

インターネット情報とデータベースの活用法

情報探索入門

薬学研究科 金子周司

本日のメニュー

- 学術情報とインターネット
- ライフサイエンス辞書の紹介
- 演習のツボと落とし穴

- スライドは授業ページでPDFを公開

あんた誰？



研究と文献

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

研究の価値は「世界初」

学術情報＝文献？

- 書籍
- 学術雑誌
- データベース
- ウェブページ

永続性は保証される

永続性は保証されない

文献検索

一次情報
(冊子体)



二次情報 (抄録検索)

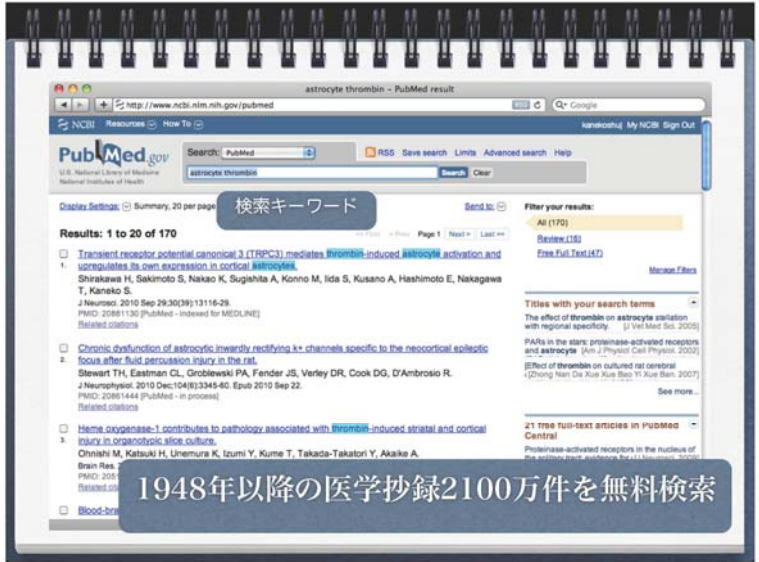


一次情報
(電子出版)



インターネットと学術情報

- 学術情報は本来インターネット向き
- 多種類, 少数数, 世界共通
- 医学・生命科学では90年代に電子化
- 抄録DBと電子編集・出版のリンク
- 遺伝子や物質DBとの統合検索



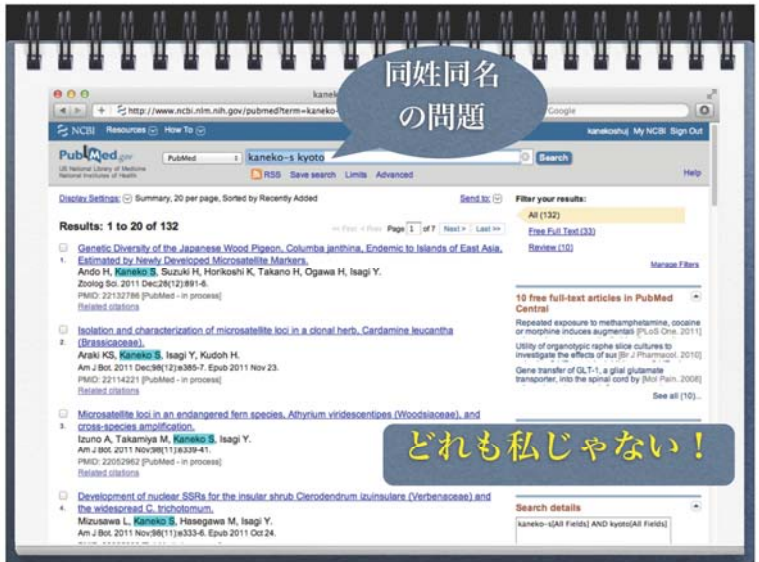
1948年以降の医学抄録2100万件を無料検索



全文へのリンク

著者名での検索

関連する論文

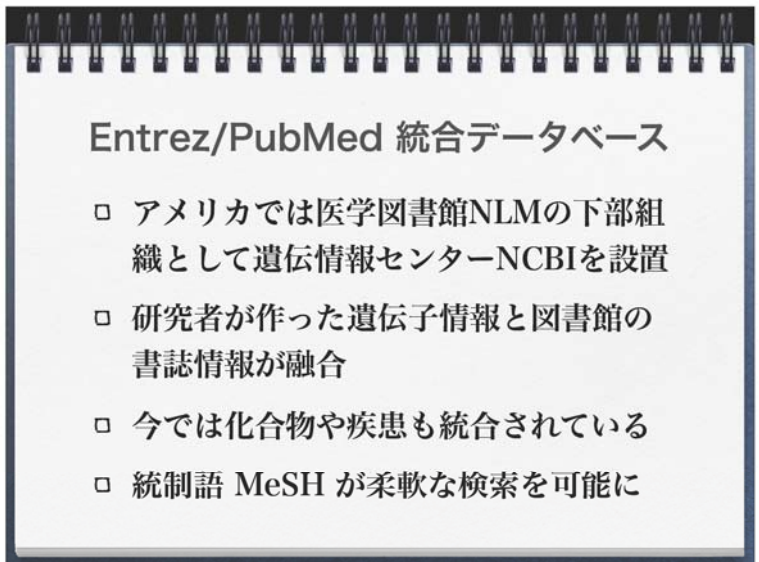


同姓同名の問題

どれも私じゃない!

