

Web 機械翻訳再現性評価の研究

宮澤信一郎[†] 岡田勇^{††} 宮崎瑞之[†]
横山 晶一^{†††} 江原暉将^{††††} 太田敏澄^{†††††}

機械翻訳再現性評価とは、Web ページを機械翻訳 (Machine Translation, 以降 MT と略) した時に、Web 画面の構成を正しく再現できるか否かの評価である。MT の翻訳品質の評価研究はよく行われているのに対して、翻訳再現性の評価の研究は、我々のグループの他には、筆者らの知る限りまだ行われていない。本研究では MT サイトの再現性評価を数理解析的評価も含めて行った。その結果、各 MT サイトで再現性にかかなりの開きがあることが判明した。再現性評価の研究では、作成した標準テストサイトを使用して、既存の MT サイトにどの程度の再現性があるかを評価し、MT システムの更なる開発に際して、注意点、改善点、制約条件について資するようにすることを目的としている。本論文では日英版の MT サイトを対象に行ったが、提案する手法は特定の言語に依存しない汎用性を有している。

Evaluating Transferability of Machine Translation of the Web

Shinichiro Miyazawa[†], Isamu Okada^{††},
Mitsuyuki Miyazaki[†], Shoichi Yokoyama^{†††},
Terumasa Ehara^{††††} and Toshizumi ohta^{†††††}

Evaluating transferability of machine translation (hereafter “MT”) refers to assessing whether or not the composition of Web screens can be properly reproduced by using MT on the Web. Presently, studies on evaluating the translation quality of MT are commonly conducted, but those on evaluating transferability of the translation have yet to be carried out. Therefore, this study evaluated transferability of MT sites, which included mathematical analysis evaluation. The result shows that transferability varies considerably depending on each MT site. The objective of this study is twofold: to evaluate the extent to which the existing MT sites are transferred, employing the Standard Test Site prepared by our group; and to identify areas that call for attention, improvement, and constraint for further development of MT systems. Although this evaluation study was conducted only for Japanese-English MT sites, the suggested method has the versatility of not being dependent on any specific languages.

1. はじめに

国際コミュニケーションの重要性はますます増加しているが、その中で MT サイトに対する需要と期待が増している。しかし MT の性能はまだ期待されているレベルには至っていない。我々は最新の MT サイトの比較を通じて、翻訳品質の評価では無く、①MT システムの機能評価[1][3]と、②Web MT の再現性評価 (機械翻訳した時に、Web 画面の構成を正しく再現できるかの評価) の研究を行っている。MT 品質の評価は以前より多くの研究があるが、再現性評価と機能評価はまだ研究が端緒についたばかりである[4]。今回は MT システムの再現性評価について述べる。

2. 再現性評価の目的

再現性評価は、Web ページを機械翻訳した後、元の Web の諸属性が、翻訳後の Web でも再現できるか否かの評価である。複雑に構成された Web 文書の機械翻訳はインターネット特有の困難さがある。ページやフォント、レイアウトなどにきちんと対応しているか、ボタンなどをきちんと表せるか、画像に文字が埋めこまれているようなものに対処できるか等について、最新の機械翻訳サイトについて評価を行った。今までの我々の研究で、これらの諸点が翻訳時に翻訳されなかったり、脱落したり、変質することが頻出する場合があることが分かっている[2]。その場合は重要な情報を見逃したり、誤認してしまう場合があり、重大な結果を招くので、この面の評価から発展させて、再現性の標準テストサイトを構築し、標準的、網羅的に評価する研究を続けてきた。

3. 標準テストサイト

Web 機械翻訳の再現性に関して、各 MT サイトを同一の基準で標準的にテストし評価するために、我々は Web 上に再現性標準テストサイトを構築した。現在は日本語対象の翻訳テストサイトである。英文対象の翻訳テストサイトは開発中であり、開発完了後は、英語サイトも含めてテストと評価を行えるようにする予定である。

[†] 秀明大学 英語情報マネジメント学部
Faculty of English and IT Management, Shumei University

^{††} 創価大学経営学部
Faculty of Business Administration, Soka University

^{†††} 山形大学 工学部
Faculty of Engineering, Yamagata University

^{††††} 山梨英和大学 人間文化学部
Faculty of Humanities, Yamanashi Eiwa College

^{†††††} 電気通信大学大学院 情報システム学研究科
Graduate School Information Systems, University of Electro-Communications,

