけばよいのか。その費用はどう分担するのか。要するに、どのようなビジネスモデルをつくれればよいのかということがいま問われていて、それはサイエンスコミュニティ全体として検討しなければいけない問題ではないかと思います。私どものはじめた「新着論文レビュー」がそのひとつのモデルケースになればと思っています。

謝辞

最後に、所属するライフサイエンス統合データベースセンターのみなさんに感謝します。とくに、9 月末までセンター長を務められていた高木利久先生にはたいへんお世話になりました。また、川本祥子さん、坊農秀雅さん、中尾光輝さんにはさまざまにお手伝いいただきました。心から感謝いたします。