

学術リソースとしてのWebサイト引用方法について

On Citation of a Web Site as an Academic Resource

伊藤 博文[†] (愛知大学法科大学院)

要旨

近時、学術リソースとしてインターネット上のWebページが引用される場面が多く見られるが、その引用方法は多様であり、統一することが望ましい。引用方法の問題点を指摘し、私案としての引用方法を提示し、規格化への一歩を踏み出すための提案を行う。

キーワード：学術リソース，引用方法，URL，URI，検索エンジン表記法

1. はじめに

本稿は、インターネット上の学術リソース (Academic Resource) を引用する場合の表記方法について検討し、望ましい引用方法を新たに提案することを目的とする¹。

昨今さまざまな情報がインターネット上で公開されており、学術リソースも例外ではない。学術論文の多くは、Webサイトからの情報リソースを典拠引用している。しかしながら、その引用方法にはさまざまな形式があり、一致を見ていない。

そこで、インターネット上のコンテンツ

や情報リソースを引用する場合の引用方法を新たに考案するために、現状で行われている引用方法を検討しつつ、引用そのものの意義を再検討して、新たな引用方法を提案したい²。

2. 引用の意義

まず、引用の意義を確かめることから始めたい。【引用】という語の意味は、広辞苑³によると「自分の説のよりどころとして他の文章や事例または古人の語を引くこと。」とされる。また、もう一例を掲げる。アメリカ法を学ぶにあたり最も権威のあ

[†] 愛知大学法科大学院教授。以下のメールアドレスに忌憚なき意見や批判を送付していただければ幸いである。mailto: hirofumi@lawschool.aichi-u.ac.jp

¹ 本稿においては、混乱を避けるため、敢えて従来型の引用方法をとる。なお引用文中URLの最終アクセス確認日は2012年11月30日である。

² この分野の先駆的な研究としては、指宿信「ネット文献の引用方法について—学術資源としてのネットの可能性—」およびその掲げる参考文献等参照。available at <http://www.ne.jp/asahi/coffee/house/ARG/054.html#c>

³ 広辞苑第六版 岩波書店 2008年

るアメリカ法律学辞典である Black's Law Dictionary では、日本語の引用にあたる Citation を次のように説明している⁴。

引用 3 特定の主張を根拠付ける若しくは矛盾を指摘する判例，制定法，または条約といった，法的先例若しくは権威。

つまり、我々は、他人の言葉を引く行為そのものと、同時にその典拠を示すことも引用と呼んでいるのである。引用という意味は多義的に用いられ、第1は、Quotationの意味であり、他所に存在する文章等の一部をそのまま書き写して、自らの文書の中で使うことであり、第2は、Citationの意味であり、引用物の典拠を示すこと若しくはその表示方法である。

換言すれば、日本語の引用という語が、QuotationとCitationという意味を混在させているのである。論点を明確にするためにも、本稿では、第2の意味つまりCitationの意味での典拠表示としての引用（典拠表示）について論じることとする。

2.1 引用の目的

では、なぜ引用（典拠表示）をするかと

いう基本的な問題から始めたい。引用を行う場面は、学术论文がもっぱらであり、引用をすることの意味は、(1) 自己の主張を裏付けるため、(2) 権威付け、(3) 後発研究への経路を示すことの3つに集約できよう。

(1)、(2)は、相互関連性の強いものであり引用者の自己満足的な要素が強いが、(3) 後発研究への経路を示すこと、つまり「バックトラック（後述2.2）」を可能にすることで、客観的な検証が行われ初めて引用が意味を持つ。よって、重要なのは、バックトラックである。原典にたどり着く指標となることを引用の主目的と考えるべきであろう。

これにより、引用・転載した記述が、本当に存在するかを確認し、適切なコンテキストの中で引用・転載されているかを証明するという引用の持つ義務を果たせるといえよう。

2.2 バックトラック

引用の主目的は、参照典拠を示し後発研究への経路を示すことである。典拠を示すということは、読者が典拠を確認しようとしたときに、確実に原典である典拠物まで読者がたどり着ける道筋を示すこと、つまりトレース可能な表記が求められるのである。これをバックトラック（Back Track）と呼ぶこととする。

これは時間軸からは、当然過去へ戻るこ

⁴ Bryan A. Garner, Black's Law Dictionary 4th Pocket Edition, WEST 2011 at 115. 原文を以下に掲げる。Citation: 3. A reference to a legal precedent or authority, such as a case, statute, or treaties, that either substantiates or contradicts a given position.