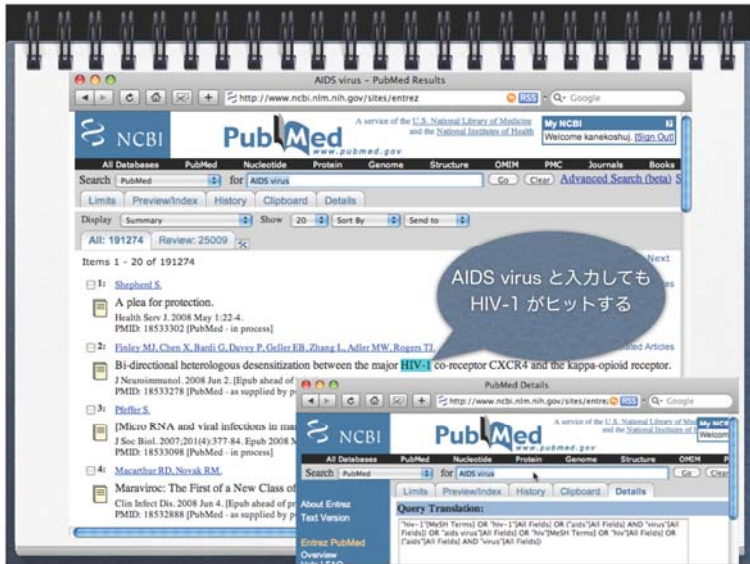


Entrez統合検索



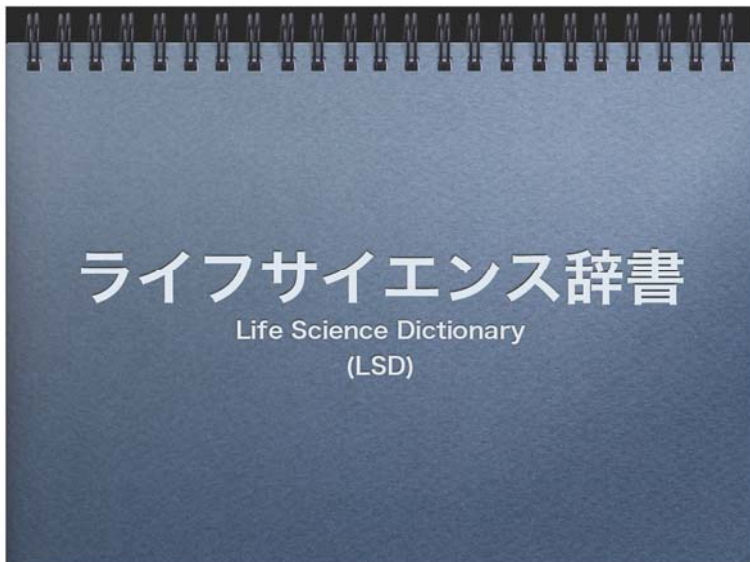
Entrez/PubMed 統合データベース

- アメリカでは医学図書館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



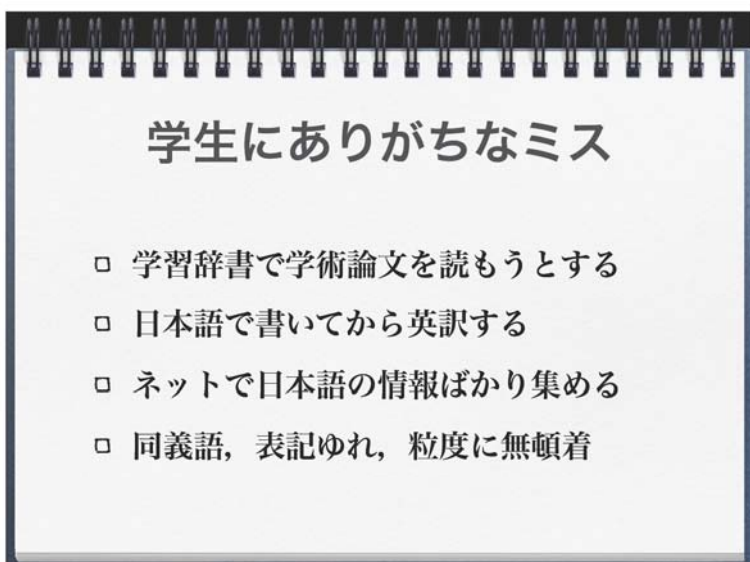
情報検索における現実的問題

- 表記のゆれ “たばこ” “タバコ” “煙草”
- 同義語 “後天性免疫不全症候群” “エイズ”
- 多義語, 特に略語 “AV” “CD”
- 階層性 “京都” “京都府” “京都市”
- 英語と日本語
“vector” “ベクター” “ベクトル”



