

第79回
(2014.11.12、
11.19、11.26)

『インターネット情報およびデータベースとその活用法』 金子周司教授(薬学研究科)

●第7回講義(11月12日)...受講生数:15名

・配布物:スライド

http://www.kulib.kyoto-u.ac.jp/modules/support/content/2014tansaku/INDB_20141112.pdf

◇講義目的 <インターネット情報およびデータベースとその活用法>

・インターネットやデータベースの現状を理解し、得られる情報の効用や特性について述べることができる。

◇講義内容

○学術情報とインターネット

- ・研究と文献(literature): 研究の価値は「originality(世界初)」。研究には「1.研究テーマを決める、2.文献情報から仮説を立てる、3.実験や調査をする、4.結果を文献と合わせて考察する、5.論文を書いて公表する」といったプロセスがあるが、2.のプロセスでまず何が既に研究されているかを漏らさず文献調査することが大事。
- ・「論文」の一般的構造: 1.表題と抄録(後に二次情報として検索される情報となる)、2.序論、3.実験方法、4.結果、5.考察、6.引用文献。このフォーマットを用いることで、他人による検証が可能となる。
- ・文献=学術情報?: 学術情報は図書や雑誌といった永続性が保証されている文献のほかにも、データベースやウェブページのように書き換え可能で、必ずしも永続性が保証されていない情報もある。学術情報の形が変わってきている。
- ・電子書籍(Kindle/iPad): 電子書籍の普及が進んでいる。電子書籍では全文検索やマーキングも可能であるし、1台のデバイスの中に大量の書籍を入れておくことができる。
- ・学術情報は本来、インターネット向き: 世界共通、早さ、少部数、論文数が膨大、検索できないと意味がない、引用等がリンクできると便利、等の特性から、学術情報はいち早く電子化されるようになった。
- ・文献検索: 従来は書誌情報DBの二次情報から冊子の一次情報にあたっていたが、現在はダイレクトに電子的な形で一次情報にアクセスできるようになった。

・PubMed

:医学図書館 NLM の下部組織として遺伝情報センターNCBIを設置

:研究者が作った遺伝子情報と図書館の書誌情報が融合

:今では化合物や疾患も統合されている

:統制語 MeSH が柔軟な検索を可能に

- ・学術論文が出版されるまで: Peer-Review による内容保証。採択率は1~3割程度が多い。また編集作業も電子的に行われることがほとんど。
- ・論文と学術誌のステイタス: 他者に読まれて引用(Citation)されることが価値を決める(Impact Factor という指標)。研究者が投稿誌を決める際に利用される。Journal of Citation Reports にて調べられる。
- ・CCライセンス: 従来は著作権は出版社にあり、作者にはないことが多かった。しかし、現在作者が決められる著作権の意思表示の取り組みがある。(creative commons ライセンス) PLOS.org など一部はCC表示で再利用を広く認めている。
- ・冊子体のない Open Access ジャーナル: 従来の情報を得るものが買う購読モデルから発信者が負担する Open Access へ。PLOS.org では読者の動向がリアルタイムに表示される。

○ライフサイエンス辞書 Life Science Dictionary (LSD)

- ・なぜ電子辞書を作ったか: 卒業研究や英語での論文執筆や国際学会の発表など英語で苦労する学生・大学院生が多い。需要はあっても大変なので電子用語集は誰もつくりたがらない。
- ・オンライン辞書サービス WebLSD: 英和・和英・出現頻度・発音・用例・共起検索・PubMed連携
- ・どうやって作ったか
 - : 1. Native speaker の PubMed 抄録などから大量のテキストコーパスを作成(fair use に該当する)、2. 単語の出現頻度や隣接数を数えて用語を収集(英語、日本語それぞれに)、3. 同じ意味をもつ語句を対訳とする、4. MeSH ツリーと照合する(上位語下位語を整理)。
 - : 専門用語は数多くの同義語を持っているため、表記のゆれが大きい。統制語を用いて制御する。
 - : 関連性の算出(suggest 機能)、シソーラスとしても用いることができる。
- ・応用例: マウスオーバー辞書(複合語検索も可能)、Kindle 辞書と PDF 辞書、iPad コンテンツ(無料の英単語帳アプリ)、