ととなり、既に滅失してしまっている情報源へと道筋を辿ることも想定される。引用表記においては、確実にターゲットとなる情報源へと導く方法が必要であり、可能な限りの客観的な方策を講ずることが必要となる。

3. さまざまなリソースと引用方法

ここまでの引用の意義の検討を踏まえた上で、インターネット上のリソースとして、引用はどのような形で行われているのかを見ていく。まずは、アメリカ法でのCitationについて最も権威あるBlueBook⁵とALWD Citation Manual⁶を例にして、その表記方法を検討してみたい。

3.1 引用例

BlueBookも第19版となり、これまで以上に電子文字化されたリソースへの引用に対して紙面を割いている。対応するリソースとしては、インターネット上のWebサイト、オンラインデータベース、CD-ROM、マイクロフィルム、Podcastsといった録音源なども対象としている。ここでは、インターネットで最も情報量の多いWebサイ

トの引用、つまりURLによる引用を例としたい。

まず、Basic Citation Forms (a) Internet Source (rule 18.2)に掲げられている引用例から2点を示す⁷。

Sample 1.

Santa Monica, Cal., Mun. Code ch. 3.20 (1976),
available at <http://qcode.us/codes/santamonica/>

Sample 2

Asahi Shimbun, A-Bomb Disease Ruling,
Japanese L. Blog (Mar. 27, 2009 9:29 PM),
<http://japaneselaw.blogspot.com>

上記Sample 1は、印刷文字として入手可能な文献がインターネット上でも入手可能であることを示しており、Sample 2は、ブログをURLで表記しており、特徴的なのはそのアクセス日時を表記している点である。

もう一つ例を出そう。アメリカでの法文書作成でよく使われるALWD Citation Manualも、PART4. Electronic Sourcesにおいて、インターネット上のリソース引用方法について紙面を割いて詳しく述べている。特に40.1 (d) URL⁸においてWeb上のコンテンツに対する引用方法が特徴的である。

⁵ The Harvard Law Review Association, The BlueBook - A Uniform System of Citation Nineteenth Edition, 2010, at PP.164-176.

⁶ Association of Legal Writing Directors, Darby Dickerson, ALWD Citation Manual 4th ed., Aspen Publishers 2010.

⁷ The BlueBook 19th ed., at 164.

⁸ ALWD Citation Manual 4th ed. at 343.

Sample 3

U.S. Govt. Printing Off., A 21st-Century Concept of Air and Military Operations,
[http://www.gpo.gov; select Libraries, select Catalog of U.S. Government Publications, search keywords "air and military"](http://www.gpo.gov;selectLibraries,selectCatalogofU.S.GovernmentPublications,searchkeywords%20%22airandmilitary%22) (2009).

ALWD Citation Manual に掲載されている Sample 3 は、合州国政府の刊行物の所在を示しているが、大規模なデータベースのような Web サイトから目的に辿り着く道筋として、URL のみならず、select, search-keywords "xxx" といった keystroke を使った引用方法を提示している。これは Bluebook に掲載されていない形式であり、単なる URL 表記からは一歩前進した引用形式であると評価できる。その理由は、バックトラックに対する配慮がみられるからである。

ここでの例示でわかるように、インターネット上のリソースに対する引用方法に定番があるわけではなく、現段階では試行錯誤が行われているというのが実情である。

3.2 引用の表示技術

インターネット上の情報つまりリソースは URL によって表記可能である。ここで URL という表記について触れておく。URL (Uniform Resource Locator) は、インターネット上の場所を示す表記方法のルールで

