Web API入門

中村覚 (東京大学情報基盤センター)

もくじ

- Web APIとは
 - o Web APIの事例
- Web APIの応用:マッシュアップ
 - マッシュアップ事例
- Web APIの課題

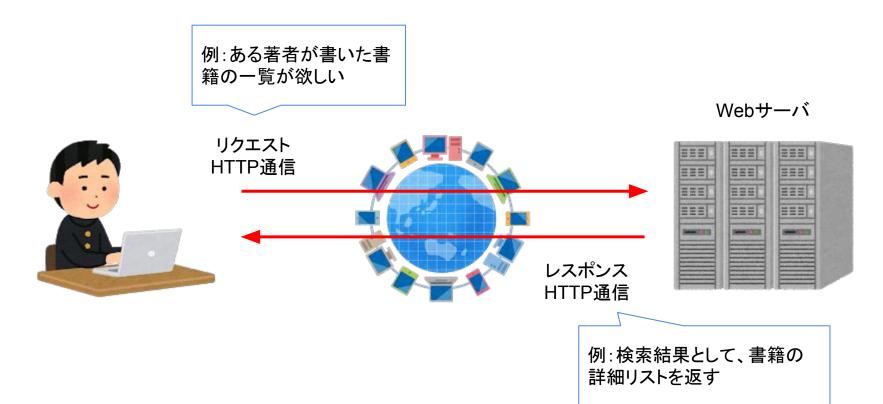
高久雅生. 第23回大図研オープンカレッジ「大学図書館員のための Web API入門」, https://www.slideshare.net/tmasao/web-api-49080729. CC BY-SA.

Web APIとは

Web APIとは

- ウェブ上で他のサービスを呼び出す方式、取り決め
 - 標準的なデータやサービスの提供を行う方法
- 具体的には:
 - ウェブを介して「要求(リクエスト)」を送り、サーバが要求された処理を実行し、「結果(レスポンス)」を返却する仕組み
 - リクエストはREST方式、レスポンスはXMLデータやJSONデータが一般的

Web APIとは(概念図)



Web APIの公開とは

- どのような形式の「要求(リクエスト)」か?
- どのような形式の「結果(レスポンス)」か?

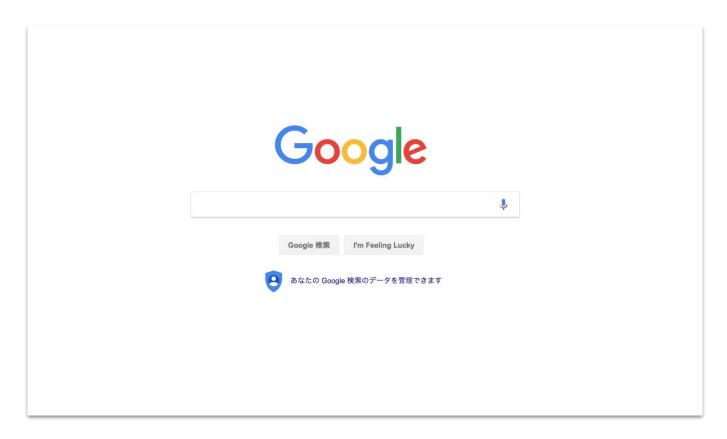
を公開すること

=> そうすることで利用方法がわかるようになる



Webサーバ

Webとは

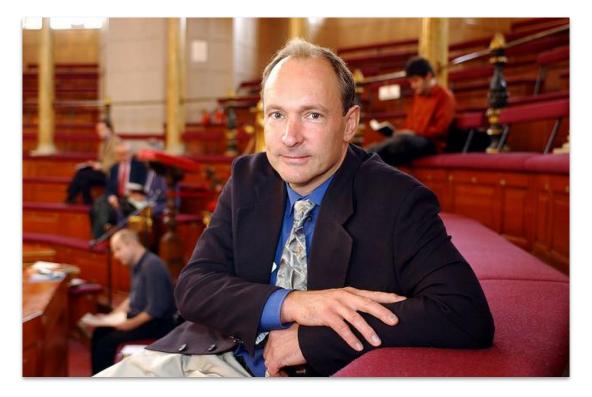


World Wide Web

- WWW (World Wide Web)
 - o ワールド・ワイド・ウェブ
 - ダブリュ・ダブリュ・ダブリュ
- または単にWeb(ウェブ)と称する
- [web](noun)
 - A network of silken thread spun especially by the larvae of various insects (as a tent caterpillar) and usually serving as a nest or shelter.



