

ウェブアーカイブとDOIの接点？

－引用における活用を視点に－

国立国会図書館 前田直俊

2017年3月27日

1. WARPについて
2. 引用における活用
3. DOIとの接点？

1. WARPについて

WARP



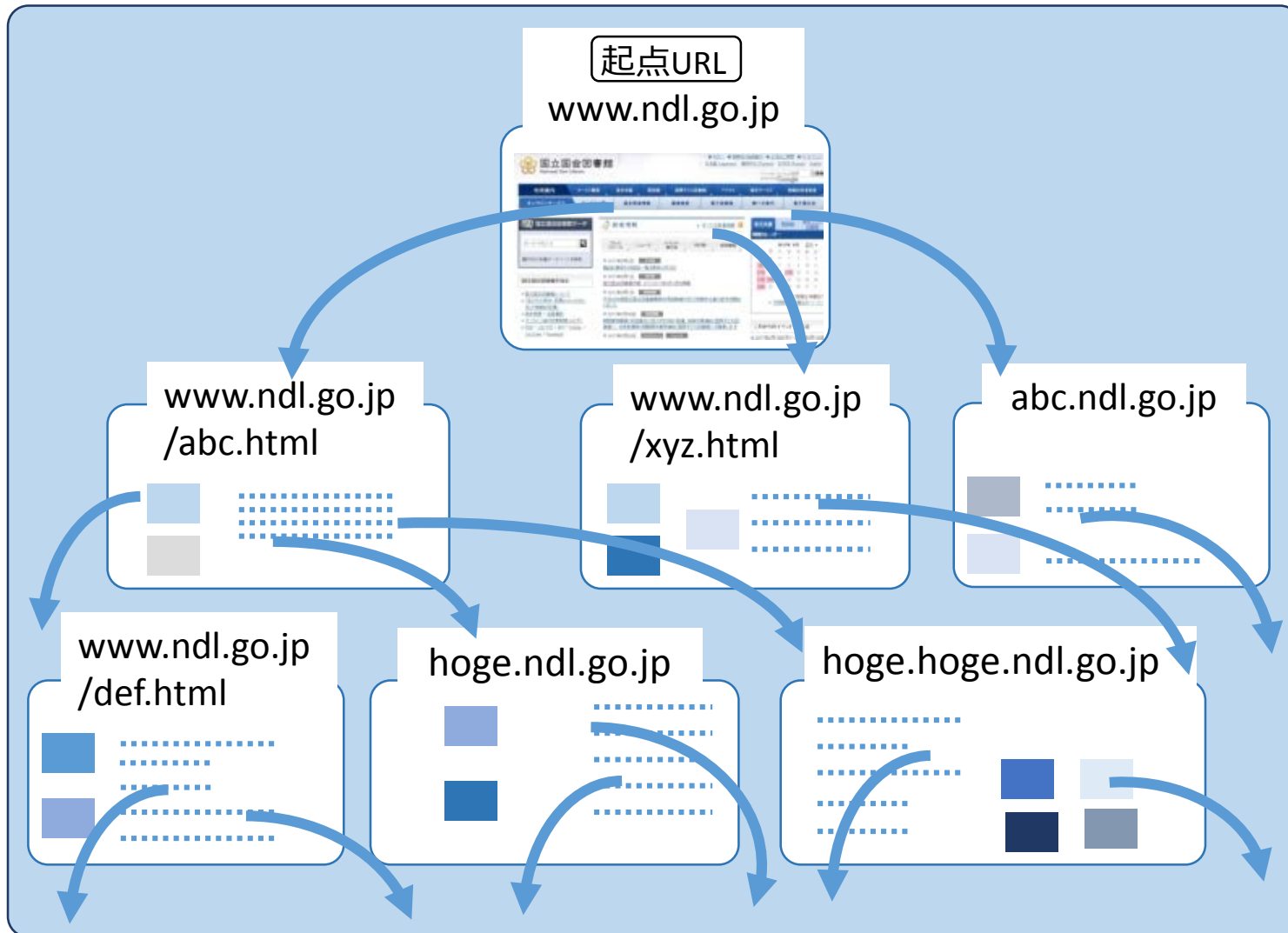
- 国内サイトを対象としたウェブアーカイブ

- 2002年に開始

- 2010年から公的機関サイトを網羅的に収集

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

サイトをまるごと収集



クローラによる収集

- ① 起点URLにあるファイルを複製
- ② リンクを解析してページ遷移
- ③ リンク先のページでも同じようにファイル複製、リンク解析
- ④ 指定した範囲の全てのファイルを複製するまで繰り返す
- ⑤ ファイルをひとまとめにして保存



ご覧いただいているのは国立国会図書館が保存した**2016年4月1日**時点のページです。このページに掲載されている情報は過去のものであり、最新のものと異なる場合がありますのでご注意ください。収集時のURLは <http://www.meti.go.jp/> (外部サイト) ですが、このURLは既に存在しない場合や異なるサイトになっている場合があります。
※このページの著作権について