あり、RFC3986⁹に規定されている。たとえば、URI のスキーム (Scheme) としては、ftp, http, ldap, mailto, news, tel, telnet, urn などがあることを下記枠内の 1.1.2 Example は示している。

1.1.2. Examples

The following example URIs illustrate several URI schemes and variations in their common syntax components:

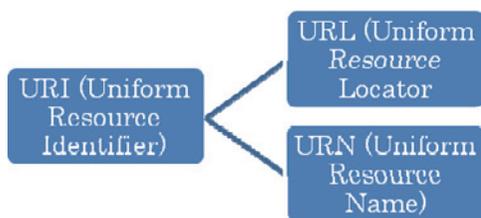
```
ftp://ftp.is.co.za/rfc/rfc1808.txt
http://www.ietf.org/rfc/rfc2396.txt
ldap://[2001:db8::7]/c=GB?objectClass=one
mailto:John.Doe@example.com
news:comp.infosystems.www.servers.unix
tel:+1-816-555-1212
telnet://192.0.2.16:80/
urn:oasis:names:specification:docbook:dtd:xml:4.
```

1.2

3.3 URL と URI について

ここで URL と URI について述べる。RFC3986 によれば、URL は、URI (Uniform Resource Identifier) の下位概念と位置づけられており、URI の一例が URL としている。インターネット上で場所を表記する場合、多くは Web 上のコンテンツを指し示すものであるため、URL と表記すれば事足りるのであるが、両者を使い分ける意味は、

⁹ RFC 3986 <http://www.ietf.org/rfc/rfc3986.txt> RFC (Request for Comment) とは、IETF (Internet Engineering Task Force) による技術仕様の保存、公開形式であり、インターネット上のネットワーク技術を標準化する作業時の仕様書となっている。<http://www.ietf.org/rfc.html>



URNの存在である。

URLが、場所を示す書き方のルールであり、Webページや画像などにアクセスするための主要な存在場所とアクセス方法を指定するものである。URN (Uniform Resource Name) は、名前を永続的に識別する書き方のルールとされる。URLがインターネット上で動的に変化する可能性のある場所を示しているのに対し、URNは、書籍のもつISBNのような静的に安定した記号を用いて情報源の特定を行うものである。たとえば、urn:isbn: 9784535510869のように表記するのである。

URIとURNの上位概念としてURIを置くが、URNが書籍名の引用といった従来の表記方法（たとえばISBN）と異なるものではなく、クリックしてもターゲットとなる情報源にアクセスできるわけではないので、結果的に普及していない。

インターネット上のリソースを表記する場合、インターネット上に存在しないリソースを参照する表記方法には意味がなく、結果的にURIとURLが同じものとなり、両者を使い分ける実益が乏しくなっている。よって、Webサイトの場所を示すた

めには、一般に普及しているURLと表記すれば事足りるのである。

3.4 リンク切れ

インターネット上のリソースは動的に動いているので、それを静的に捉えようとすると不具合が生じる。つまり、URLをクリックしても対象のWebページが表示されないという現象である。いわゆる、リンク切れ (Dead Link) の問題である。

このリンク切れの原因には、大きく分けて2種類ある。パターンAとしては、見たいWebページがインターネット上で今も見られる状態にあるが、URLが変更になったため参照できない場合である。パターンBとしては、もはや見たいWebページが削除されてしまい閲覧不能な状態になっている場合である。

まずは、パターンAである。具体例として、私のサイトのURL変遷を一例として示す。Project CaLSのサイトは、私の前任校の時代ではhttp://cals2.sozo2.ac.jp (図1)で



