

WARP で収集したウェブサイトが 1 万件を突破しました

平成 14 年から日本国内のウェブサイトを保存し続けている国立国会図書館の「インターネット資料収集保存事業（WARP：Web Archiving Project）」にて、この 13 年間に収集したウェブサイトの数（タイトル数）が、平成 27 年 10 月をもって 1 万件を突破しました。

平成 27 年 10 月末時点で保存しているファイル数は約 36 億ファイル、データ容量は約 630 テラバイトに及びます。

ウェブサイトに掲載された情報は、頻繁に更新・削除されます。国立国会図書館はこれからも、過去の情報が失われないように、また、いつまでも利用できるように、ウェブサイトを収集・保存して未来に伝えていきます。



<http://warp.da.ndl.go.jp/>

○ 平成 27 年 10 月末時点の件数／容量

タイトル数： 10,069 件
 保存件数（収集回数）： 96,545 件
 ファイル数： 約 36 億ファイル
 累積データ量： 632.9 テラバイト

■ 報道機関の方の問合せ先：
 国立国会図書館 総務部総務課広報係
 TEL：03-3506-5103（直通）

<WARPの主な収集対象>

- 国の機関、地方公共団体等のウェブサイト（法律に基づいて収集：約5,400タイトル）
国立国会図書館法に基づき、国の機関や地方公共団体、独立行政法人、国公立大学のウェブサイトを網羅的に収集しています。
- 民間のウェブサイト（許諾を得て収集：約4,600タイトル）
公益法人、私立大学、政党、国際的・文化的イベント、東日本大震災に関するウェブサイト、電子雑誌などを主な対象として、発信者の許諾が得られたものを収集しています。

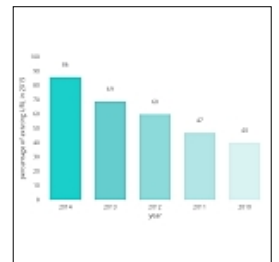
<WARPで保存しているウェブサイトの例>

- 消えた市町村サイト
平成の大合併で消失した市町村のウェブサイトを保存しています。
http://warp.da.ndl.go.jp/contents/recommend/collection/disap_localgov.html
- 東日本大震災ウェブサイトコレクション
国の機関、被災自治体のウェブサイトや、NPOやNGO、ボランティア団体、企業などによる支援活動のウェブサイトを保存しています。
<http://warp.da.ndl.go.jp/contents/recommend/collection/jpearthquake.html>

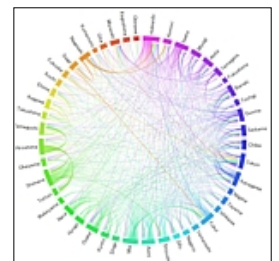


<WARPでわかるトレンド>

- 国の機関のウェブサイトは、5年で6割がなくなる（残存率調査）
収集した国の機関サイトを分析した結果、5年前（2010年）のURLは60%がなくなっていることがわかりました。
<http://warp.da.ndl.go.jp/contents/recommend/collection/linkrot.html>



- 震災後に福島県、宮城県へのリンクが増加（リンクの可視化）
都道府県ウェブサイトのリンク関係を図で示しています。ウェブ上における都道府県どうしの結びつきがわかります。
<http://warp.da.ndl.go.jp/contents/recommend/collection/preflink.html>



- 最も多い国土交通省は41テラバイト（保存サイトの容量）
WARPで保存した1万件のウェブサイトを、容量に応じて円で表しました。どのようなサイトがどのくらい保存されているのか、一目でわかります。
<http://warp.da.ndl.go.jp/contents/recommend/collection/bubble.html>