ヘルプ

保存日: 2016年4月1日

本文へ よくあるご質問 サイトマップ

文字サイズ変更 小 **中** 大



経済産業省
Ministry of Economy, Trade and Industry

サイト内検索未入力のまま送信しても検索結果は表示されませんのでご注意ください

検索 拡張検索

ホーム 経済産業省について お知らせ 政策について 統計 申請・お問合せ English

**電力の小売
全面自由化
スタート!**

2016年4月1日



**あなたに合った
電気を選ぶ
時代へ。**

詳細は資源エネルギー庁HPへ >

電力の小売 全面自由化

フォトギャラリー (3月19日~20日)

経産省 広報誌 METI Journal

平成28年度 予算等

東日本大震災 関連情報はこちら

- 動画で知る 福島の実状
- 被災地の復旧・復興に向けて
- 東京電力福島原子力発電所事故の収束および今後の廃止措置に向けて
- 福島産業復興支援

新着情報RSS配信

twitter

facebook

会見・動画配信

経済産業省 (2016年4月1日)

<http://warp.da.ndl.go.jp/info:ndljp/pid/9968513/www.meti.go.jp/>



ご覧いただいているのは国立国会図書館が保存した2016年10月1日時点のページです。このページに掲載されている情報は過去のものであり、最新のものと異なる場合がありますのでご注意ください。収集時のURLは <http://www.data.jma.go.jp/fcd/yoho/hibiten/> (外部サイト) ですが、このURLは既に存在しない場合や異なるサイトになっている場合があります。

