

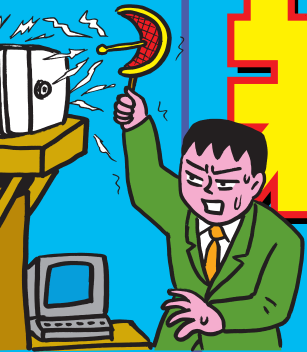
目指せ! ネットエスパー

インターネット 新検索術

インターネットの中に
潜む膨大な情報
たち。そこから欲しい
ものを見つけ出し
て活用するには、

もはや1検索サイトだけでは役に立たない。ネットの海を自在に泳ぎ、必要な情報をすぐ取り出して活用する、そんな「ネットエスパー」に変身すればインターネットの利用価値は無限大になる。この連載で「ネットエスパー」に変身するスタートを切ろう!

二木麻里(アリアドネ運営)  ariadne.ne.jp
Illust: Ebisu Yoshikazu



第10回

バイオテクノロジーの情報を選ぶ

この夏、米国国立衛生研究所(NIH)やエネルギー省(DOE)などによるヒトゲノム計画が完了すると大きく報道され続けた。30億を越す塩基配列からなるヒ

ト遺伝子の解読作業がほぼ完了するというのだから、それは大ニュースだろう。自分が利用する医薬品や医療情報について、インターネットで調べることはすっかり日

常になった。その一方でヒトゲノム・マップなどの詳細な資料や入門的な解説、遺伝子操作への懐疑の声までを、ネットの中でどのように追っていきけるだろうか。



まずは情報ソースの系統を把握しよう

話題のヒトゲノム計画を知りたいとする。まず、情報分布の構造を考えよう。調べものに先立つ一般的な把握として、こういう分野で思い浮かぶ情報ソースは4とおりある。1つ目は「政府や公的機関」による情報。医学や生化学であれば厚生省や国立遺伝学研究所、またバイオ産業という角度であれば通産省も入る。2つ目は「民間企業や商業サービス」によるもの。これはバイオテクノロジー開発企業や製薬会社などが挙げられる。3つ目は「学術機関」。大学や医学部付属の研究所などだ。そして4つ目に「個人の発信する情報」。バイオ研究者や愛好家によるページから、私財による非営利の研究機関などまでがここに入る。ある意味でもっともネットらしい領域だろう。ドメインで粗く分類するなら governmental、commercial、educational、organizational という系

統だ。日本のサイトであれば go.jp、co.jp、ac.jp、or.jp または ne.jp ということになる。もちろん実際にこのとおり gov、com、edu、org というドメイン区分で情報の内容が分かれているとは限らない。だが、資料の性質には概して違いがある。優劣ではなく、差異があるのだ。

たとえば公的なサイトでは白書や統計資料、政府報告などの大規模なデータが入手できる。またその分野で新しく制定された法律などのニュースが掲載されている場合もある。しかし、こういった資料は加工前の基礎データが手に入る一方でしばしば使いにくい。これはデータ量が多いことや、サイト設計のユーザーインターフェイスの良し悪しといった問題もあるが、なにより生の基礎データというものが持つ本質的な読みとときにくさがあるからだ。逆に民間企業が発信する情報はこうした大規

模なデータを使いやすく分類して提供している。より実用性が高く、公的機関の持つ制約がないぶん、データの解釈なども明解な場合がある。しかし一方でその企業に不利益をもたらすような情報は掲載されにくいし、アクセスが有料のことも多い。

学術機関は非営利の資料としての客観性がある程度期待できるが、学内限定など別の壁もある。個人の発信する情報については、情報の質と量の幅広さが最大の特徴だろう。情報の質に対する評価にもっとも慎重だと熟練を要する領域だ。

この4種類に、「報道」というジャーナリスティックな領域をつけ加えるのもよい。オンラインニュースなどだ。ドメイン区分では現れない部分だが、多層的な情報が平易に編集されているからわかりやすい。時系列に全容を把握しにくく、断片的だともいえるのだが。