電子カルテからの情報抽出。

・LSDプロジェクト: コーパスに基づく専門用語の分析、論文執筆を助ける共起検索、専門英語教育への活用、シソーラスからテキストマイニングへ。

○演習に向けて(ネットの落とし穴)

■データベース演習: 課題の概要、目標の紹介。代表的データベースを使えるようになること、検索した情報の要約ができること、情報が少ない場合はその原因を考えられることが目標。

・データベースの紹介: 京都大学教育研究活動データベース、CiNii Articles、J-GLOBAL、Web of Science、KAKEN

・研究業績検索のポイント

- :複数の情報源をあたる
- :本当にその人かどうか確かめる(同姓同名も多い)
- :いつの情報か、発信者は誰か確かめる
- :情報を整理する(重要な情報は何か)
- :情報が少ない場合の理由を考える
- :ネット検索もしてみる

・ネット検索の落とし穴: 日本語と英語(必要なら英語でも調べる)、同姓同名(場所や内容で判断)

※実は、もうひとつの顔:金子先生は危険ドラッグ捜査にも関わっている。ニュース情報をあたると、危険ドラッグに関する情報が出てくる。大学の研究者の多面性がうかがえる。

■インターネット演習: 課題の概要、目標および過去問題の紹介。情報やサイトの信頼性、検索語による結果の差異などを考察できるようになることが目標。出典を正しくレポートに記入する。

・ネット検索の基本:過去の演習課題例から。(正しいキーワードは何か、情報源はどこか)

・ネット検索での落とし穴

- :流言・風説(鵜呑みにしない。客観的・科学的な証拠を探す。)
- :ネット上の学術情報(二次情報の更新スピードが速い、永続性が少ない、内容の保証が少ない。)
- :ランク変動(オプション指定で期間を限定する。キーワードの選択。)
- :リンク切れ(有料データベースを調べる。キャッシュを調べる。)
- :伝達の範囲(英語で全て探せるかというところでもない。)

・情報検索のポイント

- :情報源(一次情報)までさかのぼる
- :日本語の情報だけで満足しない
- :いつの誰の情報か確かめる
- :重要な情報は何か考える
- :ネットだけに頼らない(書籍、専門家)
- :真実や正解があるとは限らない

●第8回演習(11月19日)・・・受講者数:15名

◇演習ページ URL: http://www.kulib.kyoto-u.ac.jp/modules/support/index.php?content_id=39

■例題1

雑誌『CELL』の2013年のインパクトファクターを求めよ。計算式も書くこと。

・Cellに載った論文数:2011年=338本,2012年=415本

・Cell掲載論文が2013年に他の論文に引用された回数:2011年=12,624回,2012年=12,312回

・ $(\quad + \quad) \div (\quad + \quad) \doteq (\quad)$ (小数点第4位四捨五入)

○目的:実際に計算することでインパクトファクターへの理解を深める。

■例題2

2006年、『Cell』誌に掲載された山中伸弥先生の論文について、書誌情報と被引用数を調べなさい。

最初に「京都大学教育研究活動データベース」で書誌情報を確認し、次に下記のデータベース等の中から二つ以上を選んで被引用数を調べること。

また、それぞれの被引用数に違いがあれば、その理由についても考察しなさい。

- ・Web of Science Core Collection
- ・CiNii Articles
- ・J-GLOBAL
- ・Google Scholar