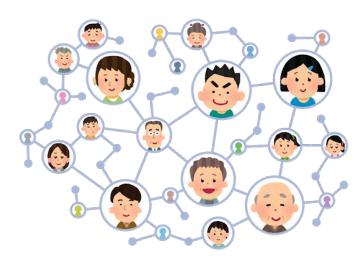
ティム・バーナード=リー(Tim Berbers-Lee)



https://www.flickr.com/photos/knightfoundation/2467553359/

Webの3要素

- HTTP, URL, HTMLの3つの枠組みを用いたWebの実現
- HTTPレイヤでのデータ転送とドキュメントフォーマット指定
- URI指定によるリンクを通じたハイパーメディアの実現



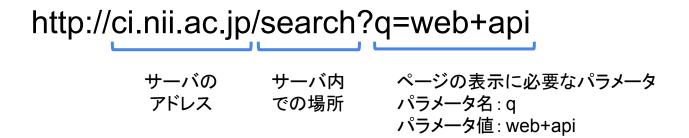
HTTPとは

- HyperText Transfer Protocol
 - 基本的なWebサーバとのやり取りを行う方式

GET /index.html HTTP/1.0

URLとは

- URL(Universal Resource Locator)
- Web上でのアドレスを指し示すアドレス(識別子)として機能する

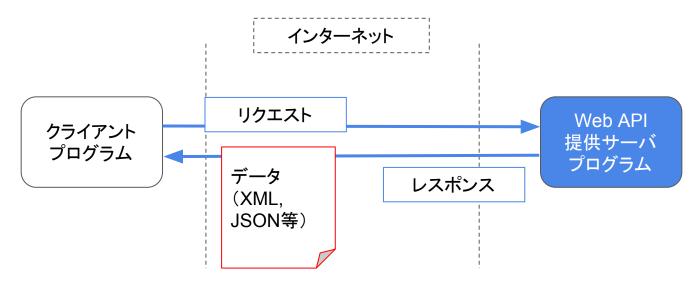


APIとは

- Application Programming Interface
- アプリケーションソフトウェアの開発にあたって使われる、標準的なモジュール群、 その利用法
 - API = 開発環境における標準
 - 典型的には、特定のOS環境のソフトウェアは共通する標準 APIを用いて開発される
 - Windows
 - Apple iOS