現段階での再現性標準テストサイトは次のような構造を有している。図 1 に標準テストサイトトップ画面を示す。ここでは標準テストサイトの目的と概要について記述している。図 2 は標準テストサイトインストラクション画面である。標準テストサイトの使用方法、テスト評価方法について記述している。左側のフレームはテスト項目一覧で、テストしたい項目をクリックすると、図 3 のような具体的なテスト項目が表示される。この例ではフレーム関係のテストである。この画面では、テスト対象、テスト画面のソース、再現が不成功の場合の原因、原理的に再現可能か否かの情報が記載されている。「模範翻訳へ」ボタンをクリックすると、模範的な翻訳画面が表示される。「テスト開始」ボタンをクリックすると、図 4 の実際のテスト画面が表示される。この例ではトップフレーム、中央フレーム、左フレームから構成されたテスト画面が表示されている。これからも分かるように、純粹にそのテスト項目をテストできるように、テスト画面は出来るだけ単純に作られている。図 5 は、このフレームテスト画面を、ある翻訳サイトで翻訳した場合の結果である。フレームが再現されていないことが分かる。

3.1 テスト項目細目

我々の開発した標準テストサイトにおける、テスト項目の細目は次の通りである。

(1) ページ設定関係

タイトル、ヘッダ

(2) フォント関係

フォントの再現（イタリック、強調、アンダーライン、フォントの種類、フォントサイズ）、訂正線の表示、タグを使ったフォントの再現（テキストの強調、サブスクリプト、スーパースクリプト、等幅フォント、BIG タグを使ったフォントサイズ、SMALL タグを使ったフォントサイズ）

(3) イメージ関係

クリカブルマップ、PDF（これはイメージだけでないので、独立した関係にすべきかもしれない）、文字画像（gif）

(4) レイアウト関係

マーキー、改行、センタリング、スペース、水平線、アスキーアート

(5) ハイパーリンク関係

リンク名の翻訳、正しくリンクできたか

(6) リスト関係

見出し、連番、行頭記号、リストタイトル

(7) テーブル関係

見出し体言止め、複数行のタイトルボックス

(8) フォーム関係

選択リスト、ドロップダウン、ボタン

(9) フレーム関係

(10) タグミス関係

省略可能な終了タグ、タグのスペルミス・不一致

(11) JavaScript 関係

最終更新日、スクロール文字（ステータスバー）、スクロール文字（ページ）、Java アプレット

(12) CGI 関係

(13) 効果関係

アンカーの文字、アドレス

3.2 各関係の説明

ここでは、各細目の詳細について説明する。

1) ページ設定関係

ウェブページのタイトルは、ブラウザのタイトルバー部分に表示される。Web 製作者はタイトルによって、ページ内容などの記載事項を端的に表現することが可能となる。また、閲覧者はタイトルを読むことにより、該当ページを瞬時に見分けることが、ある程度可能となる。該当ページ内にもタイトルが大きく表示されているのが普通であるが、これらは文字画像が使われていることが多く、現在の技術では翻訳は困難である。そのため、HTML のタイトル部分の翻訳処理を正確に行う必要性は極めて高い。そこで、ページ設定関連ではタイトルが正確に翻訳されているかテストを行う。

2) フォント関係

Web 上の文章において、文字の書体や種類、サイズといったフォントの指定は頻繁に行われる表現方法である。フォントの変更は閲覧者の注意を引くためや、重要ポイントを強調することなど、製作者が何らかの意図を持ち、必要があって行っているのであるから、翻訳結果においてもそのフォント変更が適切に反映される必要性が極めて高い。フォントの訂正線も同様である。

書体や種類については、言語が異なるため、英語と同様のフォントが無い場合もあり得る。その場合、その言語に合った何か一般的な別のフォントに変えるだけでも効果的である。そこでこの項ではこれらフォント関連が適切に再現されているかをテストする。

3) イメージ関係

この項目では文字画像や、クリカブルマップ中の文字や、ボタン名の翻訳や、図の再現ができるかをテストすることが可能である。Web によっては頁の大半を文字画像で占めているものもあり、これらが機械翻訳できないと、ほとんど翻訳できない頁が存在することになり、機械翻訳にとって大きな課題である。