なぜ電子辞書を作ったか

- 卒業研究に配属される4回生が最初に苦しむのは, 論文の英語
- 大学院生が苦しむのは, 英語での論文執筆と国際学会での発表
- みんな意外とできないのが, 文献調査
- 研究者は誰も作らない電子用語集



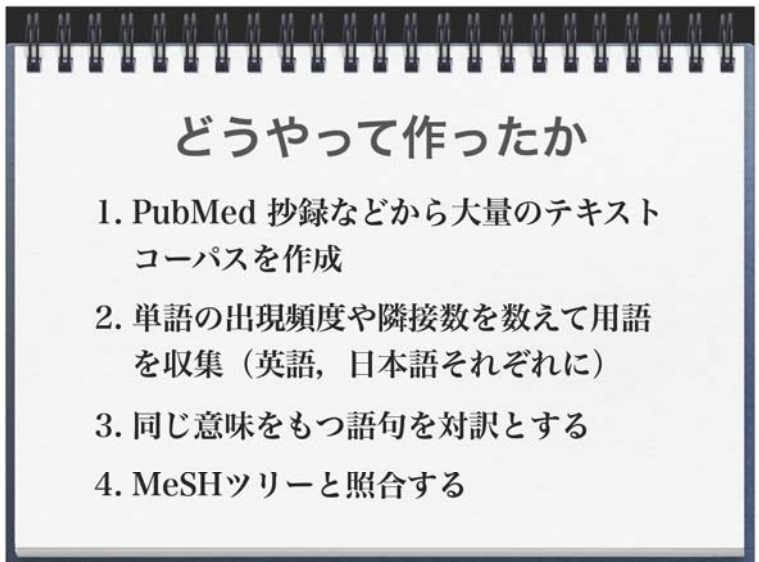
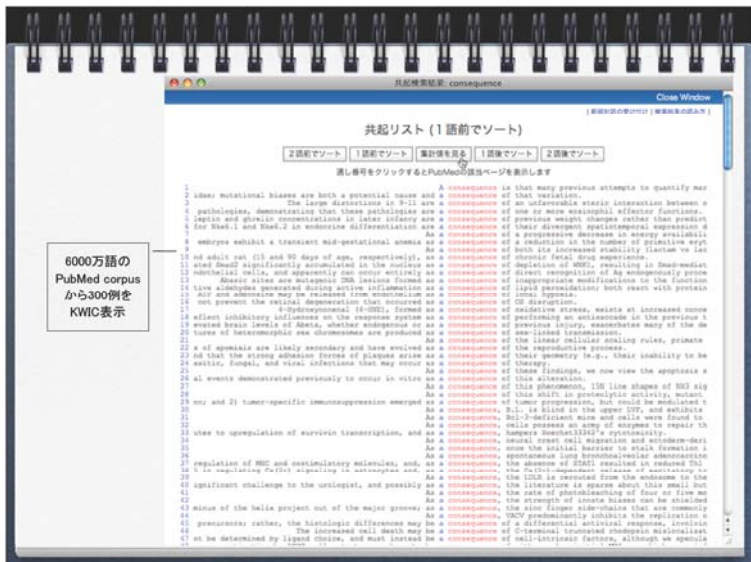
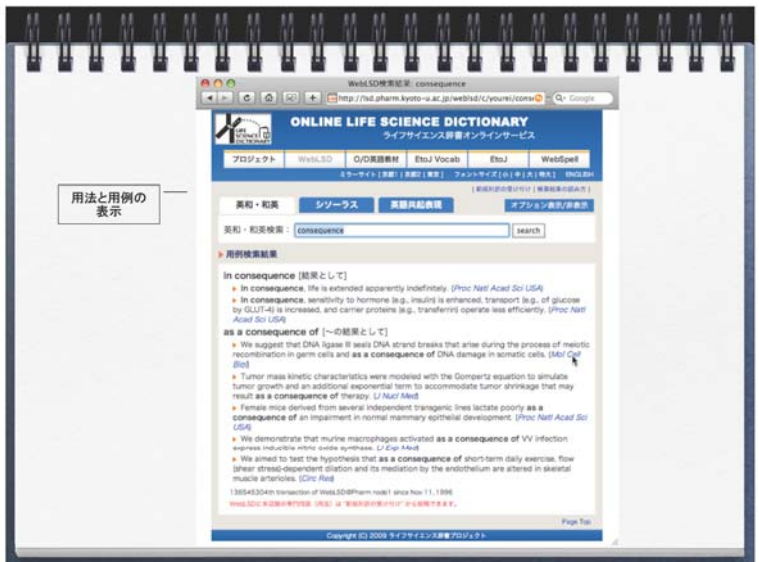
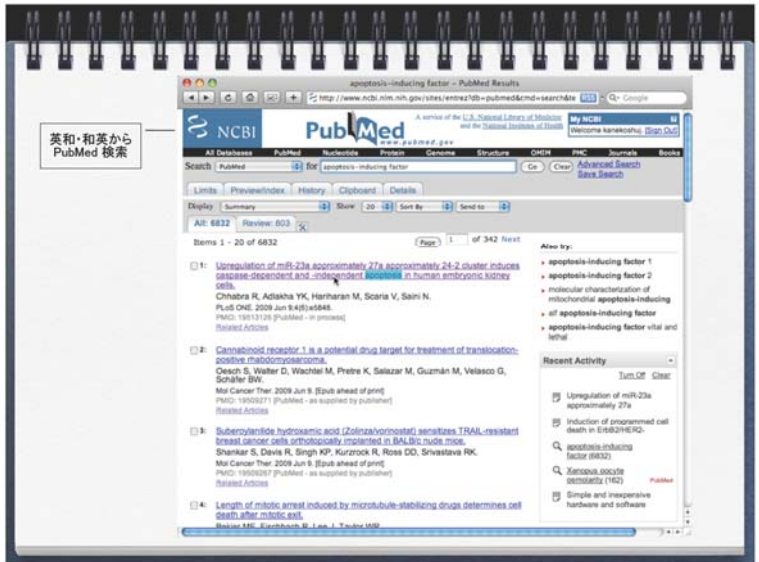
- 学習辞書で学術論文を読もうとする
- 日本語で書いてから英訳する
- ネットで日本語の情報ばかり集める
- 同義語, 表記ゆれ, 粒度に無頓着