※このページの著作権について

ヘルプ

保存日: 2016年10月1日

■ 本文へ ■ ENGLISH ■ ご意見・ご感想 ■ サイトマップ

キーワードを入力し検索ボタンを押下ください。

検索

ホーム 防災情報 各種データ・資料 知識・解説 気象庁について 案内・申請

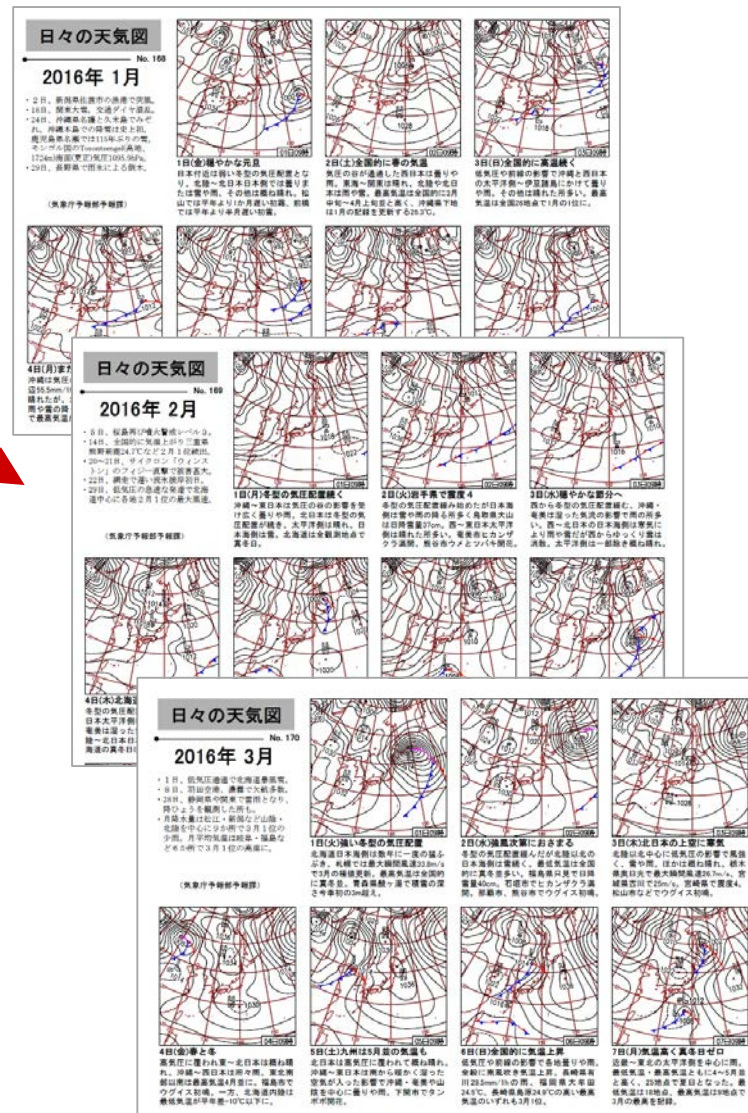
ホーム > 各種データ・資料 > 日々の天気図

日々の天気図

最新の「日々の天気図」は翌月の25日頃に掲載します。

「日々の天気図」をご覧になりたい月をクリックしてください。ファイルはすべてPDF形式です。

年	標準版				拡大版			
平成28年 (2016年)	1月[1.3MB]	2月[1.3MB]	3月[1.3MB]	4月[1.3MB]	1月[1.4MB]	2月[1.4MB]	3月[1.4MB]	4月[1.4MB]
	5月[1.3MB]	6月[1.3MB]	7月[1.3MB]	8月[1.3MB]	5月[1.4MB]	6月[1.4MB]	7月[1.4MB]	8月[1.4MB]
平成27年 (2015年)	1月[1.3MB]	2月[1.1MB]	3月[1.3MB]	4月[1.3MB]	1月[1.4MB]	2月[1.2MB]	3月[1.4MB]	4月[1.4MB]
	5月[1.4MB]	6月[1.3MB]	7月[1.4MB]	8月[1.4MB]	9月[1.5MB]	10月[1.5MB]	11月[1.4MB]	12月[1.4MB]
平成26年 (2014年)	1月[1.2MB]	2月[1.1MB]	3月[1.2MB]	4月[1.2MB]	1月[1.3MB]	2月[1.2MB]	3月[1.3MB]	4月[1.3MB]
	5月[1.3MB]	6月[1.2MB]	7月[1.2MB]	8月[1.3MB]	9月[1.3MB]	10月[1.3MB]	11月[1.3MB]	12月[1.4MB]
平成25年 (2013年)	1月[1.2MB]	2月[1.1MB]	3月[1.2MB]	4月[1.2MB]	1月[1.3MB]	2月[1.2MB]	3月[1.3MB]	4月[1.3MB]
	5月[1.3MB]	6月[1.2MB]	7月[1.2MB]	8月[1.2MB]	9月[1.3MB]	10月[1.4MB]	11月[1.3MB]	12月[1.3MB]
平成24年 (2012年)	1月[1.4MB]	2月[1.3MB]	3月[1.5MB]	4月[1.4MB]	1月[1.4MB]	2月[1.4MB]	3月[1.5MB]	4月[1.4MB]
	5月[1.3MB]	6月[1.2MB]	7月[1.2MB]	8月[1.2MB]	9月[1.2MB]	10月[1.3MB]	11月[1.3MB]	12月[1.3MB]
平成23年 (2011年)	1月[1.4MB]	2月[1.3MB]	3月[1.4MB]	4月[1.4MB]	1月[1.4MB]	2月[1.4MB]	3月[1.5MB]	4月[1.5MB]
	5月[1.6MB]	6月[1.5MB]	7月[1.5MB]	8月[1.5MB]	9月[1.7MB]	10月[1.5MB]	11月[1.4MB]	12月[1.5MB]
平成22年	1月[1.4MB]	2月[1.4MB]	3月[1.5MB]	4月[1.5MB]	1月[1.5MB]	2月[1.5MB]	3月[1.6MB]	4月[1.6MB]
	5月[1.5MB]	6月[1.5MB]	7月[1.5MB]	8月[1.5MB]	9月[1.5MB]	10月[1.5MB]	11月[1.5MB]	12月[1.5MB]



気象庁 (2016年10月1日)

<http://warp.da.ndl.go.jp/info:ndljp/pid/10202790/www.data.jma.go.jp/fcd/yoho/hibiten/>

WARP Web Archiving Project

ご覧いただいているのは国立国会図書館が保存した**2016年10月5日**時点のページです。このページに掲載されている情報は過去のものであり、最新のものと異なる場合がありますのでご注意ください。収集時のURLは <http://db.cger.nies.go.jp/DL/10.17595/20160901.001.html> (外部サイト) ですが、このURLは既に存在しない場合や異なるサイトになっている場合があります。

※このページの著作権について

ヘルプ

保存日: 2016年10月5日

地球環境データベース

国立研究開発法人国立環境研究所 地球環境研究センター

ホーム データベース 速報値 解析支援 関連データ データ検索 English

波照間ステーションにおける大気 CO₂ 濃度の連続観測データセット

DOI: 10.17595/20160901.001

一括ダウンロード

- HAT.CO2.Hourly.ver1.1.zip (md5: b4b249da926afe7ed838994b2a6dcea9)

個別ダウンロード

- HAT.CO2.Hourly.1993.ver1.1.txt (md5: 85676cafe52e950e0377b87abb390fab)
- HAT.CO2.Hourly.1994.ver1.1.txt (md5: 0a7b4bbe0dea02e3361d4d8d7ef61019)
- HAT.CO2.Hourly.1995.ver1.1.txt (md5: 9138de1a63a91b42fc29a6a34cd468bd)
- HAT.CO2.Hourly.1996.ver1.1.txt (md5: e41372dfbd72fc284083e94c3b8a15a8)
- HAT.CO2.Hourly.1997.ver1.1.txt (md5: 0aa047f4e27672a1539d3bf3572094a1)
- HAT.CO2.Hourly.1998.ver1.1.txt (md5: 3aad2d6ae443d185480ae058067aa9c2)
- HAT.CO2.Hourly.1999.ver1.1.txt (md5: ce507766b299b4387d80a090b4283424)
- HAT.CO2.Hourly.2000.ver1.1.txt (md5: d19b04483b38030c7d3ff0bb9b04aa0d)
- HAT.CO2.Hourly.2001.ver1.1.txt (md5: 952ee56ee790a98ec2402b7127187aaf)
- HAT.CO2.Hourly.2002.ver1.1.txt (md5: 63f4cdcc02eea9e6c69a07b9f165e81d)
- HAT.CO2.Hourly.2003.ver1.1.txt (md5: 536c9213ecf3ea82c1a68489fb23ee3d)
- HAT.CO2.Hourly.2004.ver1.1.txt (md5: 71db6ab5bf0c708f6bba3651f2613412)
- HAT.CO2.Hourly.2005.ver1.1.txt (md5: 823e54b1e73c048ce7c6c3a0620cdc11)
- HAT.CO2.Hourly.2006.ver1.1.txt (md5: 7be1c55997849160ffdcddab9ca85f50)
- HAT.CO2.Hourly.2007.ver1.1.txt (md5: 468fc5a4d25060ef04bb8aa668570400)
- HAT.CO2.Hourly.2008.ver1.1.txt (md5: ad4061e7bb7c445f23d4d4add33a1619)
- HAT.CO2.Hourly.2009.ver1.1.txt (md5: 6649bd6941c51164b0bb3797ec501dcc)
- HAT.CO2.Hourly.2010.ver1.1.txt (md5: ada37d63e4a2aec362b5936eef6f6f42)