4) レイアウト関係

レイアウト関連では、マーカーが方向も含めて正しく再現されているかに加え、改行、センタリングの再現についてのテストが可能である。また水平線やアスキーアート（文字を使って表した図）の再現テストも行う。マーカーは右、左、左右のいずれかの指定によりタグで囲まれた範囲の文面を動的に指定方向へ移動させる事ができる。また、改行やセンタリングの利用によって様々なレイアウトができる。これらは製作者が意図的に行っているので翻訳結果においても同じように反映されたほうが良い。

5) ハイパーリンク関係

ハイパーリンクにおけるリンク名が正しく翻訳され、正確にリンクされているかのテストが可能である。正しく翻訳されなかった場合として、以下のことが考えられる。

- (1) 翻訳されていないリンクがある。
- (2) リンクの順番が変わる。
- (3) 1個のリンクが複数のリンクに翻訳されている。
- (4) 逆に複数のリンクが翻訳後に1つになっている。
- (5) 複数のリンクが1つの文になっている。

これらの不具合が起こらないように翻訳が実行されなければならない。

6) リスト関係

リストは行頭に行頭記号（無い場合もある）を付加し、見出しとしての利用や文章のレイアウト等にも利用される。行頭に付加する記号は複数存在し、翻訳サイトでの翻訳結果においても正確に再現することが必要となる。

7) テーブル関係

セルで区切られた領域を持たせる事を可能とするテーブルタグは、文章のレイアウトのみならず、画像のレイアウトなど、HTMLの中でも頻繁に利用される。テーブルのタグ構造は単純で利用範囲も広く様々な関数的要素が用意されている。

テーブルを構成するタグには背景を色付きにする事も可能である、その際、数値での色指定も可能だが、逆に「black」など色そのものの名称を指定する事も可能である。間違っただけ名称個所を翻訳した場合、テーブルが表示されなくなる可能性も十分考えられるため、翻訳サイトのHTML解析に高い精度が要求される。

8) フォーム関係

Web ページ上でフォームを利用すると、WWW サーバとインタラクティブ(対話)ができるようになる。ユーザが情報をフォームに入力したり、いくつかあるリストから項目を選択したりするものがある。入力する場合は、その内容も同時に翻訳される必要がある。

フォームには、フォーム入力以外に、選択リスト、ドロップダウン、ボタン等があ

る。選択リストは、もともとリストをいくつか表示しておき、その中から項目を選択し入力することができるコントロールである。ドロップダウンは、右端の▼をクリックすることにより選択肢が表示され、項目を選ぶことができるコントロールである。ボタンには通常、文字や数字が書かれており、クリックすることにより選択できる。

9) フレーム関係

フレームは親ページ内にフレーム子ページを指定した情報がタグで記述されている。そのため翻訳処理を行う場合、親ページのタグをきちんと参照しないとフレームが再現されない。

10) タグミス関係

タグは Web ページを構成している要素であり、フォント情報やレイアウト、ハイパーリンク等すべての機能はタグによって表現されている。タグは始点と終点とで同じものでなければならないが、Web ページ作成者のミスによりタグが一致していない、終点タグが無い等のミスがある。また、ホームページ作成ソフトを使ってページを記述する場合であってもタグの重複や不必要なタグの記述が発生する場合がある。

Web ブラウザはそれらを不具合が起きぬよう認識して表示させており、その結果、ミスがあっても作成者が訂正を施さない場合がほとんどである。翻訳サイトは、タグにミスがある場合も、Web ブラウザと同じようにタグを認識し表示させる必要がある。また、タグが一致しないからといって翻訳が中断されることがあってはならず、テストの必要性は高い。

11) JavaScript 関係

視覚的に閲覧者の興味を引く手法として JavaScript は非常に多くのサイトで用いられており、JavaScript 中の文字は正しく翻訳される必要がある。また JavaScript で記述できるものの中で、Web ページの最終更新日を自動表示させる方法は多くのサイトで用いられている。このようなページを翻訳する際に JavaScript を実行して更新日を翻訳した日時に書き換えてしまう可能性がある。JavaScript は種類が多いため、全てのテストは不可能であるが、頻繁に利用されるものは翻訳できてほしい。最終更新日以外にも良く利用されるものとして、ステータスバー文字のスクロールや、ページ中のスクロール文字、Java アプレットをテストした。JavaScript はページの情報が変わってしまうため、ページとして表示されているものをそのまま再現できるかが問題である。