オンライン辞書サービス WebLSD

出現頻度 音声 対訳 解説 関連語や用例を表示

英和・和英・活用辞書

WebLSD2010
英和 103,143語
和英 116,740語
音声 16,146語
例文 22,261文



どうやって作ったか

1. PubMed 抄録などから大量のテキストコーパスを作成
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

耳で覚えるライフサイエンス英語

スクリーンショット: iTunesライブラリ。プレイリストには「耳で覚えるライフサイエンス英語」というオーディオファイルがリストアップされている。単語と発音の練習用と思われる。

LSDプロジェクト

- コーパスに基づく専門用語の分析
- 電子辞書を情報ポータルに
- シソーラスから連想検索へ
- テキストマイニングへの応用

自然言語処理などへの応用を支援します

演習のツボと落とし穴

データベース演習

- 課題：京大に在籍する1名の教員について、専門研究の業績（学位論文、最近の著作物や論文と被引用数）と 科研費取得状況を調査する。
- 目標
 - 代表的データベースを使えるようになる。
 - 情報を要約することができる。
 - 情報が無い原因を考えられる。

Read & Researchmap (JST)

金子 周司
URL: <http://www.zshimizu.com>
所属: 京都大学
職名: 教授、教授、准教授
学位: 医学博士

研究キーワード
カルシウムチャネル(Ca²⁺)、アミノカルシウム結合タンパク質、TRPチャネル(Ca²⁺)、膜電位、Zn²⁺ハイマゼンチン(100)、神経細胞(20)、イオンチャネル(10)、2,2'-ビス(トリフルオロメチル)プロパン(10)

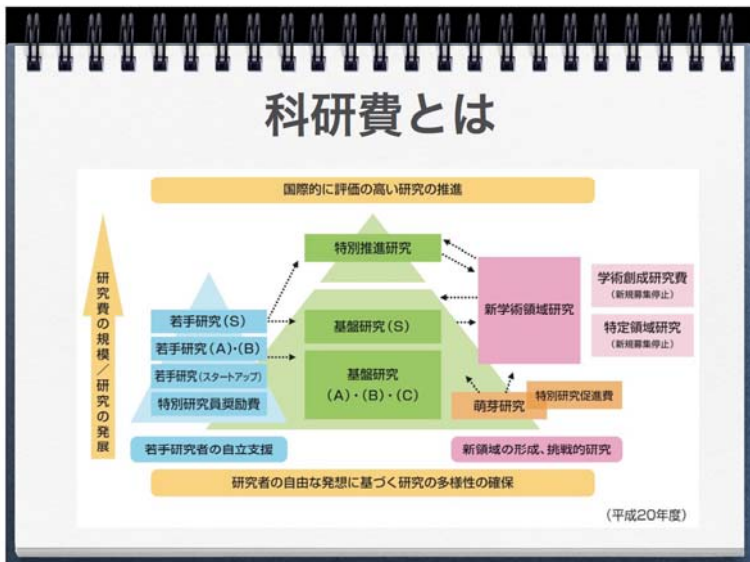
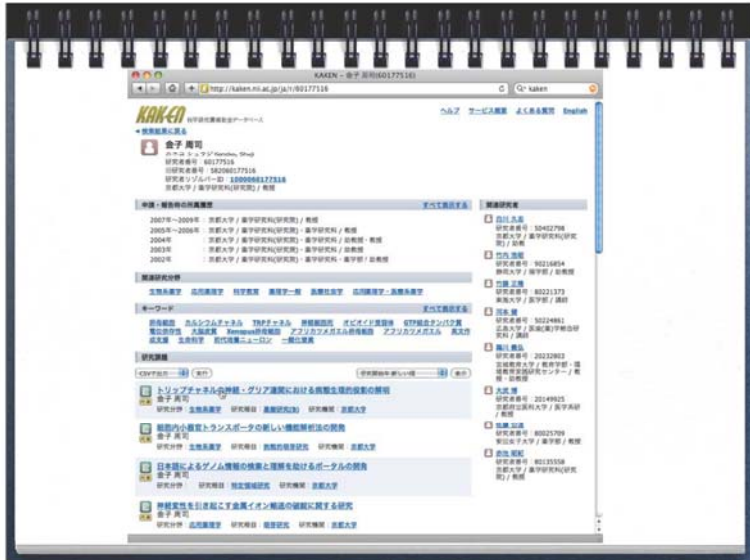
研究分野
薬学 / 生物化学(1596) /

