

2017年度 春学期・(火Ⅱ)

情報

第8回目「インターネットの利用」(続き)

19. インターネットの歴史

「情報収集」

20. データベースと検索エンジン

21. インターネット時代の著作権・ライセンス

● 最初に

講義資料は、<http://astro.u-gakugei.ac.jp/~nishiura>

→ 「西浦クンの講義室」に縮小版(PDFファイル)を置く予定

東京学芸大学 自然科学系
宇宙地球科学分野 講師
西浦 慎悟

19. インターネットの歴史

第8回目

1957年10月04日：ソビエト連邦による人類初の人工衛星スプートニク1号の打ち上げ成功。東西冷戦下の西側諸国に衝撃と危機感を与えた。

→ スプートニク・ショック

- ・ 1958年 アメリカ航空宇宙局(NASA)の設立。国防総省高等研究計画局(Advanced Research Projects Agency = ARPA)の設立。国家防衛教育法など各種教育計画、理系教育・外国語教育の重点化、奨学金制度の拡充(→ 大学進学率の飛躍的増加)。
- ・ 1959年 マーキュリー計画(アメリカの有人宇宙飛行計画)。連邦議会による科学研究予算の劇的増加(3,400万ドル→1億3,400万ドル)。
- ・ 1961年 アポロ計画の目標を月面着陸に変更。
- ・ 1971年 日本の文部省の学習指導要領改訂で理数教育のカリキュラムを現代化。

1961年：マサチューセッツ工科大学のクラインロック(Kleinrock)がパケット通信に関する最初の論文を発表。

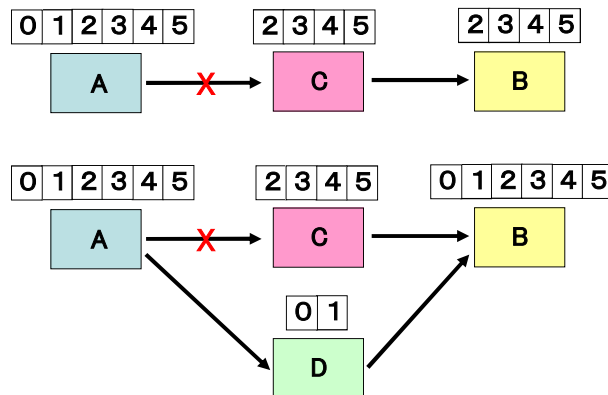
アメリカ・ユタ州でテロにより、3ヶ所の電話中継基地が破壊。軍用回線も一時的だが完全停止する。核戦争時にも使用し得る通信システムの開発が進む。

→ 分散型ネットワーク

19.インターネットの歴史

第8回目

- ・ **パケット通信** → パケット化することで、通信が寸断されても、続きのデータを別の経路・別のタイミングで送ることができる。
- ・ **分散型ネットワーク** → 情報伝達の中継点を複数用意することで、複数の通信経路を確保する。



19.インターネットの歴史

第8回目

1965年：イギリス国立物理学研究所の**ナルド・デービス**が、全国的なパケット通信網を提案するが、政府は不採用。デービスは**1970年**に、多くの学問領域にまたがるパケット通信網 **Mark I** を設計・構築し、実用性を実証する。

1966年：ARPAにて**ARPANET**の開発開始。発足当時のノードはカリフォルニア大学ロサンゼルス校 (UCLA)、UCサンタバーバラ、スタンフォード研究所 (SRI)、ユタ大学。

ARPANETの基本仕様

- 負荷共有(分散コンピュータ)
- メッセージサービス(電子メール)
- 情報の共有(掲示板/WEB)
- プログラム共有(ダウンロード)
- 遠隔ログイン

同年、ミシガン州の三つの公立大学が、教育・経済の発展に寄与することを目的に**メリット・ネットワーク**(Michigan Educational Research Information Triad)を創設。**1971年**にデモ公開される。

1969年：ARPANETにて初のパケット通信が成功。

1971年：トムリンソン(Tomlinson)が、**email**のシステムの原型を開発。

1973年：フランスで**ルイ・ブザン**がパケット交換網**CYCLADES**を設計・構築。同年、ARPANETがノルウェー王立レーダー施設、英国ロンドンカレッジ大学と国際接続。世界初の**国際間パケット通信**が成し遂げられる。