全部で4とおりないしは5とりのソースがあると考えたうえで、検索サイトでキーワードを入力し、サーチをかけよう。どんと出てくるヒット数には、今言った全系統の情報がたまって入っているはずだ。バイオテクノロジーやヒトゲノム計画のような世界的な大ニュースの場合、膨大なリターンが予想されるが、一体どう絞り込めばよいだろう。たとえば「一般的に、客観的に知りたい」というのであれば、個人サイトはばっさり落としてしまうやり方がある。判断がむずかしい領域の資料は後回しにするのだ。それでも客観的な資料は十分そう。では、最近日本語ページ版が登場したばかりのメガ検索サイトGoogle Jumpで試してみよう。もともとGoogleは多言語検索が強みの1つだが、このページ版はページ上のユーザーインターフェイス自体が日本語なので引きやす

い。国別の検索サイトだとデータベースが別々なので、たとえば英語のキーワードを入れた瞬間、しばしば英語版に飛んでしまう。だが、Google日本語版では各国の検索サイトを行き来せずに、1か所からポータルレスで検索できる。biotechnology、human、genomeと3つのキーワードを入れ、「すべての言語」を指定するとヒット数は5万2000件。Googleはandやorといったブーリアン記号を使わないシンプルな入力法をとるが、言語構成を自分で設定するカスタム機能などがあって、大きなふるい分けや「ネット調査」に向いている。いまこれを英語と日本語にカスタマイズして、おなじキーワードで引き直すと5万1600件。ここにキーワードgovを足して絞り込むと9090件。うん、減る減る。go.jpでは419件、ac.jpで1020件になる。この結果です

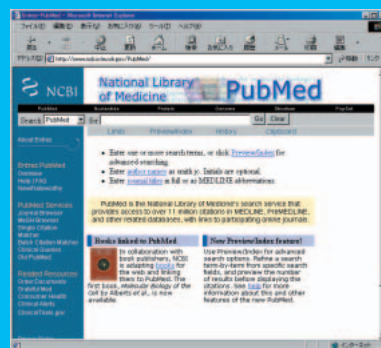
に見当たるのは、この分野の学術機関では、日本のサイトであっても英語を使っているページが多いらしいということだ。実際、いわゆる理系のページで英語のサイトは多い。国際的な共同研究上も、日本語だけではどうにもならないというプラクティカルな選択が関係しているからだろう。



Jump www.google.com/intl/ja/

医学、生科学文献の宝庫 PubMed

バイオ関連の文献資料庫として、米国の無料データベースPubMedを挙げておきたい。本文中にも出てきたNational Center of Biotechnology Information Center (NCBI)によるもので、Entrezという独自のサーチシステムをそなえている。内部だけでなく外部のオンラインテキストやジャーナル類にも豊富にリンクしており、日付指定なども行える。ジャーナルは一覧リストにもなっている。その機能性の高さや、素早い公開など、専門文献データベースのなかでも画期的といわれたものの1つだ。用語集としても活用でき、人名、雑誌名などでの検索が可能である。なお、ヒトゲノムのマップもある。蛋白質構造など数種のデータベースを組み合わせであり、検索窓でGenomeを指定して引くと出てくる。



PubMed
Jump01 www.ncbi.nlm.nih.gov/PubMed/



Homo sapiens Map Viewより
Jump02 www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?ORG=hum&MAPS=cntg-r



3 リソースの並行利用をしよう

いよいよ具体的にヒトゲノム計画、つまり Human Genome Project を知る資料をそろえよう。

① 政府資料をゲットする

この計画そのものがネットで発信されているなら、もちろんまずそれを押さえる。だが政府サイトのオリジナル資料をいきなり頭からすべて読もうなどという苦しいことは考えないでおく。

② ニュース資料を使う

ヒトゲノム計画がいつから始まったかなどの時系列を政府のサイトで読み、そのサイトを開いたまま、信頼できる報道サイ

トに行く。そしてその年次をコンパクトに区切って検索を行う。そして過去の記事のある程度連続して読んでくる。ニュースのバックナンバーを連続して読むのは紙媒体ではかなり苦勞するが、これはオンラインデータの強みの1つだろう。まことに快適なのでぜひ活用したい。

③ 学術系のサイトでこの分野のコアな情報を拾う

これは多角的なリンクをたどるための「ゲート」だ。この「ゲート」の探し方はいくつかあるが、今回は研究計画なので、①のオリジナルサイトから人または関連機関でたどれそうだ。