図1 旧URL http://cals2.sozo2.ac.jp



図2 新URL <http://cals.aichi-u.ac.jp>

あった。2004年に愛知大学に移ったことにより同一のサイト名が、<http://cals.aichi-u.ac.jp> (図2) へと変わった。よってURLとして<http://cals2.sozo2.ac.jp>を典拠としている私の過去の論文は、インターネット上で入手可能なPDFファイル¹⁰も含め、全てリンク切れになっている。旧URLから新URLへと誘導する場合、旧URL上のサイトに移転先を明示する若しくは旧URLを自動的に新URLへと導くWebページを置けば解決となるが、既に所属していない組織のサーバーにそのようなWebページを置くことは難しい。よってドメイン名の多様化にともなう変化もあわせて、URLが変化することは避けられず、リンク切れは今後も頻発することとなる。

パターンAに対して最も効果的なのは検索エンジンの活用である。コンテンツがインターネット上に存在しているがたどり着

¹⁰ 私の書いた論文はPDFファイル型式で下記URLにて公開している。
<http://cals.aichi-u.ac.jp/project/PN0160.html>。

けないのであれば、検索エンジンを使い拾い出すことが可能である。検索エンジンが全てのコンテンツをインデックス化しているわけではないので、必ずたどり着けるという保証はないが、かなりの確率でたどり着けるはずである。この点は次項において詳しく述べる。

次にパターンBである。インターネット上にもはや原典となるコンテンツそのものが存在しない場合にはアクセスしようがないのは当然である。しかし、このパターンBにおいてもバックトラックの方法がないわけではない。

インターネット上にコンテンツが存在しないというのは、著作者自らはもはや公開していないが、インターネット上には別の形で残されている可能性もある。それは、検索エンジンの持つキャッシュ、他者によるコンテンツの転載、インターネット上にあるアーカイブサイト(後述4.2, 図3参照)を利用する方法である。

これらの方法を使うことを前提としてどのような引用方法が望ましいのか次項で説明したい。

4. 新しい引用方法

ここでは、リンク切れによりバックトラックが不能となってしまう問題への対応策を考えていきながら、引用方法としての望ましい姿を検討していく。

4.1 引用の静と動

ここで新たな視点として電子文字という観点を視野に入れる。電子文字とはコンピュータ等の画面で表示される電気的な信号に置き換えられた文字であり、紙等に記載・印刷された静的な状態の印刷文字への対立概念である¹¹。

インターネット上のリソースを引用するということは、電子文字（動）を印刷文字（静）に置き換えること、つまり動的に変化し得る電子文字情報を静的で変化しない印刷文字にするという側面を持つこととなる。よって、常に変化し続ける電子文字を静的に捉えるということは、ある一時点での情報を印刷文字として静的に表記することにならざるを得ない。これは当然のことながら動的な電子文字の変化に対応はできない。既述のようにURLでも、ドメイン名の変化に伴いバックトラックを失うことは現実に多発している現象である。当然のこととして、日々変化し続けるインターネット上のコンテンツを紙面上で表記することには限界がある。

よって引用方法を考慮する場合、その引用方法が電子文字で行われるのか印刷文字で行われるのかも分けて考える必要がある。つまり、論文や書籍のように印刷され

紙面に表記されるには印刷文字用の引用表記が必要であり、オンラインで閲覧するコンテンツの場合は動的に対応できる引用が必要であり電子文字用の引用表記が必要となる。

印刷文字での引用表記には、一時点のURLを表記してその後の変化に対応すべく可能な限りのフォローをすべきであろう。一方、電子文字による引用では、Webサイト上のリンクを想定し、動的に変化可能な形式を想定すべきである。つまり、検索エンジンの力を借りて表記することも必要となる。電子文字の特徴である動的変化に対応できる機能を活かした引用方法が必要である。

そこで、引用を動的引用と静的引用と区別し、動的引用では、電子文字独自の機能を反映させた手法を用いるべきである。

4.2 過去に遡る技術

バックトラックでの大きな問題の一つであるリンク切れへの対策の一つには、インターネットアーカイブを利用する方法がある。

既述のリンク切れパターンBにおいては、もはやインターネット上に存在し得ないリソースを参照することになり、バックトラックができない状態であった。しかし、インターネット上には過去のWebサイト上のデータを全て保存しアーカイブ化しているサイトがある。それは、図3に