19. インターネットの歴史

第8回目

- 1974年：サーフ(Cerf)とカーン(Kahn)がネットワーク間の伝送制御手順(TCP)の仕様を発表。
- 1977年：クロッカー(Crocker)とバイタル(Vittal)が電子メールの仕様を発表。
- 1978年：インターネット運営に関する国際協力会議(International Cooperation Board)、インターネット・コンフィギュレーション管理会議(Internet Configuration Control Board)会議が開催。
- 1979年：デューク大学とノースカロライナ大学をUUCP(UNIX-to-UNIX copy connection)で接続したUSENETが開始。先駆的な研究者用ネットワーク。
- 1981年：大学や民間のコンピュータ研究グループ用のCSNETが開始。スポンサーはNSF(National Science Foundation)。1983年にはARPANETと接続。
- 1982年：東工大と慶応大間を接続。1984年には東大を加えたJUNETが開始。1986年にはCSNETに接続(日本初の海外接続)。1991年に終了。
- 1986年：スーパーコンピュータ利用を目指したNFSNETが発足。同年、日本初のIPネットワークを構築するWIDEプロジェクトが発足。1989年にはNFSNETと接続。
- 1990年：ARPANET終了。
- 1991年：CERN(=ヨーロッパ素粒子研究所、スイス)のバーナード・リーが、ハイパーテキスト記述言語HTMLやそれを伝送するプロトコルHTTPなど、WWW(World Wide Web)のシステムを開発。
- 1995年：NFSNET終了。インターネットの完全商業化。

20. データベースと検索エンジン

第8回目

● データベース(database)

追加・削除・検索・抽出などの管理・運用を容易に行えるように、規則に従って整理された情報の集まり。単なる情報・データの集まりに加えて、管理・運用を行うための機能やアプリケーションを備えたものを「データベース」と呼ぶ。

第二次世界大戦後に、米軍が、膨大な資料を効率的に管理・運用するために、それらの情報を一つに基地に集め、その基地を「データベース」と呼んだことが、名称の由来とされる。

● データベースの例

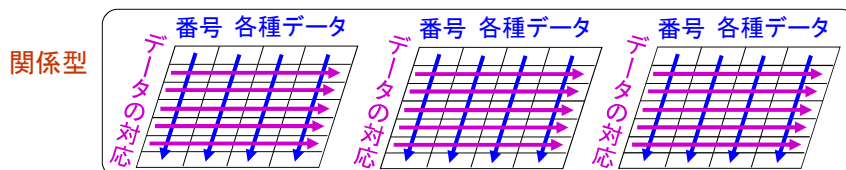
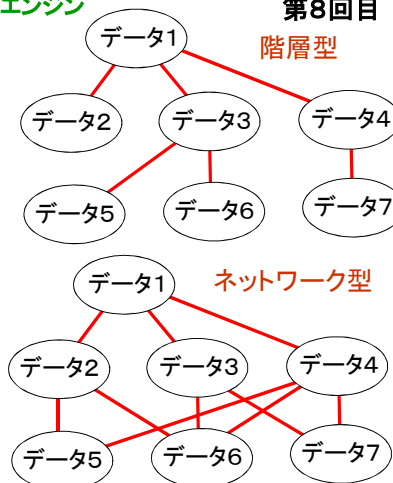
- ・住所録：氏名(家族構成)、住所、郵便番号、電話番号、職業、職場の連絡先
- ・カルテ：患者名、住所、職業、連絡先、症状、対処記録、所見
- ・会社四季報：企業名、所在地、代表社名、連絡先、社員数、株価
- ・蔵書目録(図書館)：タイトル、著者名、出版社、出版年月日
- ・音楽曲(カラオケ)：タイトル、歌手名、歌詞、ジャンル、関連番組
- ・気象・台風(気象庁)：年月日、天気、気温、湿度、風向き、台風の位置、風速
- ・検索エンジン(インターネット)：URL、各種キーワード、ファイル形式

20. データベースと検索エンジン

第8回目 階層型

● データベースのモデル

- ・ **階層型データベースモデル**：ヒエラルキカル・データベースモデルとも呼ばれる。データ構成を、上位階層から下位階層に向けて、1対多で対応させた形式。
- ・ **ネットワーク型データベースモデル**：上位階層のデータと下位階層のデータ同士を多対多で対応させた形式。
- ・ **関係型データベースモデル**：リレーショナル・データベースモデルともいう。1970年にIBMのエドガー・F・コッド(Edgar F. Codd)が考案したモデルであり、数学の集合論と述語論理に基づいている。互いに関係した複数の表のイメージで捉えられることが多い。現在の主流。