J-GLOBAL - 金子 周司 [研究者]

Read & Researchmap 研究者情報の登録、更新はご自身で行ってください

English

英語も見ましょう

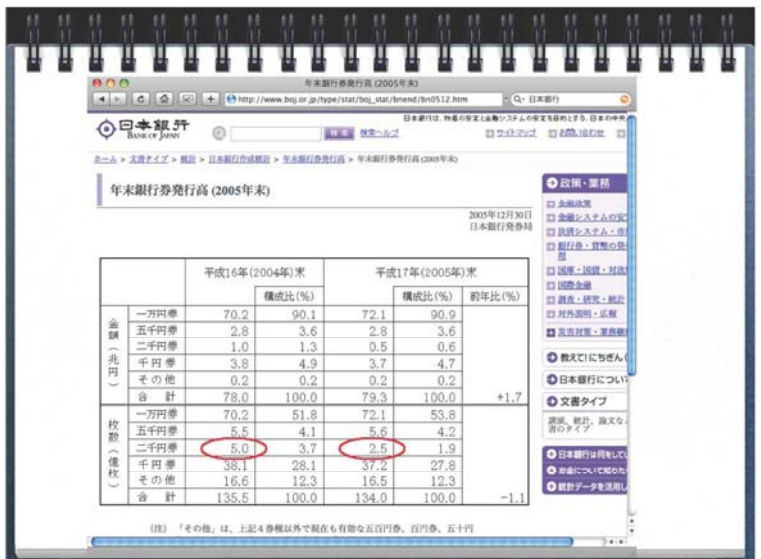
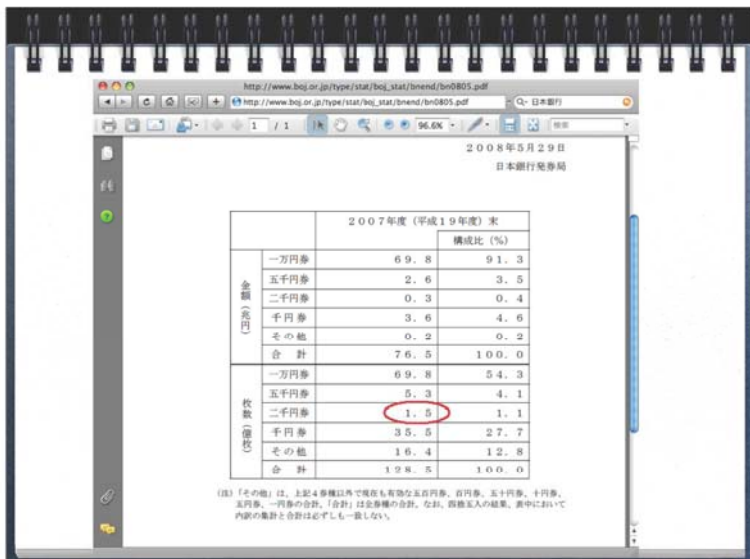