Web APIとは

- Web + API = Web API
- ウェブ上で他のサービスを呼び出す方式、取り決め
 - 標準的なデータやサービスの提供を行う方法



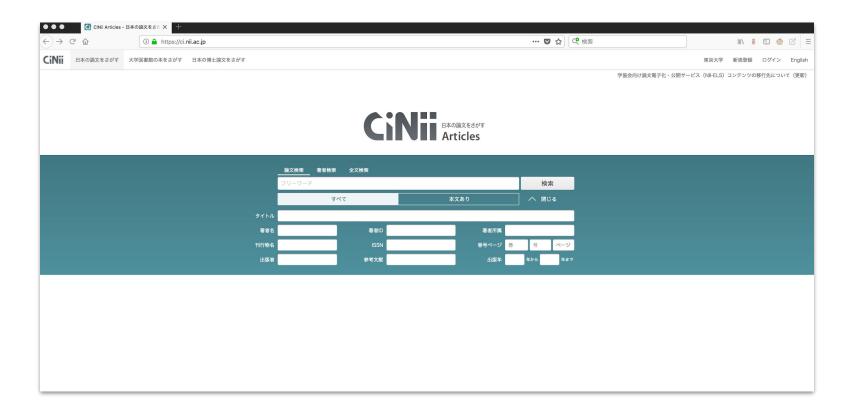
Web APIの歴史

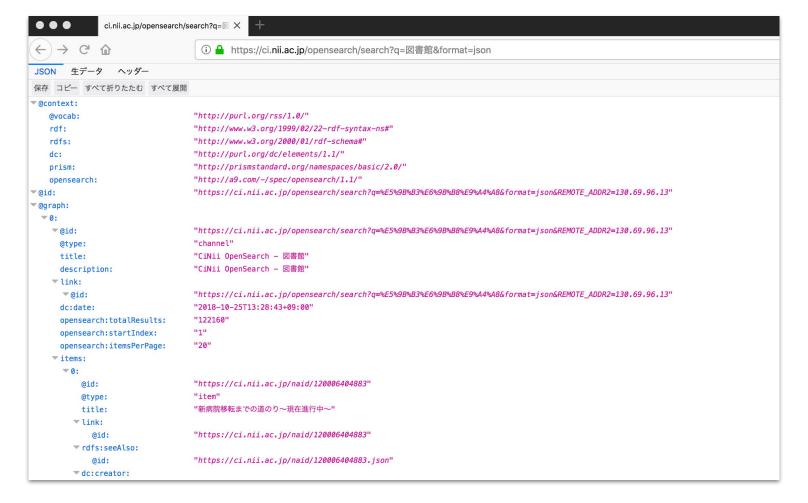
- インターネット普及前には...
 - そもそも、同一環境(計算機や OS)同士で通信する仕組みはある程度必要
 - 遠隔のサービスを呼び出す仕掛け
 - ウェブ以前には: RPC(Remote Procedure Call)
 - オープン環境においても、通信コマンドやデータ形式の標準化(分野ごとに)
- これらの方式を受け継いで、Web上でも行うための手段
 - SOAP方式
 - …一般的にはさほど普及しなかった
 - …(面倒くさいらしい)

Web APIの歴史

- Web API: 別名 Web Servicesとも呼ばれる
- 2002年頃から一般的に使われ始める
 - Amazon, Googleなどが代表例
- きっかけはWeb 2.0
 - 2005/2006年頃に爆発的に流行
 - 当時流行っていて先進的なサービスをまとめた用語
 - その技術要素のひとつが AJAXという手法
 - 前提として必要なデータのやり取りの方式 = Web API
 - Google Mapsがその代表格
 - つまり、Google Maps APIが前提となって、AJAXを実現させた

実際のWeb APIのやり取り





2つのやり取りの方式

SOAP方式

- 呼び出すパラメータ類を XML形式のデータにまとめ、それを受け渡す方式
- W3C(World Wide Web Consortium)が規格化
- 企業による初期のWeb API提供形態

REST方式

- 呼び出すパラメータ類は URL内のクエリとして発行し、データ結果のみを受け取る方式
- ブラウザからも簡単に呼び出したり、テストしたりできるため、今ではこちらが主流に

RESTとは

- Webの基本的なやりとり、コンテンツの取得(GET)、更新(POST)といったやりとりをまとめたお約束
- Web APIとの関連では:
 - コンテンツの取得はGETメソッドを使う
 - コンテンツの更新はPOSTメソッドを使う
 - といった、ブラウザとウェブサーバのやり取りの仕方を指す。
- RESTの範囲を外れる例:
 - データの取得にPOSTを使ったり、クッキーを前提としたページ遷移を必要とするサービス
- 要は、GETリクエストとパラメータによって、必要なデータを取得する仕組み

URLとWeb APIリクエストの関係

- Web APIにおいてはURLが基本的な単位となって、それに対する「操作」(処理)が やり取りを決める
 - URLに様々なパラメータを与える
- 例:

http://ci.nii.ac.jp/opensearch/search?

q=図書館&

format=atom&

sort=7

URLとパラメータ

- Web APIのサービス単位で異なる
- CiNii Articlesの場合
 - q=クエリ
 - format=データ形式
 - o sort=ソート方式
 - 詳細はヘルプを参照...