9. Advisory service:
 - If you need scientific advice or expert opinion regarding the contents or scientific validity of this data set or Products derived from it, we can provide an advisory service based on an individual contract, different from the above one.
 - Depending on the extent of the help required, we can offer collaboration and/or supervision of the work based on the present data set. If you want to use our advisory service, contact the Data Provider.

TIME	YEAR	MONTH	DAY	HOUR	CO ₂	FLAG
0	1993	1	1	0	999.99	010
1	1993	1	1	1	999.99	010
2	1993	1	1	2	999.99	010
3	1993	1	1	3	999.99	010
4	1993	1	1	4	999.99	010
5	1993	1	1	5	999.99	010
6	1993	1	1	6	999.99	010
7	1993	1	1	7	999.99	010
8	1993	1	1	8	999.99	010
9	1993	1	1	9	999.99	010
10	1993	1	1	10	999.99	010
11	1993	1	1	11	999.99	010
12	1993	1	1	12	999.99	010
13	1993	1	1	13	999.99	010
14	1993	1	1	14	999.99	010
15	1993	1	1	15	999.99	010
16	1993	1	1	16	999.99	010
17	1993	1	1	17	999.99	010
18	1993	1	1	18	999.99	010
19	1993	1	1	19	999.99	010
20	1993	1	1	20	999.99	010
21	1993	1	1	21	999.99	010
22	1993	1	1	22	999.99	010
23	1993	1	1	23	999.99	010
24	1993	1	2	0	999.99	010
25	1993	1	2	1	999.99	010
26	1993	1	2	2	999.99	010
27	1993	1	2	3	999.99	010
28	1993	1	2	4	999.99	010
29	1993	1	2	5	999.99	010
30	1993	1	2	6	999.99	010
31	1993	1	2	7	999.99	010
32	1993	1	2	8	999.99	010
33	1993	1	2	9	999.99	010
34	1993	1	2	10	999.99	010
35	1993	1	2	11	999.99	010
36	1993	1	2	12	999.99	010
37	1993	1	2	13	999.99	010
38	1993	1	2	14	999.99	010
39	1993	1	2	15	999.99	010
40	1993	1	2	16	999.99	010
41	1993	1	2	17	999.99	010
42	1993	1	2	18	999.99	010
43	1993	1	2	19	999.99	010
44	1993	1	2	20	999.99	010
45	1993	1	2	21	999.99	010
46	1993	1	2	22	999.99	010
47	1993	1	2	23	999.99	010
48	1993	1	3	0	999.99	010
49	1993	1	3	1	999.99	010
50	1993	1	3	2	999.99	010
51	1993	1	3	3	999.99	010
52	1993	1	3	4	999.99	010
53	1993	1	3	5	999.99	010
54	1993	1	3	6	999.99	010
55	1993	1	3	7	999.99	010
56	1993	1	3	8	999.99	010
57	1993	1	3	9	999.99	010

国立環境研究所 (2016年10月5日)