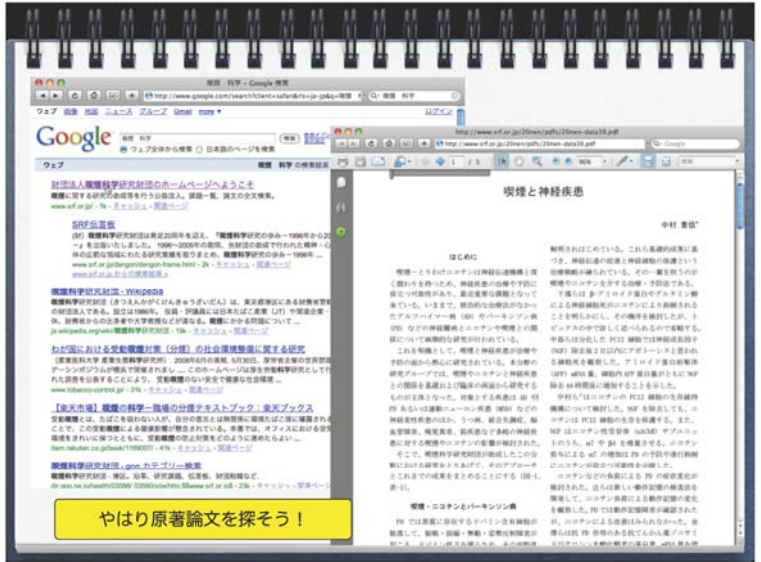
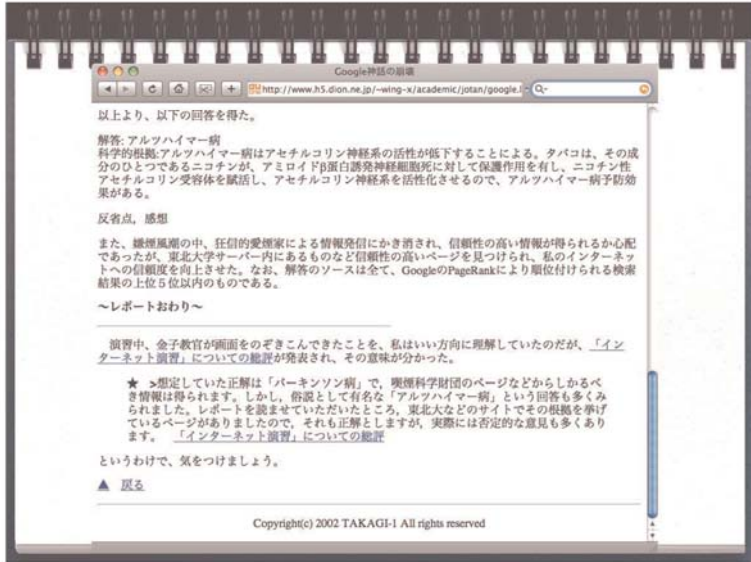
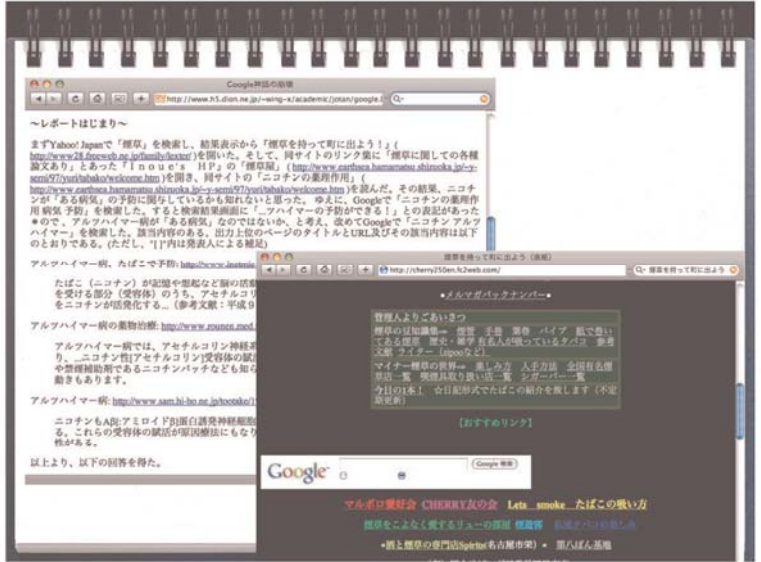
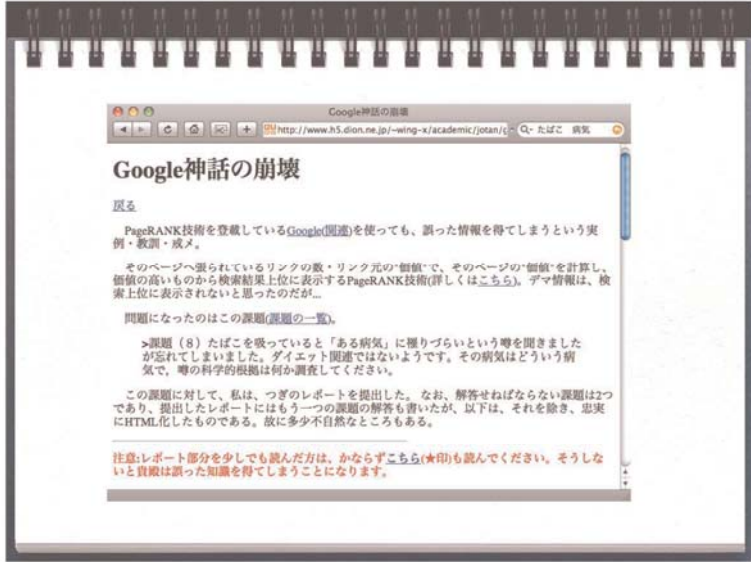