「CiNii Articles論文検索のOpenSearch」へルプより



Web APIの事例

Web APIの種類

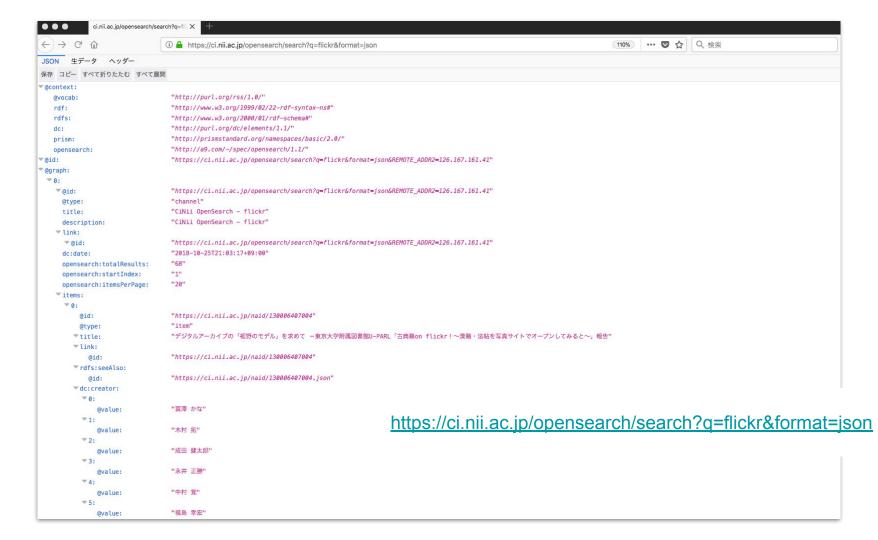
- 様々なサービスがWeb APIとして提供
 - 例えば、手元で使ったことがあるものだけでも、書籍検索、論文検索、画像検索、地図情報、音声 認識、顔認識、など
 - Web APIの紹介サイトProgrammable-Webには、600近いカテゴリが用意され、13,000を超える Web APIサービスが紹介されている
 - https://www.programmableweb.com/category

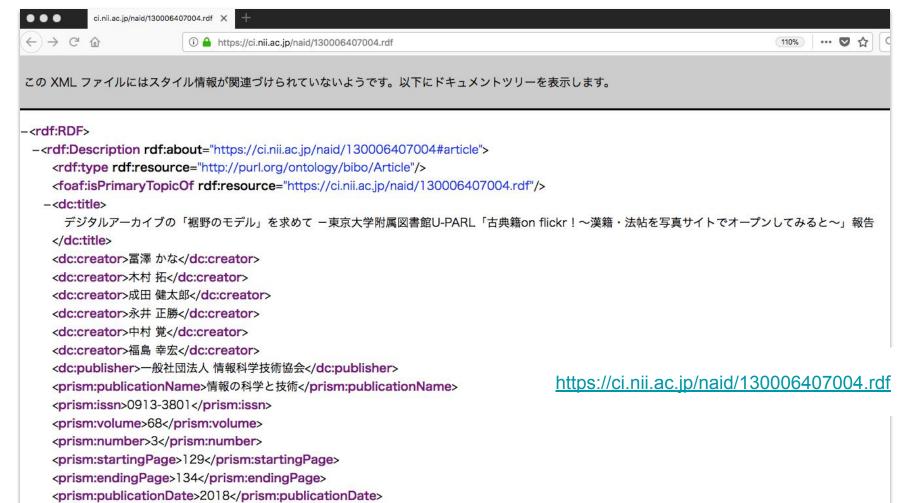
Web APIの事例紹介

- CiNii (Articles/Books)
- NDLサーチ
- カーリル
- (Google Maps)
- (Amazon)

CiNii Articles

- 国立情報学研究所が提供する日本の論文検索サービス
 - https://ci.nii.ac.jp/
- 検索: OpenSearch
 - https://ci.nii.ac.jp/opensearch/search?q=omeka&format=json
 - 検索条件をパラメータとしてやり取りする
 - 検索結果はAtom形式(XML)やJSON形式等を選択可能
- 詳細レコード: RDF
 - https://ci.nii.ac.jp/naid/130006407004.rdf
 - 詳細レコードのデータをRDF形式で返す