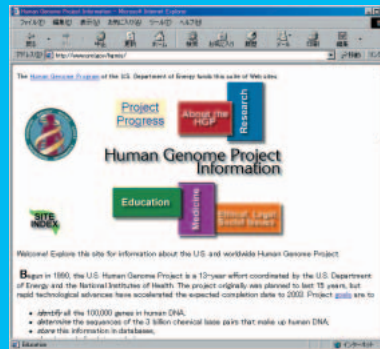
④ バイオ辞典や用語集を入手しておく

先月、辞典類のありがたみについてふれたが、文献を読むサポートツールとして画面上にいつもオンラインの辞書を開いておこう。わからない用語があったらすぐに手元で辞書を引ける。

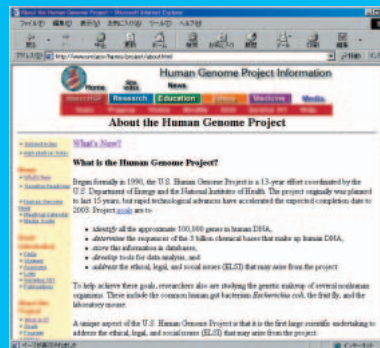
ここまできて気がつくことだが、要するに資料が紙だろうとデジタルデータだろうと、何かを知りたいときの調べ方にさして変わりはない。上のように検索方法を先に整理してしまえば、1つずつの資料を引いてくることは全然むずかしくないのだし、残暑の炎天下に図書館に行くまでもない。

4 シンプルに実行に移してみよう

まずは政府資料。biotechnology human genome project gov と入力。このキーワードで探るのが面倒なら、米国立衛生研究所 (NIH)、エネルギー省 (DOE) と固有名詞をさらに追加してみよう。実際、固有名詞というのはキーワードとしてもっとも強力なもの1つだ。Google で英語のみを指定してこれを引くと、検索時間0.06秒で冒頭に次のサイトが出てきた。Human Genome Project Information (ヒトゲノム計画の情報) **Jump01**。The Human Genome Program of the U.S. Department of Energy funds this suite of Web sites. とある。そのものずばりが出てきたようだ。サイトに行ってみると、資料がひとまとめにしてある。Project Progress、About the HGP、Research、Educationなどの項目があって、手回しのよいことに Ethical、Legal、Social Issues まである。この計画中3パーセントから5パーセントの予算は、遺伝子情報を使用するにあたっての倫理的、法律的、社会的側面を研究するのにあてられているそうだ。やれやれ。



Human Genome Project Information
Jump01 www.ornl.gov/hgmis/



About the HGP
Jump02 www.ornl.gov/hgmis/project/about.html



次にAbout the HGP [Jump02](#)を読むと計画開始は1990年とある。オーケー。ニュースサイトに行こう。オーソドックスに、たとえばワシントンポスト [Jump03](#)を選ぶ。さすがに1990年当時のバックナンバーはないが、1997年以降は読める。サイト表紙の下部に時間指定の窓もあるから、必要に応じて範囲を区切ればよいだろう。ひとまず1997年1月1日から現在までを指定し、Human Genome Projectでサーチをかけると24件ヒットする。最新のものが一番上にくるので、最後のページから逆行して読めばよい。なお、日本国内の無料バックナンバーとしては毎日新聞 [Jump04](#)などがよく知られている。いま「ヒトゲノム計画」でサーチすると12件の関連記事を読める。

ワシントンポストに戻ろう。Human Genome Projectの計画進捗状況から読

み始めると、ぐっとわかりやすい。量が多ければ惜しまず印刷しよう。またResearchのページを見るとこの組織が発行しているニュースレターがある。この種のレターはよく見かけるが、ざっと見て一般向けであれば絶対に押さえておこう。さらにリンク先を見てみると、数は少ないが米国国立バイオテクノロジー情報センター(NCBI)など、巨大ゲートにつながっている。一般向けの情報を読んだあとなので、眼前にあるものが「自分が理解すべき範囲の情報かそうでないか」は最初の時より見当がつくはずだ。このリンクが完全に専門的で一般向けの辞書がなければ、検索サイトにもどって2、3本探してこよう。ここではBioTech's Life Science Dictionary [Jump05](#)とOnline Life Science Dictionary "WebLsd" [Jump06](#)を例に挙げておく。