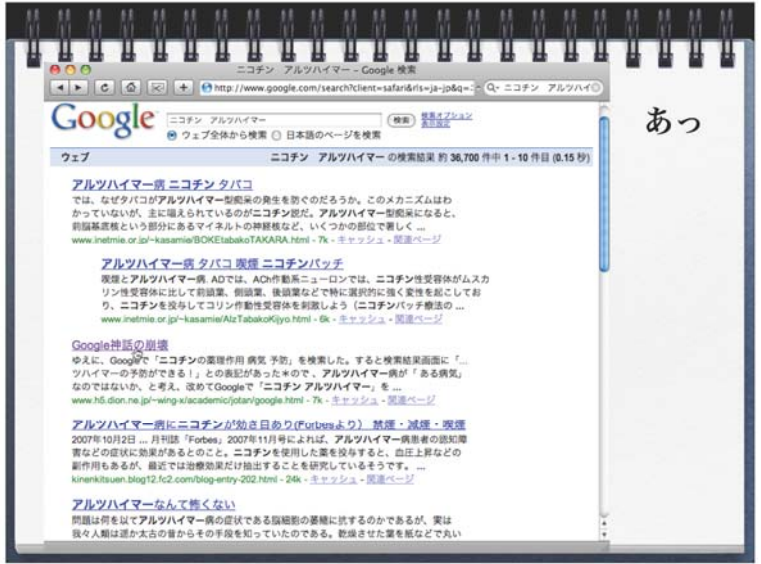
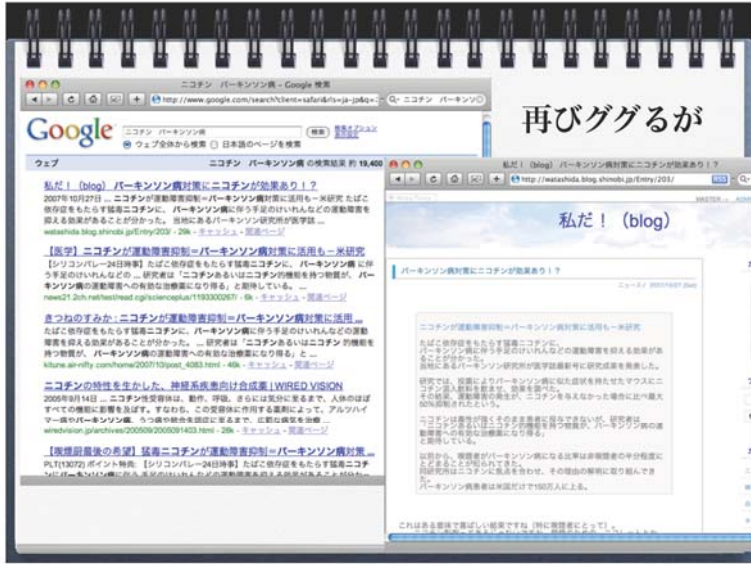
- ## 研究業績検索のポイント
- 複数の情報源をあたる
 - 本当にその人かどうか確かめる
 - 情報を整理する (重要な情報は何か)
 - 情報がない場合の理由を考える

