

## インターネット利用の情報の流通

- 電子メール
- ファイル転送 (ftp)
- ネットニュース
- **World Wide Web (WWW)**

など

## World Wide Web (WWW)

1980年代末～90年代初頭に開発

元々は **CERN** (欧州原子核研究機構) に於いて

各国の大勢の研究者同士で

研究情報を共有するための仕組みとして

開発された。

**HTTP (HyperText Transfer Protocol)**

によってデータの送受信を行なう

## HyperText

データ (テキスト) 同士が**リンク**で繋がっている  
… ハイパーリンク (hyperlink)

**HTML (HyperText Markup Language)**  
などで記述される

ハイパーリンクを手軽に辿れるソフトウェア  
(ウェブブラウザ (web browser)) の開発が

**WWW** の爆発的な普及に繋がった。

## ウェブブラウザ (web browser)

- HTML を適切に表示
- ハイパーリンクを辿れる
- 必ずしも graphical ではない  
(テキストベースブラウザというものもある)

## ウェブブラウザ (web browser)

実際に手元に届いているデータは？

- [表示] → [ページのソース]
- [右クリック] → [ページのソースを表示]

HTML で書かれた内容を  
ウェブブラウザが適切に解釈して  
表示している

ブラウザによって表示が異なることがあり得る

本授業でも後で (多分最後の2回に)

HTML で Web 文書 (ウェブページ) を

作ってみます

## URL(Uniform Resource Locator)

Internet 上の資源 (の在処) を  
特定するための文字列

ハイパーリンクでは URL を指定

参考: URI (Uniform Resource Identifier)

## URL の例

`http://ccweb.cc.sophia.ac.jp/index.html`

- スキーム名 (`http:`)
- ホスト名
  - ★ ホスト名 (`ccweb`)
  - ★ ドメイン名 (`cc.sophia.ac.jp`)
- パス名
  - ★ ディレクトリ名
  - ★ ファイル名 (`index.html`)



`http://ccweb.cc.sophia.ac.jp/  
CenterNEWS/2008no3/PCVista.pdf`

- スキーム名 (`http:`)
- ホスト名
  - ★ ホスト名 (`ccweb`)
  - ★ ドメイン名 (`cc.sophia.ac.jp`)
- パス名
  - ★ ディレクトリ名 (`CenterNEWS/2008no3/`)
  - ★ ファイル名 (`PCVista.pdf`)

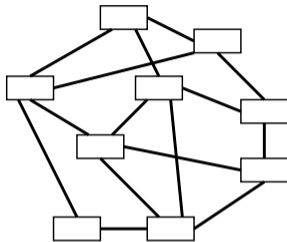
実際のネットワーク通信では、

各ネットワークインターフェイスに対し、

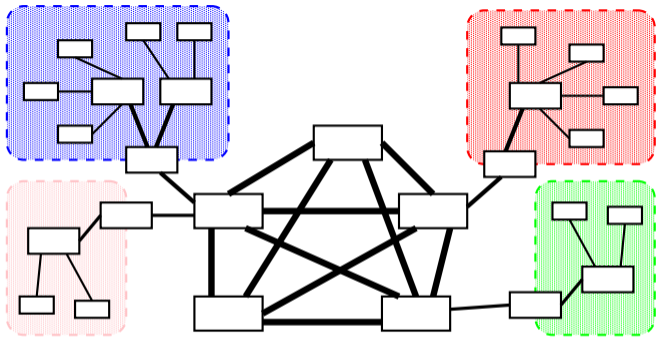
全世界 (Internet 全体) で  
一意的な番号を付けて、

その番号 (**IPaddress**) を用いて通信する。

Internet の特徴は分散型ネットワークであった。



- どこか壊れても機能する
- 誰も全貌を把握していない / 出来ない
- 全貌を把握していなくても機能する  
→ 刻々と変化するネットワーク構成に対応



ネットワーク上で遠く離れた所に  
アクセスするにはどうしたら良いか？  
… 経路制御 (**routing**)

## IPaddress と DNS(Domain Name Service)

Internet 上のネットワークインタフェースは  
IPaddress で一意的に特定されるが、  
いつでも番号で呼ぶのは扱い難い。

通常は

「ホスト名 (ホスト名 + ドメイン名)」  
で呼んでいる。

<http://ccweb.cc.sophia.ac.jp/index.html>

## IPaddress と DNS(Domain Name Service)

「IPaddress  $\longleftrightarrow$  ホスト名」の変換の仕組み

… **DNS(Domain Name Service)**

各ドメインの name server が  
自ドメイン内の変換表を持つ  
(分散管理)