○目的:例題1の復習をかね、検索結果の違いについて考察する。

■課題 京大に在籍する1名の教員(金子周司教授は除く)について、下記データベースを利用して研究業績を調査せよ。

○目的:データベースファイルや検索語の適切な選択を体得し、情報が見つからない場合の考察を行う。

○注意点:同姓同名の人物を見分ける。何が重要な情報かを考える。

■調査レポート項目

- ・1.氏名, 現職
- ・2.研究開始年がもっとも新しい科学研究費補助金の獲得状況
- ・3.最近(概ね10年以内)の論文で最も被引用数の多いものの書誌情報(日本語、英語計2件)およびその被引用数。共著者もすべて解答すること。
※被引用数が不明または0の場合は、出版年月日が最も新しい論文を答えること。
※Web of Scienceの被引用数を解答する際は、Core Collectionからの被引用数を解答せよ。
- ・4.検索結果を考察し、以下の観点で論じよ。
 - 欲しい情報は簡単に見つかったか?もし見つからなかったとしたらなぜか?
 - 同姓同名の人物が混じっているケースはなかったか?(今回調査したケースではなかった場合も、仮に同姓同名の人物が混じっていたら、どのように対処するか検討せよ)
 - データベースによって内容が異なるのは何故か?なぜ情報が少ない場合があるのか?
 - (日本語または英語論文がなかった場合)その理由。
 - その教員の研究テーマの中で、重要かつ周囲から評価されているテーマは何か?

■使用するデータベース

- ・京都大学教育研究活動データベース
- ・KAKEN (国立情報学研究所)
- ・CiNii Articles (国立情報学研究所)
- ・J-GLOBAL (科学技術振興機構)
- ・Web of Science Core Collection (トムソン・ロイター)

●第9回演習(11月26日)・・・受講者数:16名

◇演習ページURL:http://www.kulib.kyoto-u.ac.jp/modules/support/index.php?content_id=41

○例題

■例題1

京都大学で学んでいる海外からの学生の数は、全国の大学の中で第何位か。最新の順位を答えなさい。

○目的:複数の情報を比較し、その中から信頼性の高い統計情報を選択する。

■例題2

イギリス留学するにあたり、危険に備えて治安情報や現地でのトラブルの具体例を調べなさい。公的機関によるものだけでなく、留学体験談等も調査すること。

○目的:インターネットでは見つかりにくい情報もあることに気付く。検索キーワードによって、検索結果が大きく変わることを体感する。

■課題

京都大学の派遣留学制度を利用して、イギリス・シェフィールド大学に留学する場合の手続きを調査する。

1. 学内選考へ応募するために必要な書類を答えなさい。
2. シェフィールド大学への派遣留学に応募するために必要な資格・条件をすべて答えなさい。
派遣先大学の応募資格も詳細を確認すること。
3. (1)(2)の情報が掲載されているウェブページについて、解答例を参考にして以下を答えなさい。
※ウェブページの題名、ウェブサイトの名称、URLの例
 - (ウェブページの題名) 平成 26 年度 全学共通科目 "学術情報リテラシー入門:図書館と Web 情報の活用"
 - (ウェブサイトの名称) 京都大学図書館機構
 - (URL) http://www.kulib.kyoto-u.ac.jp/modules/service/index.php?content_id=3
4. シェフィールド大学の英語版公式ウェブサイト参照し、派遣留学 (※一般留学は除く) の出願方法と必要書類を簡潔にまとめなさい。なお、必要書類は英語のまま引用してかまわない。
【ヒント: 「派遣留学」 = 「exchange」】
5. (4)の情報が掲載されているウェブページについて答えなさい。
6. シェフィールド大学公式ウェブサイトの英語ページと日本語ページを比較して、情報の有用性と信頼性について考察しなさい。
【ヒント: “Information in Japanese 日本語ページ”の URL]
<http://www.sheffield.ac.uk/international/countries/asia/east-asia/japan/language>
7. (1)(6)において調べたサイト以外で、留学するにあたって役に立つと思うウェブサイトを1つ挙げ、そのサイトを評価する理由、提供されている情報の信頼性について述べなさい。
8. この課題に取り組んでの感想・反省を述べなさい。

○目的: 探すターゲットの原語で、翻訳機能などを賢く使い検索する。紙とウェブ、それぞれの特性を知り、情報を補完したうえで信頼性を見極め、価値ある情報かどうか評価する。

■ポータルサイトへのリンク

Yahoo! JAPAN (ヤフー)

goo (NTT レゾナント)

Google (グーグル)

Excite (エキサイト)

Bing (マイクロソフト)

MARS FLAG (マーズフラッグ)

はてな (はてな)

(担当: インターネット・データベース演習担当チーム)