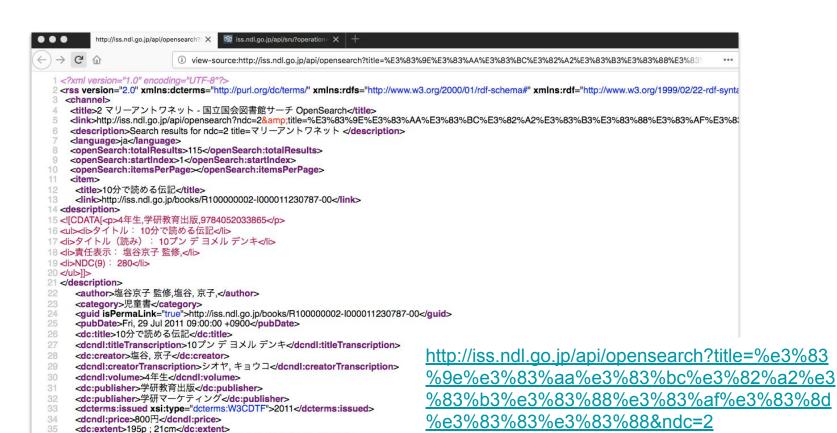




-<dc:description>

NDLサーチ

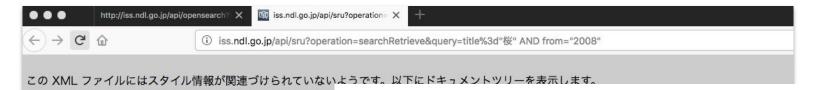
- 国立国会図書館が提供する検索サービス
 - http://iss.ndl.go.jp/
- 検索: OpenSearch、SRU/SRW
 - http://iss.ndl.go.jp/api/opensearch?title=%e3%83%9e%e3%83%aa%e3%83%bc%e3%82%a2
 %e3%83%b3%e3%83%88%e3%83%af%e3%83%8d%e3%83%83%e3%83%88&ndc=2
 - http://iss.ndl.go.jp/api/sru?operation=searchRetrieve&query=title%3d%22%e6%a1%9c%22%2
 OAND%20from=%222008%22
- データ取得系: OAI-PMH、SRU/SRW
 - DCNL(NDL版Dublin Core(XML))



<dc:identifier xsi:type="dcndl:ISBN">9784052033865</dc:identifier>

<dc:identifier xsi:type="dcndl:NDLBibID">000011230787</dc:identifier><dc:identifier xsi:type="dcndl:JPNO">21958058</dc:identifier><dc:identifier xsi:type="dcndl:NSMARCNO">123450400</dc:identifier>

36



-<searchRetrieveResponse>
 <version>1.2</version>
 <numberOfRecords>8713</numberOfRecords>
 <nextRecordPosition>201</nextRecordPosition>

http://iss.ndl.go.jp/api/sru?operation=searchRetrieve&query=title%3d%22%e6%a1%9c%22%20AND%20from=%222008%22

-<extraResponseData>

| cfacets | clst name="REPOSITORY_NO" > cint name="R100000002">4707</int > cint name="R100000001">2396</int > cint name="R000 name="R10000004">220</int > cint name="R100000049">214</int > cint name="R000000024">96</int > cint name="R100000073">5 name="R10000004">26</int > cint name="R100000067">13</int > cint name="R100000094">96</int > cint name="R100000039">66</in name="0">33</int > cint name="1">42</int > cint name="1">42</int > cint name="1">42</int > cint name="3">179</int > cint name="1915">1</int > cint name="4">78</int > cint name="1916">1</int > cint name="1915">1</int > cint name="1916">1</int > cint name="1990">1</int > cint name="2001">1</int > cint name="2000">1</int > cint name="2000">3</int > cint name="2010">3</int > ci

カーリル

- カーリル社が提供する図書館横断検索サービス
 - 独自のWeb API
- 検索・所蔵検索
 - http://api.calil.jp/check?isbn=4877511172&systemid=Tokyo_Ome&format=json
- データ形式
 - JSON形式(独自)

カーリル

```
callback({
  "session": "ae6318869192394d0d2a62164ff4a1b3",
  "books": {
   "4877511172": {
     "Tokyo_Ome": {
       "status": "Cache",
       "reserveurl": "https://www.lics-saas.nexs-service.jp/ome/webopac/searchdetail.do?biblioid=44433",
       "libkey": {
         "沢井": "貸出可",
         "小會木": "貸出可",
         "中央": "貸出可",
         "今井": "貸出可",
         "大門": "貸出可"
 },
  "continue": 0
});
```

そのほかのWeb API事例

- 興味深い事例としてはAmazon Web ServicesやTwitterなどのように主要な商用 サービスにおけるWeb APIの展開例がある
- これらでは、自身のサービスをWeb API経由で取得、操作できる
 - Twitterクライアントの操作:ツイートの投稿、編集、リプライなど(BOT)
 - Amazonにおけるインフラ系ツールを Web APIから操作:サーバインフラの増強、ツールの追加

マッシュアップの事例

Web APIの応用例(マッシュアップ)

- Web APIの典型的な応用例(アプリケーション)
- マッシュアップとは?(Mashup)
 - 複数の機能を組み合わせて、可視化したり、異なる機能同士を有機的につなげて利用する
 - Web APIの応用アプリケーションは、Web API機能をほかの何かと組みわせて扱うことが多いこと から「マッシュアップ」と呼ばれる