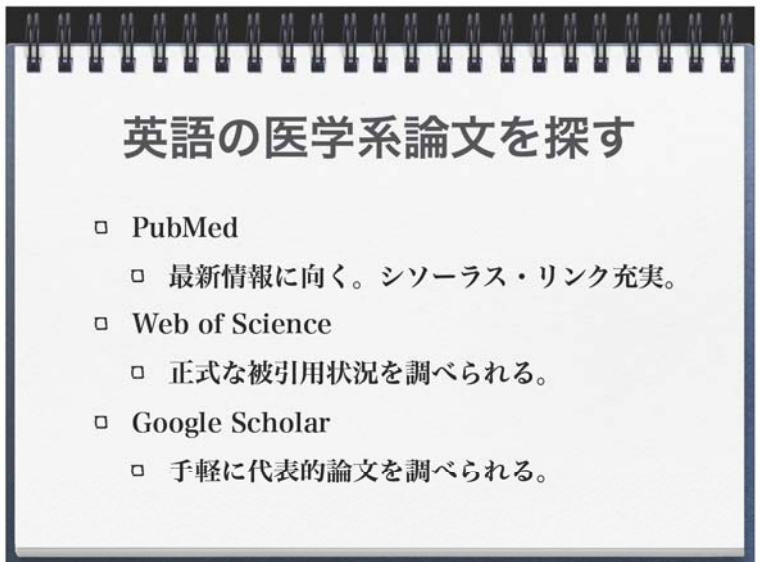
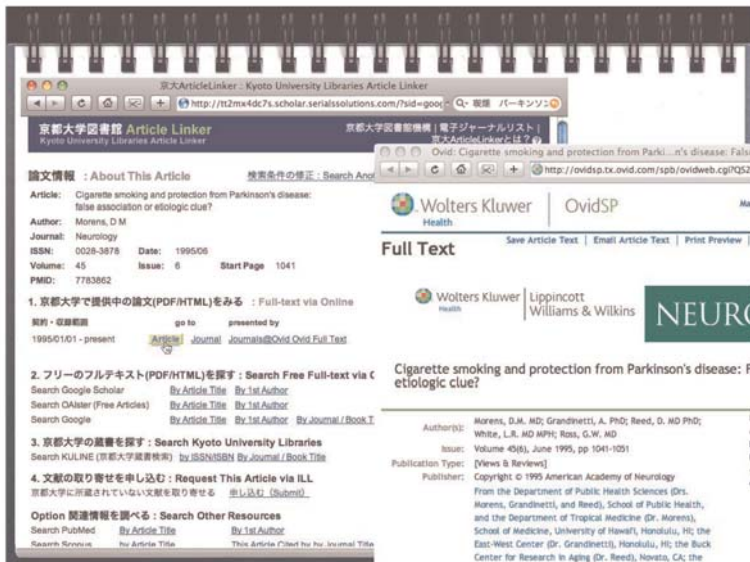
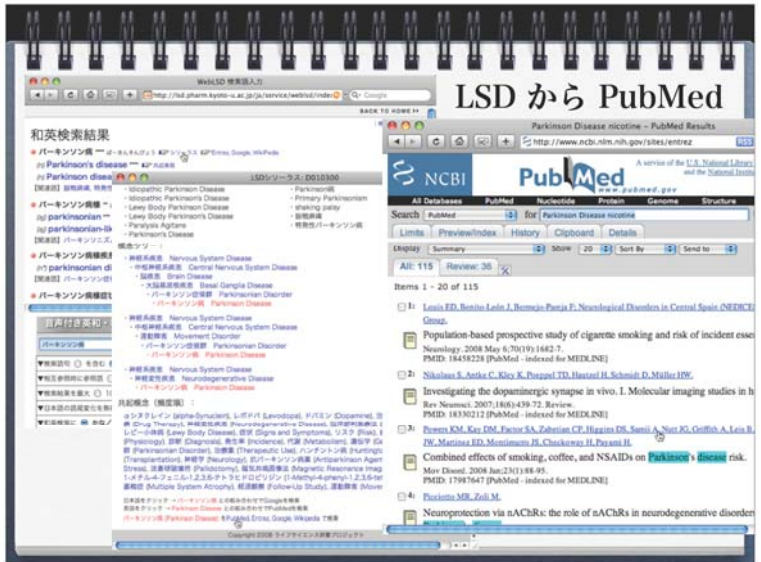
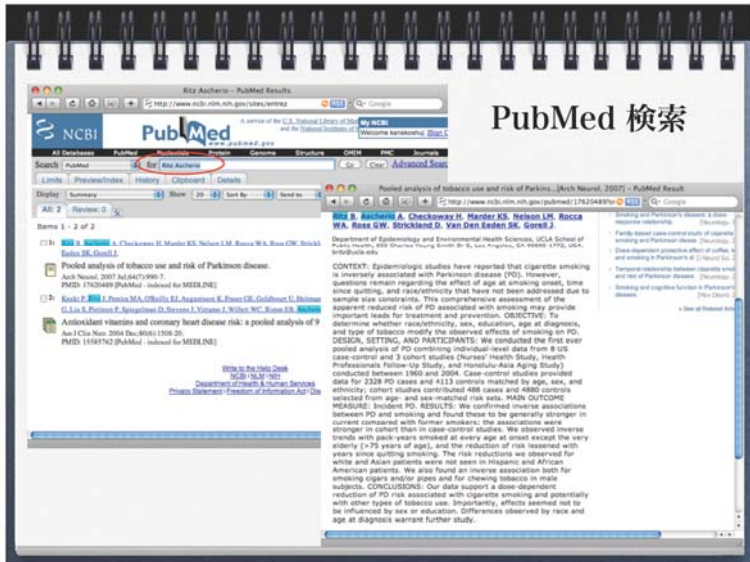
- ## インターネット演習
- 課題：演習日に授業ページで提示する課題について、インターネットを用いて情報を収集する。
 - 目標
 - 情報やサイトの信頼性、検索語による結果の差異などを考察できるようになる。

過去の問題例

1. 現在、2千円札はどのくらい出回っているか。
2. 携帯電話は国内でどれくらい使われているか。
3. タバコは百害あって一利なしと言われるが、タバコを吸っているとある病気に罹りづらいという噂を聞いた。その病気とは何か。







検索のポイント 2

- 表記のゆれ, 同義語, 言い換え表現
- ネットには諸説入り乱れている
- 科学的な情報=総説や論文を探す
- 科学的情報は英語で検索する
- 科学的情報に「定説」はない

おわり