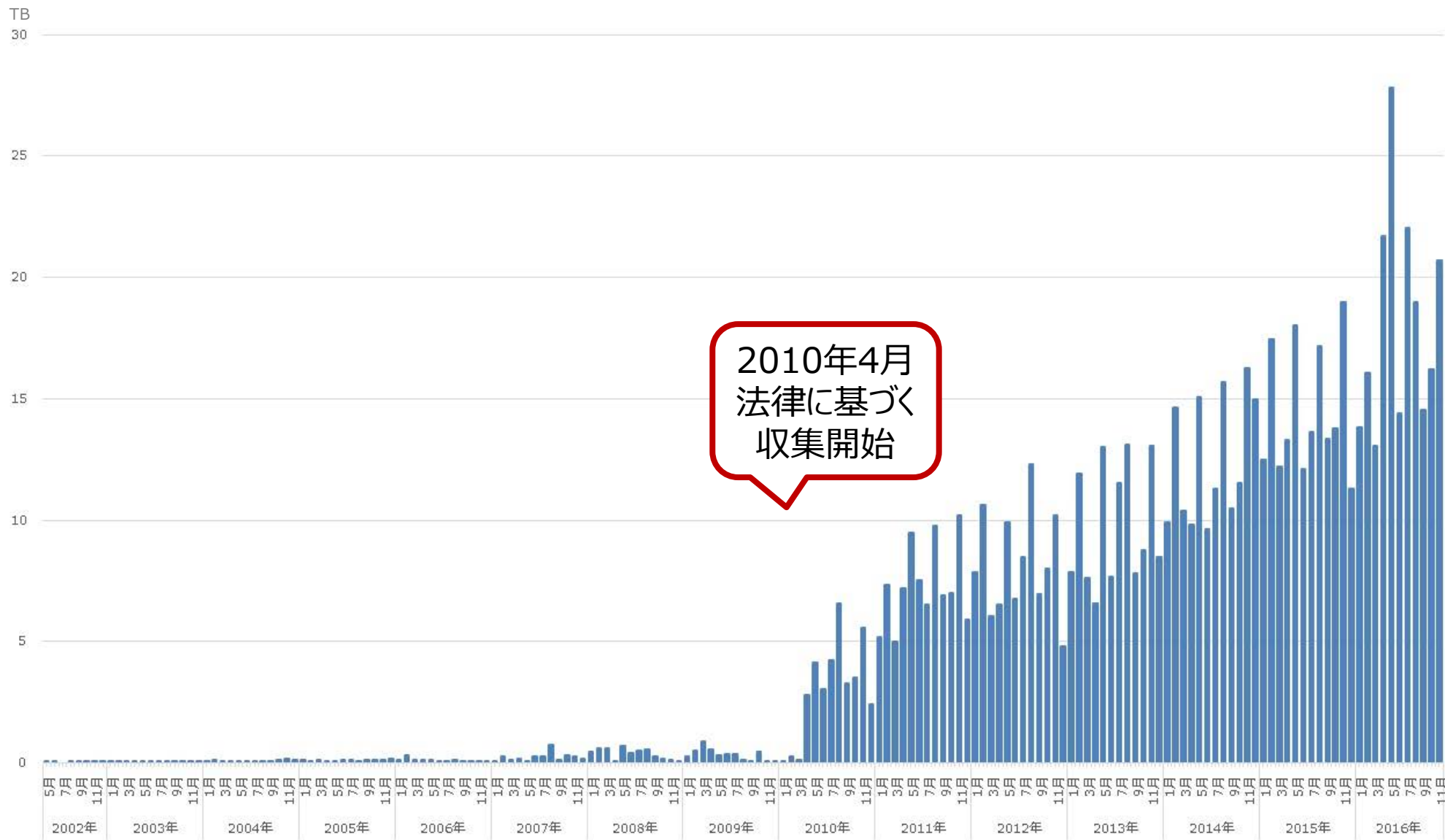
<http://warp.da.ndl.go.jp/info:ndljp/pid/10202005/db.cger.nies.go.jp/DL/10.17595/20160901.001.html>

