

<p>第8-9回 (2020.6.30, 7.7)</p>	<p>『インターネット検索の落とし穴とデータベースの活用方法』 金子周司教授 (薬学研究科)</p>
<p><b>第8回講義 (6月30日) 受講生数：17名</b></p> <p>1. 講義目標</p> <ul style="list-style-type: none"> <li>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。</li> </ul> <p>2. 配布資料</p> <ul style="list-style-type: none"> <li>PandAの「授業資料」に以下の資料をアップロード <ul style="list-style-type: none"> <li>インターネット検索の落とし穴とデータベースの活用方法スライド.pdf</li> <li>演習のテーマ解説 2020.pdf</li> </ul> </li> </ul> <p>3. 講義内容 (45分)</p> <ul style="list-style-type: none"> <li>金子先生が研究室から Zoom ライブ配信</li> <li>資料：インターネット検索の落とし穴とデータベースの活用方法スライド.pdf</li> <li>第7回授業の終わりに、PandAの「課題」に掲載した予習ビデオを見ておくよう案内</li> <li>予習ビデオを見てきたかどうか、PandAの「テスト・クイズ」を使って講義の前に小テストを実施</li> </ul>	
<p><b>学術情報とインターネット</b></p> <ul style="list-style-type: none"> <li>● <b>研究における情報の重要性</b> 研究の価値は「新しい」こと。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。</li> <li>● <b>最先端の研究者に必要なもの</b> その学問領域の基本的な知識、研究対象についての最新の知識、研究や調査をする技術、研究コミュニティ（人脈）、語学力、ICT・プレゼン能力。</li> <li>● <b>情報の検索にあたってのスキル</b> 一般の情報検索は、網羅的でなくても多くは問題ないが、先行研究の情報検索は、見落としがあるのはまずい。十分かつ最新の情報を探し出すためには、まず、図書、学術論文、データベースなど情報源の特性を理解する。そして、検索ワードや絞り込みなど検索技術を磨く必要がある。</li> <li>● <b>資料の特性</b> <ul style="list-style-type: none"> <li>書籍・学術雑誌：一次情報、査読された内容、電子化の進展、永続性がある</li> <li>データベース：二次情報、分類や索引づけされ検索しやすい、永続性がある</li> <li>インターネット（Web ページ、動画・画像）：不確かで間違った情報も含む、最新情報を掴みやすい、永続性はない</li> </ul> </li> <li>● <b>学術情報は本来、インターネット向き</b> 全世界に向けて発信可能である、速報性がある、発行が少数であっても冊子体と比べて流通にコストがかからない、論文数が膨大であっても収蔵スペースを要しない、索引付けが半自動的に行われる、引用等がリンクできる、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。</li> <li>● <b>学術誌のステイタス</b></li> </ul>	

Impact Factorとは、該当雑誌1論文あたり発表後2年間での被引用数の平均値。個々の論文の影響度ではなく、雑誌全体の平均値。Journal Citation Reportsにて調べられる。最近では、個別の論文について被引用回数やSNSでの評判等から影響度を測ろうとする動き(Altmetrics)も盛んである。

### 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勞する学生・大学院生が多く、需要があった。電子化テキストが入手可能になり始めていたため、電子的な制作方法であれば最大効率で実現できると考えた。

- **内容**

1億語のコーパス(英語論文抄録)を解析してつくった医学辞書。単語の前後の文も表示することで、単語の使い方が分かりやすい。

### しくみから考えるネット検索での落とし穴

- **言葉の「ゆれ」**

- ・ 事物を表す言葉にはゆれがある
- ・ 検索語の単純一致では多くの情報を見逃す
- ・ 検索システムで語の「ゆれ」を吸収する仕組みが必要。Googleなど「賢い」システムでは実装されているが、すべての情報検索に実装されているとは限らない

- **概念や事物を表すことば**

- ・ 情報検索は「ことば」で行われるから、用語法を知ることが大切
- ・ 概念の範囲は場合によって曖昧なので、正しく限定する必要がある
- ・ 同姓同名など、固有名詞の重複もよくある
- ・ 専門的なことは日本語だけでの調査では不十分