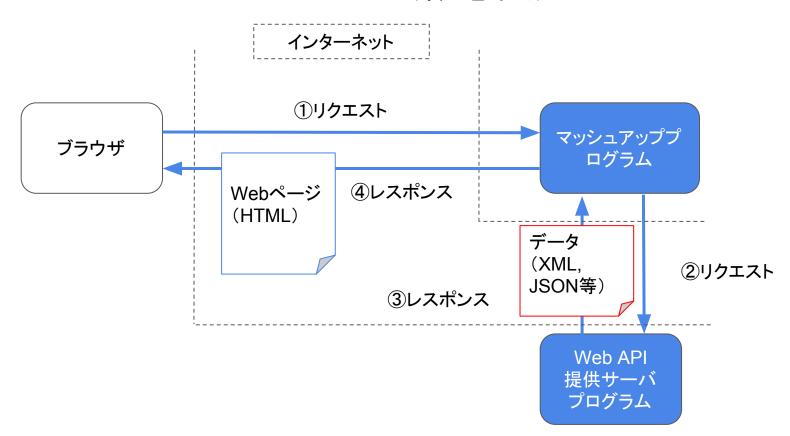
マッシュアップの種類

- (付加価値サービス)
- マッシュアップ (Mashup)
 - サーバサイド型マッシュアップ
 - Ajax型マッシュアップ
 - クライアントサイド
 - データ集約型マッシュアップ
 - ハーベスト型

サーバサイド型マッシュアップ

- 最も古典的なタイプの応用例
- 様々な類型がある
 - ひとつだけでなく、複数のサービスを組み合わせる
 - Ajax型との組み合わせも

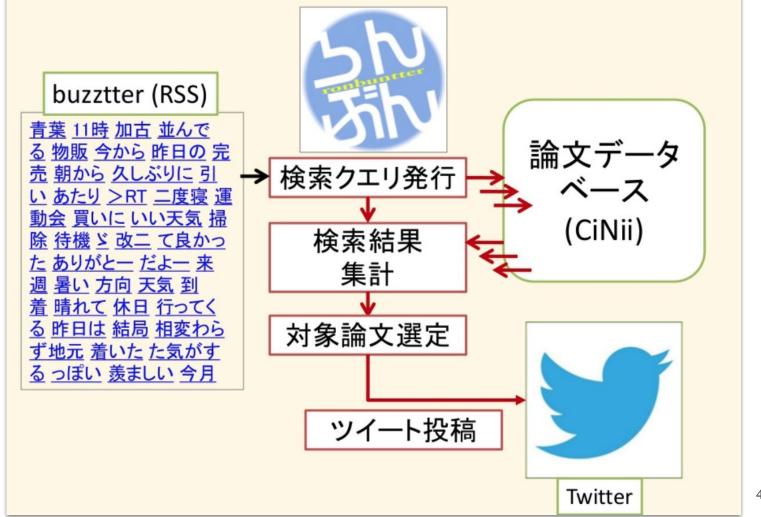
サーバサイド型マッシュアップ(概念図)



論文った一



処理の 流れ

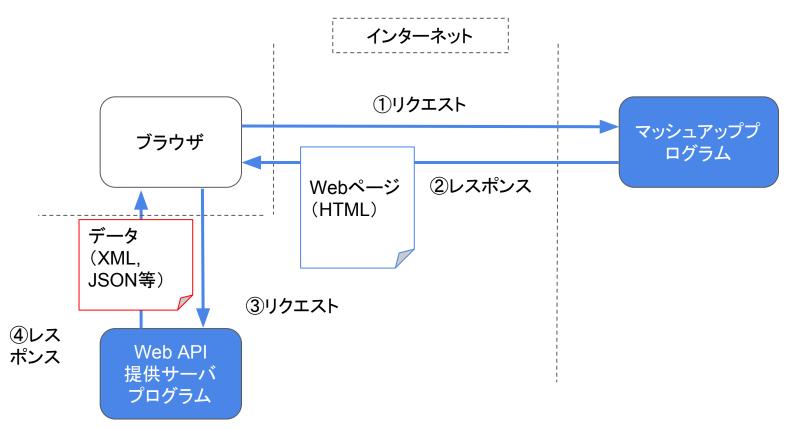


Ajax型マッシュアップ

● ブラウザ上で実行されるJavaScriptを主体としたアプリケーション

 サーバ側でまとめて実行された結果が返ってくるのではなく、個別のWeb APIへの リクエストがひとつずつ実行された分ごとに返ってくるため、見た目の実行時間を早 める利点がある