## WWW 上の情報検索

WWW での情報共有機能は、  
情報検索の的確化・効率化によって  
一挙に世界を変えるに至った、  
と言っても過言ではない。

## 情報検索のコツ

そこにどんなデータ (文書) があるかを  
想像せよ。

キーワード検索なら、

自分が見付けたいデータ (文書) が  
どんなキーワードを含んでいそうかを  
想像せよ。

それが人の書いた文書なら、

どんな文章だろうか、と想像せよ。



ネットワークの向こうには人間がいる

→ ネットワークは人間社会である

→ 社会には守るべき規範 (慣習) がある

- ネットワーク倫理 (Nethics)
- ネットワークエチケット (Netiquette)

## 実習:

都道府県庁所在地の一覧を探したい。

WWW 上の検索で一覧データを見付けよ。

→ **Google** を使う

昔は次のように良く言われた。

「自分が見付けたいデータ(文書)が  
どんなキーワードを含んでいそうかを  
想像せよ」

→ 思い付く限りの県名・県庁所在地名を  
検索キーワードに並べて指定せよ

しかし、現在ではそこまでしなくても  
良い検索結果が得られることが多い

## 検索結果の順位付け方法

以前は主に、

そのページに含まれるキーワードの  
多寡・質で順位付け

- 検索で上位に来ることだけを狙った  
ページが増加 (宣伝業者など)
- 検索結果の品質の低下
- **WWW** のデータベースとしての品質の低下
- **Google** の登場

## Google の Pagerank™

WWW 上の全てのページの  
重要度 (**Pagerank**) を判定

基本原理:

「多くの良質なページから  
リンクされているページは、  
やはり良質なページである」

→ キーワードを含むページを  
重要度の順に提示

## Google の Pagerank™

計算法の原理:

リンクをランダムに辿りながら大勢で動き回る  
→ 人が集まるページが重要なページ

- 線型代数 (行列計算) の固有値問題
- 高速な数値計算法の利用

これにより高速かつ有効な検索結果が得られた

→ WWW そのものが巨大なデータベースに

## メディアとしての WWW の特性

- 誰でも手軽に参加できる
- 強力な検索機能

→ 少数者・新規参入者も  
多数者・既得権益者と同等

→ (相対的には) 少数者に力を与えるメディア

## レポート課題

次のような事柄について調べよ。

例:

- **ARPANET** 以来のインターネットの歴史
- インターネット上での情報の流通
- **Google** の **Pagerank** について

今後も授業内容に関連した課題例を追加予定

詳細は追ってお知らせします



インターネットとWWWの話は一旦ここまで

今日の後半からは、

## エディタによるテキスト作成

について

電子データは  
プレーンテキスト (plain text)

が基本

プレーンテキストで済むものは  
プレーンテキストで済ませる

- 誰でも (どんな環境でも) 確実に読める
- 容量が小さく、  
コンピュータ・ネットワークへの負荷が少ない
- 引用・検索・自動定型処理が出来るなど  
使い回しが効く

## エディタとワープロソフト

- エディタ (editor):  
純粋にテキストファイルを作る
- ワープロ (word processor):  
レイアウト・文字飾り等の情報を含んだ  
データを作る

→ 目的によって使い分けよ

## エディタの特徴

「テキストファイルを作る」  
ということに専念したソフトウェア

- 使い方は非常にシンプル
- 一般に軽快・高速
- テキストファイル作成のための機能が充実
  - ★ コピー・ペースト
  - ★ 検索・置換
- レイアウト・文字飾り等の機能はない

ここでは、

「エディタの使い方」を学ぶというよりも、

「エディタを使うこと」を学ぼう。

**“EmEditor”** を使う

## 課題:

都道府県庁所在地の一覧を

都道府県名:都道府県庁所在地名

(1行1都道府県)の形式の

テキストデータとして作成せよ。

ファイル名: kencho.txt

- WWW 上の検索によって  
一覧データを見つけて保存
- 保存したファイルをエディタで編集

## EmEditor の使い方

- **起動**: [スタート] → [プログラム]  
→ [Tools] → [EmEditor]
- **終了**: [Alt+Shift+x]  
または [ウィンドウ右上の x]
- **開く**: [Ctrl+o]  
または [ファイル] → [開く]  
(複数のファイルを同時に開いて編集できる)

## EmEditor の使い方

- 保存していない変更があると  
タブのファイル名に \* が付く
- 上書き保存: [Ctrl+s]  
または [ファイル] → [上書き保存]
- 名前を付けて保存:  
[ファイル] → [名前を付けて保存]
- 変更前の版は残して、  
変更後の版を別ファイル名で保存したい時  
開いて即「新ファイル名で保存」が賢明



## EmEditor の使い方

- **保存して閉じる: [Ctrl+e]**  
または [ファイル] → [保存して閉じる]
- **すべて保存して閉じる: [Ctrl+Shift+e]**  
または [ファイル]  
→ [すべて保存して閉じる]
- **終了: [Alt+Shift+x]**  
または [ウィンドウ右上の×]