- **シソーラス(同義語辞書)**

- ・ 概念をあらわす統制語を決め、同義語を寄せ集める
- ・ 概念の上位下位によりツリーをつくる
- ・ 概念同士の関係性を数値化し、一定形式で記述する
- ・ 人工知能に必要なリソース

- **検索スキルを上げるために**

- ・ 複数の情報源を使いこなす
- ・ 検索語を様々に変えて検索する
- ・ 情報の信憑性や客観性を確かめる
- ・ 複数キーワードや演算子で的確に絞り込む
- ・ 重要な情報は何か考え整理する(被引用調査の必要性)
- ・ 一次情報にまで遡って調べる
- ・ 情報が得られないなら理由を考えて見逃しを防ぐ

### 演習に向けて

- ・ 演習について、演習で用いるデータベース・課題の概要についての解説
- ・ 金子先生ご自身を例に、ZOOMの画面共有機能を使って、実際に検索の例を示した

#### 4. テーマ設定の説明(15分)

- ・ 演習補助者が学術情報メディアセンター南館303からZoomライブ配信
- ・ 資料：演習のテーマ解説2020.pdf

#### 第9回授業までにすること

- ・ 自分が調査したいテーマを決める
  - ・ そのテーマを研究している京大の教員を探す
- ※京大に適切な教員がいなければ、他大学の教員でも構わない

#### テーマ設定のコツ

- ・ そのテーマでレポートを書くことを想定して設定する
- ・ 最新の話題すぎると研究成果がまだ出ていない場合がある
- ・ テーマが広すぎる場合は、他のキーワードと組み合わせてみる（後述のキーワード例を参考として提示）
- ・ テーマが狭すぎる場合はその上位概念を考える

#### テーマ設定のためのキーワード例

- |          |          |             |
|----------|----------|-------------|
| ・ グローバル化 | ・ 気候変動   | ・ ノーベル賞     |
| ・ 多様性    | ・ 災害復興   | ・ オープンサイエンス |
| ・ 格差     | ・ レジリエンス | ・ AI（人工知能）  |
| ・ 人権     | ・ 発達障害   | ・ 人類史       |
| ・ 景観     | ・ 生活習慣病  | ・ SNS       |
| ・ 食料     | ・ 感染症    | ・ インターネット   |

#### 5. 個人ワーク(30分)

- ・ 各自テーマ設定および教員選択に取り組んでもらった
- ・ 質問はZoomのチャット機能にて受付けた  
主な質問は次のとおり
- ・ 理系と文系のどちらのテーマ設定が取り組みやすいか
  - 分野によっては英語論文・日本語論文が見つからない場合がある

## 第9回演習（7月7日） 受講者数：13名

### 1. 演習目標

- ・ 目的に応じてデータベースを適切に使い分けられるようになる
- ・ 検索語を変えるなど、検索方法を工夫して必要な情報を収集できるようになる

### 2. 配布資料

- ・ PandAの「授業資料」に以下の資料をアップロード
  - 演習のポイント解説 2020.pdf
  - 20200707\_総合演習について.pdf

### 3. 演習説明（30分）

- ・ 演習補助者が学術情報メディアセンター南館303からZoomライブ配信
- ・ 資料：演習のポイント解説 2020.pdf

### 前回までのおさらい

- ・ 学術情報の生産サイクル
- ・ 図書・雑誌の違い
- ・ データベースについて

### 本日の演習課題のポイント解説

- ・ 課題1-4について
- ・ 課題全体について

### 課題

設定したテーマに沿って、京大に在籍する教員を1人選び、下記のデータベースを利用して研究業績を調査せよ。（京大に適切な教員がいなければ、他大学の教員でも構わない）

なお、レポートには下記の調査項目について記載すること。

### 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  
※研究代表者となっている研究課題を答えよ。研究代表者となっている研究課題がない場合は、研究分担者となっている研究課題を答えよ。それもなければ「なし」とせよ。
3. 2010年以降の論文で最も被引用数の多いものの書誌情報（日本語1件、英語1件、計2件）およびその被引用数。筆頭著者の論文でなくてもよい。共著者がいる場合は共著者もすべて回答すること。  
※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の被引用数を回答する際はCore Collectionからの被引用数を回答せよ。  
※ヒットしなければ「なし」とせよ。
4. 教員に関する2010年以降の新聞記事  
※ヒットしなければ「なし」とせよ。