これまでの流れでかかった時間は1、2時間。一般教養の範囲なら十分にクリアできそうだ。さらに調べたい場合は別の方法をいろいろ試してみしてほしい。



Washingtonpost.com
[Jump03](#) www.washingtonpost.com

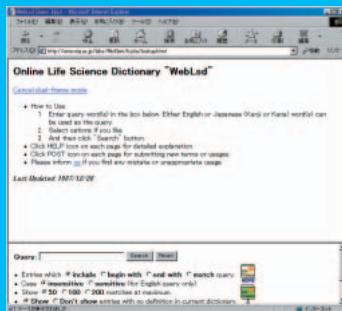


毎日新聞
[Jump04](#) www.mainichi.co.jp

おすすの科学辞書



BioTech's Life Science Dictionary
[Jump05](#) biotech.icmb.utexas.edu/
search/dict-search.html



Online Life Science Dictionary "WebLsd"
[Jump06](#) www.nig.ac.jp/labs/MolGen/
fujita/lookup.html

5 今月のポータルキット

まずは資料の性質を分類し、それを並行して使うことが大事だが、今回の最大のポイントは「個人の発信する情報を使わずにすませる」ことがもしい、もちろんケースバイケースであって、たくさんのアプローチのなかの1つのあり方として、ということだ。個人が発信している情報を否定してしまつたら、ネットの楽しさは激減する。あなた自身の活用方法を模索

して、いろいろとトライしてみたい。そしてもちろんネットに限らず、紙の本から得られる情報もお忘れなく。ともあれ頭からネットに飛び込む前に、「これとこれがあれば、たぶんここまでは調べられる」と割り切っておくのは1つの手かもしれない、といつても、思いがけない連鎖によってのめりこんでいくところがネットの醍醐味なのだけれど。

1	基本	情報ソースの系統をまず把握
2	選択	個人のサイトを「落とす」
3	設計	リソースの並行利用を準備
4	試行	シンプルに実行
5	援護	文献アーカイブ

二木麻里(ふたき まり)
上智大学外国語学部卒。翻訳家。社会・人文科学系の国内外資料を案内した総合サイトARIADNEを運営。著書に「思考のためのインターネット」(筑摩書房ちくま新書)など。



[インターネットマガジン バックナンバーアーカイブ] ご利用上の注意

このPDFファイルは、株式会社インプレスR&D(株式会社インプレスから分割)が1994年～2006年まで発行した月刊誌『インターネットマガジン』の誌面をPDF化し、「インターネットマガジン バックナンバーアーカイブ」として以下のウェブサイト「All-in-One INTERNET magazine 2.0」で公開しているものです。

<http://i.impressRD.jp/bn>

このファイルをご利用いただくにあたり、下記の注意事項を必ずお読みください。

- 記載されている内容(技術解説、URL、団体・企業名、商品名、価格、プレゼント募集、アンケートなど)は発行当時のものです。
- 収録されている内容は著作権法上の保護を受けています。著作権はそれぞれの記事の著作者(執筆者、写真の撮影者、イラストの作成者、編集部など)が保持しています。
- 著作者から許諾が得られなかった著作物は収録されていない場合があります。
- このファイルやその内容を改変したり、商用を目的として再利用することはできません。あくまで個人や企業の非商用利用での閲覧、複製、送信に限られます。
- 収録されている内容を何らかの媒体に引用としてご利用する際は、出典として媒体名および月号、該当ページ番号、発行元(株式会社インプレス R&D)、コピーライトなどの情報をご明記ください。
- オリジナルの雑誌の発行時点では、株式会社インプレス R&D(当時は株式会社インプレス)と著作権者は内容が正確なものであるように最大限に努めましたが、すべての情報が完全に正確であることは保証できません。このファイルの内容に起因する直接のおよび間接的な損害に対して、一切の責任を負いません。お客様個人の責任においてご利用ください。

このファイルに関するお問い合わせ先

株式会社インプレスR&D

All-in-One INTERNET magazine 編集部

im-info@impress.co.jp