集めているウェブサイト

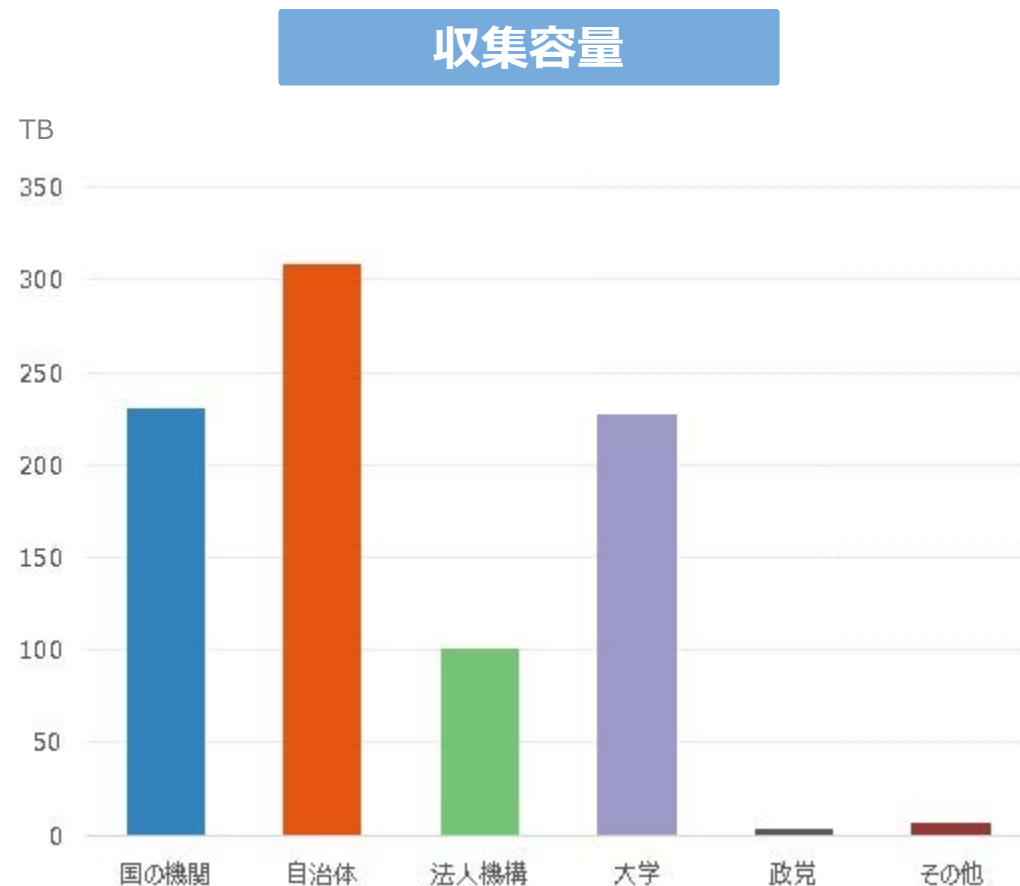
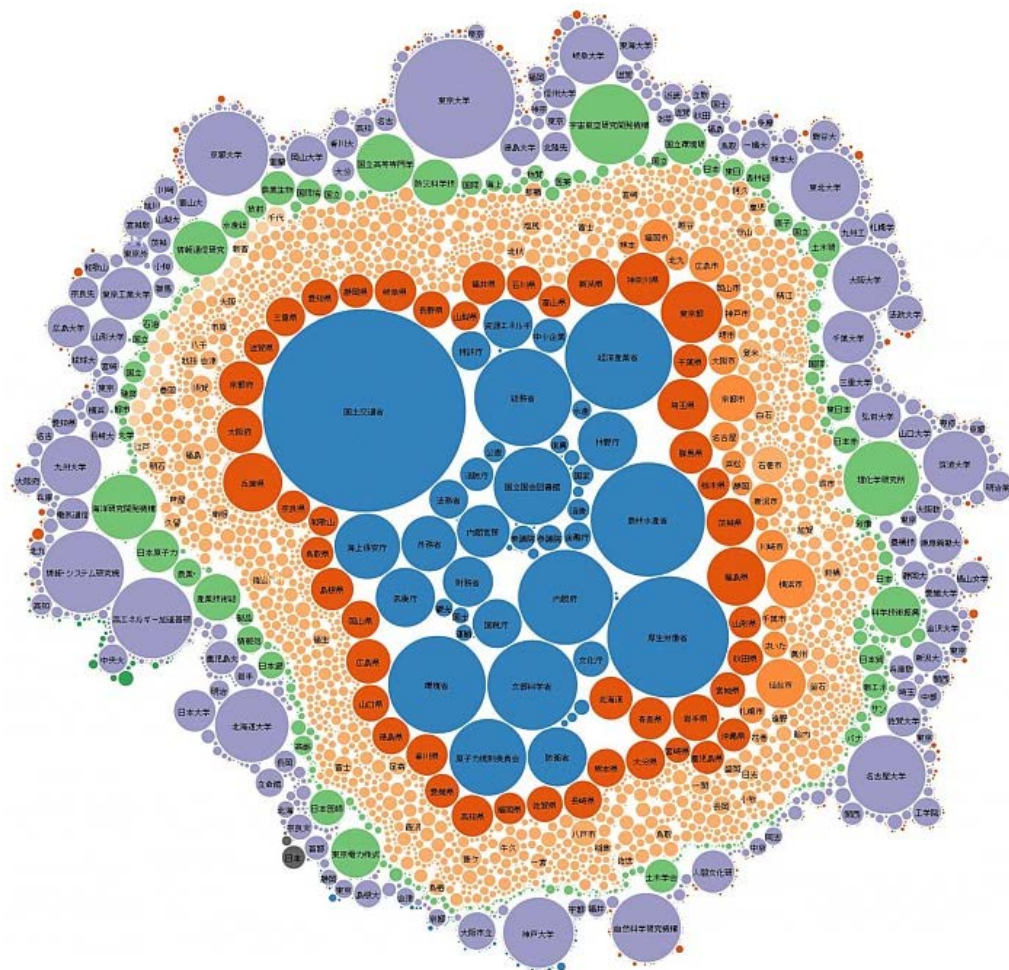
区分	根拠	対象	サイト数	収集頻度	容量
公的機関	法律	国の機関	5,400	月1回	910 テラバイト
		地方自治体		年4回	
		独立行政法人			
		国公立大学			
民間	契約	公益法人、学協会、私立大学、第三セクター、インフラ、業界団体、文化施設、政党、イベント、震災など	4,600	年1～4回	