¹¹ 電子文字化については、伊藤博文「電子文字化論再考」愛知大学情報メディアセンター紀要『COM』Vol.22/No.1第37号1頁（2012年）参照。available at <http://cals.aichi-u.ac.jp/products/articles/ElectrifyingWrittenSymbolsRevisited.pdf>

示されている Internet Archive WayBack Machine である。この Web サイトでは、あたかもタイムマシンのように、過去に遡ってインターネット上の Web サイトの経年変化を調べることができる。

このサイトは膨大な量の蓄積データから過去の Web サイトを表示してくれるが、限界もある。たとえば、前掲 17 頁の図 1 では私の過去のサイトを表示させていたものであるが、ASP (Active Server Pages) のように動的な Web サイト表示には対応できておらず、完全な意味での復元とはなっていない。しかし、文字データなどはほぼ完全な形で保存されているので、文献引用といった場面において問題はなく使えると考えている。

これ以外にも、検索エンジンの持つキャッシュを利用し、検索エンジンの検索結果一覧表に出てくるキャッシュを使うことも一方法である。そして、他者によるコ

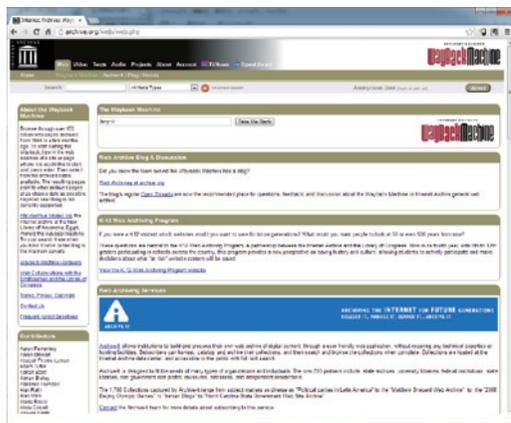


図3 Internet Archive WayBack Machine
<http://archive.org/web/web.php>

ンテンツの転載が有効な場合もある。インターネットのユーザーが、自身のブログや Web サイトに転載し掲示している場合は、これも検索エンジンにより拾い出す可能性もある。

4.3 検索エンジン最適化

検索エンジン最適化 (SEO: Search Engine Optimization)¹²は、インターネット上の検索エンジンが普及するにつれ、検索エンジン・ビジネスの発展とあいまって注目される技術である。

検索エンジンの代表格である Google が検索結果を出すアルゴリズムは¹³

- (1) クロール：Google がサイトを認識して検出する
- (2) インデックス作成：Google がサイトのインデックスを作成する
- (3) 検索結果の表示：ユーザーの検索と関連する有益なコンテンツが含まれているサイトを表示する

の3つから構成される。

検索エンジンが検索結果として上位に表示するために重要なのは、(2)の「インデックス作成」過程である。Google の検索エンジンが放つ「Googlebotはクロールした

¹² <http://ja.wikipedia.org/wiki/検索エンジン最適化>

¹³ <http://support.google.com/webmasters/bin/answer.py?hl=ja&answer=70897#1>

各ページを処理し、検出したすべての単語とページ上の場所を登録した大規模なインデックスを作成します。さらに、titleタグやalt属性などの主要なコンテンツタグや属性に含まれる情報も処理します。」¹⁴としている。

つまり検索エンジンを利用する場合も、目的となるWebサイトのタグ、特に<title>～</title>をキーワードとすることで、バックトラックする可能性が高くなる。

```
<html>
<head>
<title>抜刷を差し上げます</title>
</head>
<body bgcolor="#d2b48c" text="#08084d"
link="#000000" alink="#146f12"
vlink="#2b44f9">
:
</body>
</html>
```

たとえば、上枠内は私のサイトの1ページ (<http://cals.aichi-u.ac.jp/project/PN0160.html>) のソースファイルの一部である。このWebページのタイトルは、タグ<title>とタグ</title>に挟まれた文字列「抜刷を差し上げます」である。この文字列を検索エンジンにキーワードとして入れれば、上位にランクされて検索結果表示がなされる。