20. データベースと検索エンジン

第8回目

● 検索エンジン (search engine)

インターネット上に存在するサイト、ページ、画像、動画、音声、記事などを検索する機能やプログラムのこと。近年では、検索以外にも様々なサービスを加えたポータルサイトとなっていることが多い。

● 検索エンジンの種類

- ・ **ディレクトリ型検索エンジン**：手動作業によって構築されたWEBページ内のディレクトリを検索するシステム。人の手によって作成されるため、特定のテーマや地域に限定した場合、質の高いサイトを検索可能。インターネット普及の初期は検索エンジンの主体だったが、サイト数が膨大になった現在では、主流ではなくなっている。
- ・ **ロボット型検索エンジン**：クローラ(crawler)やボット(bot)、スパイダー、ロボット、と呼ばれるプログラムを用いて、サイト上の文書や画像などを自動的に・周期的に取得・データベース化した検索サイト。検索結果の表示順の決定も自動化されており、このアルゴリズムは、各検索サイトでは非公開とされている。現在の検索エンジンの主流。
- ・ **メタ型検索エンジン**：横断型検索エンジンとも呼ばれる。ひとつの検索ワードを、複数の検索エンジンで検索する。

● 検索式

検索内容を、検索ワードと論理演算子・正規表現などを用いて論理式形式で表現したもの。検索サイトやアプリケーションによって、記述方法が異なるので、注意が必要。

- ・ **論理演算子、論理式** : 「AND」「OR」「NOT」を基本要素とする論理演算用の演算子および式のこと。

例) 「X=20 OR Y=15」 → 「X=20」または「Y=15」

「(X NOT 20) OR (Y=15 AND Z=10)」 → 「Xが20でない」または「Y=15 かつ Z=10」

- ・ **正規表現 (regular expression)** : 正則表現とも呼ばれる。文字列の集合を一つの文字列で表現する方法。

例) 「AB*CDE」 → 「AB」と「CDE」の間に任意の文字列がある、つまり、「ABCDE」「ABZWCDE」「AB48RCDE」「ABほえほえCDE」などは全て該当する。

論理演算子、論理式、正規表現は、プログラミング言語やワード・プロセッサ、表計算ソフトなどで普通に使用されている。

● Google による検索の例

- ・ 「ほえほえ」 → 「ほえほえ」を含むサイトが表示される。
- ・ 「ほえほえ うふふ」 → 「ほえほえ」と「うふふ」を含むサイトを表示。
スペースで検索ワードを区切る
- ・ 「ほえほえ OR うふふ」 → 「ほえほえ」か「うふふ」を含むサイトを表示。
“OR”で検索ワードを区切る
- ・ 「ほえほえ -うふふ」 → 「ほえほえ」を含み、「うふふ」を含まないサイトを表示。
“-”を検索ワードの前に付ける
- ・ 「ほえほえ*うふふ」 → “ほえほえ”と“うふふ”の間に任意の文字を含むサイトを表示する。例えば「ほえほえあははうふふ」など。
任意の文字列として“*”を使う
- ・ 「cache:http://hoehoe.net」 → キャッシュされているhttp://hoehoe.netを表示する。
- ・ 「intitle:hoehoe」「intext:hoehoe」 → タイトル、本文中に“hoehoe”を含むサイトを表示。
- ・ 「5+20*log(100)」 → 数式を入力して検索すると、計算結果を表示。
- ・ 「y=2+log(x)+sin(x)」「z=3*x-5*y+2」 → 入力した式の2/3次元グラフを表示。

→ 他にも「簡単に」単位換算、翻訳、日の出・日没時刻など様々な情報にアクセスできる。

21. インターネット時代の著作権・ライセンス

第8回目

注) 本担当者は、著作権をはじめとする法律の専門家ではない。
本内容は、くれぐれも参考程度に心中に留めておき、必要があれば、専門家に相談すること。

● よくある先生や研究者の勘違い

~~「商用目的で無く、教育や研究が目的であれば、著作権に抵触しない。」~~

~~→ だから、著作物に対して複製・改変などをやっても構わない~~

学校教員：「天文台が公開している天体画像を使って、教材を作りました」

(天文台のWEBページに「商用以外の研究・教育・普及活動には無許可で使ってOK」と記載)