2017年2月現在

月ごとの収集容量



種類ごとの収集容量

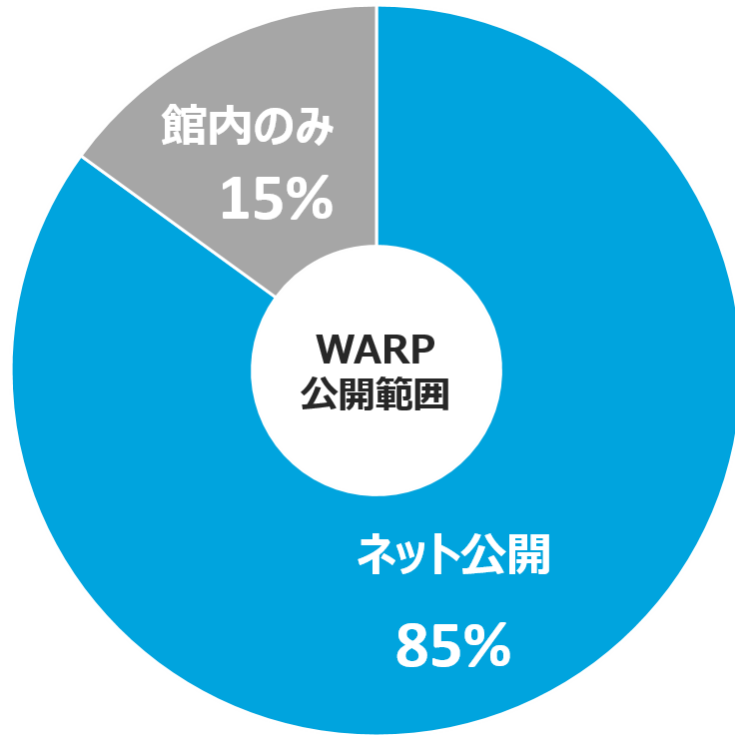


2016年11月現在

保存した 1 万サイトの可視化

<http://warp.da.ndl.go.jp/contents/recommend/collection/bubble.html>

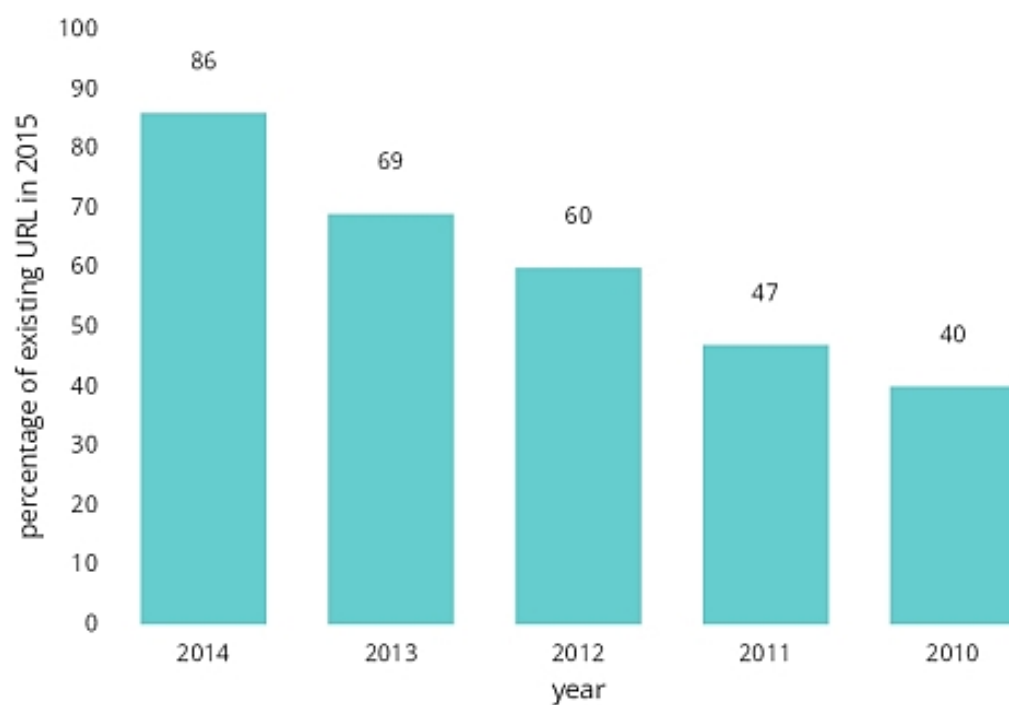
(かなりの) オープンアーカイブ



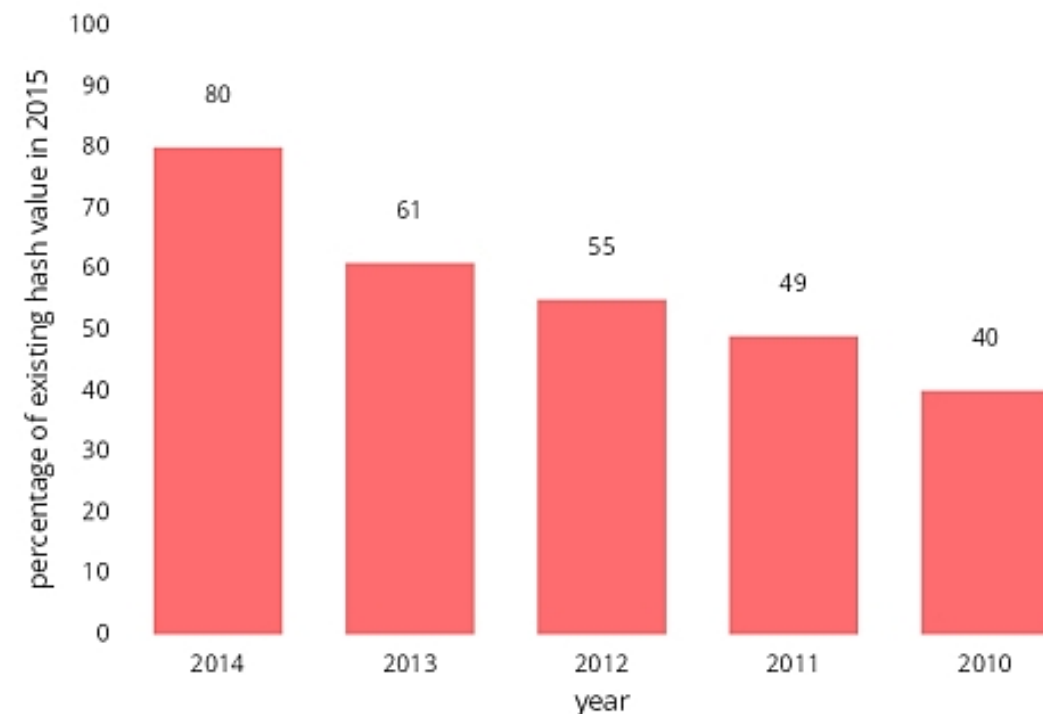
- ・ネット公開には発信者の許諾が必要
- ・許諾を得て85%をネット公開

2. 引用における活用

ウェブコンテンツはどのくらい消えていくのか



URLの残存率



内容の残存率

国の機関サイトの残存率調査

<http://warp.da.ndl.go.jp/contents/recommend/collection/linkrot.html>

<http://current.ndl.go.jp/e1757>

引用文献におけるリンク切れ/内容変化



Zittrain, Jonathan. et al. Perma: Scoping and Addressing the Problem of Link and Reference Rot in Legal Citations. Legal Information Management. 2014, vol.14, no.2, p. 88-99, <http://dx.doi.org/10.1017/S1472669614000255>

リンク先、引用でのウェブアーカイブ活用



英語版ウィキペディアの外部リンクのうち、リンク切れになったものをWayback Machine（IA）へのリンクに切り替える試み



中央省庁や自治体のサイトで古いページを消す際に、サイト管理者自らがWARPにリンクしてアクセスを保証する動き



学術論文の引用文献におけるリンク切れを、複数のウェブアーカイブを使って技術的に解決するプロジェクト



学術出版物の著者、編集者が自らウェブコンテンツのスナップショットを指定・保存し、永続的アクセスを保証するサービス



引用文献のリンク切れ/内容変化を防ぐため、引用者自らウェブコンテンツを指定・保存し、永続的アクセスを保証するサービス

perma.cc

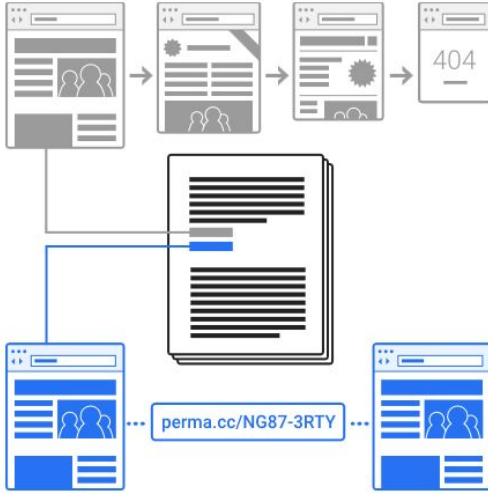
Perma.cc ∞

About Perma.cc Guide Blog Sign up Log in

Websites change. Perma Links don't.

Perma.cc helps scholars, journals, courts, and others create permanent records of the web sources they cite.