5. 検索結果を考察し、以下の観点で論じよ。

- ・ 同性同名の人物はどのデータベースで混じっていたか、またどのようにして調査対象人物を絞り込むことができたか。
- ・ それぞれのデータベースの特徴を比較してまとめよ。
- ・ 論文データベースと新聞データベースで得られる情報に違いはあったか。違いがある場合はなぜか。
- ・ その教員の研究テーマの中で、最新の研究テーマ、周囲から評価されているテーマを挙げよ。また社会的にどのような点で評価されているか述べよ。

**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)
- ・ researchmap (科学技術振興機構)
- ・ Web of Science Core Collection (Clarivate Analytics)
- ・ 聞蔵Ⅱ ビジュアル (朝日新聞)
- ・ 毎索 (毎日新聞)
- ・ ヨミダス歴史館 (読売新聞)
- ・ 日経テレコン (日本経済新聞・京都新聞)

※新聞データベースは、7/7-13は附図利用支援掛を通して、同時アクセス数を以下のとおり増やしてもらった

- ・ 聞蔵Ⅱ ビジュアル：+50 アクセス
- ・ 毎索：+30 アクセス
- ・ ヨミダス歴史館：50 アクセス
- ・ 日経テレコン：無制限

**4. 演習時間(45分)**

- ・ 各自課題に取り組んでもらった
- ・ 質問は Zoom のチャット機能にて受付けた  
主な質問は次のとおり
- ・ データベースへアクセスできない
- ・ CiNii Articles の被引用数の確認方法
  - 演習補助者がそれぞれ個別に対応した

**5. 課題発表 (10分)**

- ・ 一人一分程度で以下の内容を発表
  - 自分が設定したテーマ
  - 選んだ教員名、現職
  - 業績 (研究内容) を簡潔に紹介
- ・ 発表者からは課題に取り組んだ感想として、次のようなコメントがあった
  - 新聞データベースで検索する際、キーワードを「京都大学」ではなく「京大」とするとヒット件数が増えた
  - 文系研究者の英語論文は、Web of Science Core Collection ではなかなかヒットしない

- ・ 後日、北村先生から受講者宛てにメールにて以下のようなアドバイスがあった
  - 人文社会科学系の英語論文に関しては、KULINE 左側のメニュー「データベース」からアクセスできる「ProQuest ResearchLibrary」を試してみてください

## 6. 第10回授業の説明(5分)

- ・ 第10回演習補助者が吉田南総合図書館から Zoom ライブ配信
- ・ 資料：20200707\_総合演習について.pdf

### 全体を通して：昨年度からの変更点・感想・反省点など

#### ● 補助内容

スライド資料作成、課題作成、課題ページ等 (PandA) 作成、演習補助 (Zoom のチャット機能を使用)、採点

#### ● 2019年度からの変更点

新型コロナウイルス感染症対策に伴い Zoom を用いたオンライン授業を実施

#### 第8回

- ・ 金子先生の講義は、先生の研究室から Zoom でライブ配信
  - ※ 北村先生及び演習補助者は学術情報メディアセンター南館 303 から参加
- ・ オンライン授業であること、受講者数の少なさから、グループワークは行わず、個人ワークを中心とした
- ・ テーマは自分の興味があるものとした。考えるヒントとして、演習補助者が用意したキーワードを提示した

#### 第9回

- ・ 学術情報メディアセンター南館 303 から Zoom でライブ配信
  - ※ 金子先生は都合により欠席
- ・ 課題の成果をアウトプットする機会を設けるため、一人一分程度で簡単に課題発表をする時間を設けた
- ・ オンラインでの視聴を前提として、課題の解説資料を抜本的に見直した

#### ● 感想・反省点

- ・ 課題提出者は15名、平均点146点(15点満点)、得点率97%
- ・ 第9回の演習では一人一分程度で課題発表の予定だったが、2、3名の発表で時間切れとなった。データベースへのアクセスが上手くいかない受講者への対応に時間がかかったことが主な原因である。発表内容は、他の受講者の参考となるものであったため、次年度はより多くの人に発表してもらえるよう工夫したい。

(担当：インターネット・データベース演習担当チーム)