→ 教材を、学校で児童相手に使ったら、評判が良かった！

→ 先生どうしの研究会で紹介・発表 → 違反

→ 専門機関で少年の更生目的で使用 → 違反

「先生同士の研究会」や「少年の更生目的の機関」は天文台のWEBページで示された目的・活動に合致しない。 → 事前に天文台に連絡を！

21. インターネット時代の著作権・ライセンス

第8回目

● 様々な検索エンジン・検索サイト

歌詞検索、画像検索、動画検索、ニュース検索、図書検索、論文検索などなど

ニュース・ブログ検索：ニュースやブログなどのウェブサイトの更新情報や配信を行うために、RSSというフォーマットが使用されている(複数規格が混在)。

図書検索：ISBN(International Standard Book Number = 国際標準図書番号)によるデータベース管理。グループ、出版社、書名、チェック・ディジットからなる。

CiNii (Citation Information by National Institute of Informatics = サイニイ)：
国立情報学研究所(=NII)が運営する学術論文や学術図書などの情報データベース。

→ 検索結果をどこまで自由に使用してよいのか？

● 様々な形での情報・データの提供

・ 電子図書 / 電子書籍：基本的に電磁的に記録されたデータ(ソフトウェア)を指すが、閲覧のために専用の端末(ハードウェア)が必要なことも多い。

・ 電子文書

・ 電子画像

・ 電子音声 / 電子動画

これらには常に、違法コピー、違法アップロード、違法ダウンロードなどの問題が付いてくる。

→ 入手した情報をどこまで自由に使用してよいのか？入手してよいのか？

● ライセンス (license)

許可、認可、承認、またはそれを示す書面・書類。他者が持つ特許や技術を使うための法的な許可。許可・承認を与える側を**ライセンサー (licenser)**、受ける側を**ライセンシー (licensee)**という。

例) NASAが公開している天体画像などは、字面通りに、商用以外の教育・研究目的であれば、**クレジットタイトル**明記の元で使用可と言われている。

クレジットタイトル: 書籍でいう「奥付」にあたるもの。
著作者の名前などのこと。

- 教育・研究目的の場合、まず、「**引用元を明記**」することが重要。
読者・閲覧者が容易に引用元に辿り付けるだけの情報を記す。
その意味で容易に変容するURLは微妙な情報(西浦私見)。
- 教育・研究目的であっても、過度な引用などについては、著作権などの侵害と見なされる。特に、他者の執筆・作成部分を、自分のオリジナルであるかのように記述することは重大な違反行為である。

● パブリック・ドメイン (public domain)

著作物や知的創作物に対して、知的財産権が発生していない、または、消滅している状態をこのように表現することがある。

ただし、国によって**著作権**や**意匠権**、**所有権**、**人格権**など様々な要素が複雑な関係にあるため、決して、無制限な状態ではないことに留意する必要がある。

- ・ **青空文庫**: 1997年に**富田倫生**によって開設された、著作権が消滅した文学作品を公開している**電子図書館**。
- ・ **ウィキソース**: 著作権が消滅した文書を電子化・公開するための計画で、アメリカのウィキメディア財団が運営。

● GFDL (GNU Free Documentation License)

● GPL (GNU General Public License)

著作権者が「**無断で複製可**」「**無断で改変可**」「**無断で頒布・販売可**、ただし、**頒布された、もしくは、購入した者に対して先の許可を与えること**」としたライセンス。GFDLは文書、GPLはソフトウェアに対するもの。ただし、各国の法(例えば日本)と整合しないこともあり、問題が無い訳ではない。

コピーレフト (copyleft): 著作権を維持したまま、二次的著作物も含めて全ての者が、著作物を利用・再配布・改変できるべき、という考え方。

本日の演習

第8回目

班・グループ毎に集まって、テーマとして与えられた天体やそれに関連した様々な情報に関して、検索し、情報を収集・報告せよ。

→ 班の代表者1名が、WEB CLASSを用いて、担当者まで検索した事柄を簡潔にまとめて、報告せよ。

(方法) WEBブラウザで「www.u-gakugei.ac.jp」より「学芸ポータル」へアクセス。

→ 「WEB CLASS」→ 「情報(月5)」→ 右下の「管理者」

→ ポータル/メッセージを通して、担当者にメッセージを送ることができる。

[報告する事柄]

- 1) 班名・グループ名(アルファベット)
 - 2) 代表者の学生番号と氏名
 - 3) 本日、参加したメンバーの学生番号と氏名
 - 4) テーマの天体名
 - 5) 今日、検索して分ったことを、10～20行程度で、箇条書きで報告する。
- 送信が終わったら、各班・グループ毎に終了してOK