Perma.cc is simple, free to use, and is built and supported by libraries.



Sign up and use Perma.cc for free

How can my library get involved?

ハーバード大学が開発し米国の約120の法律関係図書館が参加

サービス構築の理由

- DOIは引用規則“Bluebook”に記載がなく法律分野では牽引力ない
- 多様な発信者が存在するウェブコンテンツへのDOI付与は現実的でない
- IAやWebCiteは単独民間セクターで継続性の保証がない

(Zittrain. op. cit.)

<https://perma.cc/>

perma.cc

Perma.cc is a simple way to preserve your links.

Find any URL.



Add it to Perma.cc.



Create a Perma Link.



Cite safely.



<https://perma.cc/>

3. DOIとの接点？

ウェブアーカイブとDOIの類似点

	考え方	活用例
ウェブアーカイブ	・「アクセスをいつまでも」	・リンク切れの代替リンク先 ・引用文献
DOI	・「アクセスをいつまでも」	・引用文献

ウェブアーカイブとDOIの相違点

	考え方	対象	主体
ウェブアーカイブ	・時間軸でコンテンツを固定化 ・アーカイブURLは永続的	・ウェブ全体	アーカイブ機関
DOI	・コンテンツは消さない ・URLの変更をリゾルブ	・学術論文 ・刊行物 ・研究データ	コンテンツ発信者

接点 -その1-

- ・DOIコンテンツは公開され続けるのが原則ですが、公開が取り下げられる場合も・・・ランディングページから先のアクセス保証？

cf. 「DOI登録後にデータ提供者が公開を取り下げた場合でも、DOIのランディングページは維持してメタデータの公開を継続し、アクセスが保証されるようにする。」
(研究データへのDOI登録ガイドライン - 3.3.2 DOI 登録後にデータの公開をやめる場合)

接点 -その2-

- ・アーカイブに対する考え方？
(特に第三者コンテンツ)

cf. 「DOI名のプレフィックスの中でオブジェクトを一意に識別する責任は登録者とその登録機関にあります。」
(DOI Handbook - 2.1 番号の割り当て)

接点 -その3-

- ・アクセス保証も視野に入れた、URLの引用ガイドライン？
（オリジナルURL、DOI、アーカイブURL、並列、補完…）

cf. 「電子媒体の文献は、内容が改変されてもそれがいつ行われたか把握しがたいし、時間経過とともにアクセスした場所から消失することも起こる。そのことは十分認識しておかねばならないが（中略）実際に参照したそのものを書くのがよい…」
（SIST02 - 参照文献の書き方 解説）