Ajax型マッシュアップ(概念図)





データ集約型マッシュアップ

- Web APIを通じてデータ収集し活用できるようにしたマッシュアップサービス
- RSSリーダーなどはこの形態が多い
- 機関リポジトリなどにおけるメタデータハーベスティング
 - 複数のRSSサイト => RSSフィード閲覧サービス
 - 複数の機関リポジトリ => 論文の横断検索サービス



Web APIとデータ形式

Web APIと標準化

- 様々な標準規格化への欲望
- 多様なサービス vs ニーズ vs 標準化
- 標準化の難しさ vs 標準の乱立
- 検索・データ取得用Web API標準の例
 - SRU, OpenSearch, OAI-PMH, OpenURL, SPARQL, ...

データ形式と標準化

- Web APIにおいては、基本的には、リクエストとレスポンスの2つの標準化がある
- レスポンス:機械可読なデータ形式
- データ形式:
 - o XML
 - JSON
 - o TSV

XML(Extensible Markup Language)とは

- タグ要素とその内容テキストを指定し、機械的に読み取り可能とする形式
 - <親要素>
 - <要素 属性="値"></要素>
 - </親要素>
- なぜXMLか
 - 階層関係(木構造が入れ子になったデータ)を扱いやすい
 - 多言語対応(Unicode):日本語はじめ、多くの言語に対応
 - ほとんどのプログラミング環境に対応するライブラリが完備されていて、簡単に利用可能
- XMLの欠点
 - データ量が増える(タグ要素を入れ込む分)
- XMLの利点
 - メタ言語:様々な形式の新しい種類の言語を規定できる
 - XHTML, HTML5, SVG, RSS, Atom, ...

JSON(JavaScript Object Notation)とは

- ブラウザ上で動作するJavaScript言語が標準対応する形式
 - XMLの複雑でデータ量が多くなる欠点を、タグを使わず、より簡易に書けるようにしたデータ形式

```
(著者": ["高久雅生"],"タイトル": "Web APIの過去・現在・未来"
```

Web APIの課題

Web APIのご利益

- サービスの「プラットフォーム化」
 - 第三者にサービス提供する方式を標準化し、利用を促進できる
 - 近年の多デバイス環境下(PC、タブレット、スマートフォン)では、自サービスのプラットフォーム化が必然になってしまうという側面も ...
- サービスの多様化
 - Web APIにより、通常の利用形態を超えた利用を作り出すことができる

Web APIの寿命

- Web APIの永続性はそれほど長くないことに注意が必要
 - GoogleサーチAPI(2010年に終了)
 - Yahoo! Web検索API(2013年に終了)
 - Twitter(2013年に大幅な更新)
- 必然的に、マッシュアップ側の寿命に影響を与える!
- 利用する場合には
 - ある程度、代替サービスを念頭におく必要がある
 - Youtube vs 二コ二コ動画
- 提供する場合には
 - 例えば、検索サービスを単純に更新するだけであっても、 Web APIにおいては大きな影響を与える こととなる
 - 最初に公開する際のリクエスト形式、データモデルの設計

Web APIの制限、認証

- Web API利用は機械可読(プログラムによるやり取り)が前提であるため、過負荷 状態になる危険もある
 - 人手アクセスの数倍から数十倍
- 同時利用に制限をかけることが一般的
 - 1時間に1,000アクセスまで
 - 1時間に3,000アクセスまで
 - 1日に100,000アクセスまで、等
- 加えて、アクセス元のユーザを識別するために認証をつけることもある
 - o Twitterなど

Web APIの提供にあたって

- URLとデータモデルの決定が非常に重要
 - どのURLにリクエストを送ると、どんなデータが戻ってくるか?
 - 「クールなURIは変わらない」
 - cf. https://www.kanzaki.com/docs/Style/URI
- Web API vs データファイルの提供
 - 数万件くらいの単純なデータなら、データファイル全件をダウンロードできるようにしておくだけでも 十分
 - 逆に、データ量が膨大だったり、複雑な条件指定を行う必要があるものでは Web API化する 恩恵も
 - 「CSV最強…」by 吉本龍司
 - 「Excelで扱えない規模からがビッグデータ ...」by 喜連川勝

ご静聴ありがとうございました。