¹⁴ <http://support.google.com/webmasters/bin/answer.py?hl=ja&answer=70897#1>

検索エンジンにも種類があり、どの検索エンジンを使うかによっても検索結果は異なる。また、それに近い形での二次的なサイトを見つけることも可能となる。つまり、ターゲットとなるサイトが消滅していてもそれを複製したサイト、一部を引用したサイトなどを見つけ出すことも可能となるのである。

つまり検索エンジン最適化の働きを上手く利用すれば、かなりの確率でバックトラックができるのであるから、この機能を引用方法に入れることが望ましいといえる。

4.4 アクセス日時表示

動的に変化する情報を静的に引用表記するのであれば、表記時の時間を示す必要がある。

確かに、ブログなどのように刻々と変化するリソースでは、時分秒まで特定する必要があるだろうが、通常は年月日までで表記すれば十分である。しかし年月日表示にしても、下枠内のようにアメリカ式表記とイギ

```
2012年12月21日 を表記すると
1. アメリカ式表記
   Dec 21, 2012 or 12/21/12
2. イギリス式表記
   21 Dec 2102 or 21/12/12
3. 本稿での提案
   20121221
```

リス式表記では異なる。これらを統一する意味からも、年（西暦4桁半角数字）＋月（2桁半角数字）＋日（2桁半角数字）で表記するのが好ましい。この表記方法の一番のメリットは、ファイル名などに利用したとき必ず時系列にコンピュータ上で並べ替えることができることである。

また日時表示では、“last visited on”とか“accessed on”を付加することなく、端的に8桁の半角数字だけを表示することが望ましいといえる。

4.5 あらたな引用方法

これまで述べてきた引用方法以外のものとしては、引用表記を行う時点での原典を画像で保存することも一案である。つまり、Webサイト上で原典となるページ画面のハードコピーを取って、それを画像として保存し典拠表示する方法である。この画像ファイルをサーバー上に置いてクリックすれば大きな画像を表示できるようにす

る。動的引用においては、むしろこのようなアプローチが好ましいともいえる。

まとめとして、引用形式として含ませる要素には、検索エンジン（Search）と検索キーワード（Titleタグ内の文字列）を表記し¹⁵、最終アクセス時刻を表示することも必要であろう。以上をまとめると、提案したいのは上記のような引用方法である。



¹⁵ 検索キーワードを表記するといっても、面倒な操作を伴うことも事実である。ブラウザ上で右クリックしてWebサイトのソースを表示させ、このソース内から<Title>タグ内にある文字列をコピー＆ペーストする作業が必要となる。こうした作業も将来的にはMacroで簡単に引き出せるようになることが望ましいであろう。

5. おわりに

知大学に感謝したい。

インターネット技術には無限の可能性がある。現在起きているリンク切れの問題、インターネット上の情報氾濫とその情報の玉石混淆問題は、なんらかの技術により解決されるであろうと予想される。インターネットを形成してきた繋ぐ技術が目指したものが、分散した情報へどのようなアクセス経路を示すかというアプローチであったのが変化しつつある。それは、昨今のクラウド化のような情報リソースの一局集中化により、より新たな情報の存在形式が生まれ、より確実に情報源へたどり着ける方法が確立されることが望ましいからであろう。

また、本稿の扱った引用についての議論が、さまざまな制約から、限られたコンテンツのみを対象とすることに留まっており、あらゆる学術リソースを網羅しているとはいえない。これは今後の研究に委ねざるを得ない。そのためにも本稿の提起する問題点や手法がその研究の一助になれば幸いである。もっと多様な視点から伝統的な文献引用という手法に考察の光が当てられて、次世代の引用方法が生まれてくることを期待したい。

本稿は、愛知大学研究助成共同研究B「大学教育における学生との双方向通信のあり方」による研究助成による研究成果である。このような研究助成を与えていただいた愛