12) CGI 関係

HTMLはWebサーバに蓄積してある文書を単に送出するだけであるのに対し、CGIは、Web ブラウザからの要求に応じて、Web サーバがプログラムを起動し、その処理結果に従って動的に文書を生成し送出する仕組みである。具体的にはアクセスカウンタや掲示板、ショッピングカート等があり、現在の Web ページでは多く使用されている技術なので、機械翻訳可能であることが望まれる。

13) 効果関係

効果関係とは、SEO(Search Engine Optimization)対策に効果があるタグを言う。今回はアンカーとアドレスタグをテストする。アンカーは他ページへのリンクである。メニューやリンクリストに付けられたアイコン等の画像をアンカータグで囲むことで、alt 属性を有効に活用でき SEO 効果がある。アドレスタグは Web 文書に関する連絡先・問合せ先を表す。住所だけでなく、ホームページのリンク先・メールアドレス・電話番号・FAX 等の連絡を取るために必要な情報を書くことが出来、イタリック体で表示され、SEO 効果がある。

1.4) 複合型テストページ

上記までは単一のタグのテストであるが、複合型テストは複数のタグで囲まれた場合のテストページである。テスト対象は二つある。重複タグのテストは、テキストが同種のタグで囲まれた時の再現性のテストである。もう一つの異関係タグの混在のテストは、テキストが異種のタグで囲まれた場合のテストである。

4. テスト結果

本節において、主要MTサイトの再現性評価を行う。

4.1 調査対象機械翻訳サイト

我々は今回の再現性評価の再テストを行うにあたり、下記6つのサイトを選出した。主な理由として、いずれも日本国内を含め、世界的に有名なサイトが運営する機械翻訳サイトである為、多くの人々が利用するからである。

- (1) Google 機械翻訳サイト：<http://translate.google.co.jp/?hl=ja&tab=wT#>
- (2) Infoseek 機械翻訳サイト：<http://translation.infoseek.co.jp/?ac=Web&lng=en>
- (3) Livedoor 機械翻訳サイト：<http://translate.livedoor.com/>
- (4) Nifty 機械翻訳サイト：<http://honyaku.nifty.com/>
- (5) OCN 機械翻訳サイト：<http://www.ocn.ne.jp/translation/webpage.html>
- (6) Yahoo! Japan 機械翻訳サイト：<http://honyaku.yahoo.co.jp/url/>

これらのサイトに対し、我々の開発した「標準テストサイト」を翻訳した結果、再現性に問題のあった web ページが多く見られた。問題のあるページの例を図 5 に示す。また表 2「再現性評価一覧」に結果を示す。

4.2 テスト結果分析

1) ページ設定関係

タイトルの翻訳は2つの翻訳サイトでしか再現されていなかった。タイトル部分はタグ<title>-</title>で囲まれているだけで翻訳は難しいことではなく、十分に翻訳可能である。再現できていない翻訳サイトはタイトル部分の重要性を認識していないと思われる。それらの翻訳サイトにおいてはタイトル部分の翻訳に注意を向ける必要がある。

2) フォント関係

(1) フォントの再現 (イタリック, ボールド, アンダーライン, フォントの種類, フォントサイズ)

4つの翻訳サイトで再現されていたが、一部は不完全である。フォントはタグを認識することにより再現可能である。フォントの変更は必要があつて行われているので、翻訳結果でもそれを反映する必要がある。

(2) 訂正線

2つの翻訳サイトでしか再現されていなかった。訂正線はタグ<s>-</s>で囲まれているだけで翻訳は難しいことではなく、十分に翻訳可能である。

(3) タグを使ったフォントの再現

テキストの強調, サブスクリプト, スーパースクリプト, 等幅フォント, BIG タグを使ったフォントサイズ, SMALL タグを使ったフォントサイズである。半分くらいの翻訳サイトで再現されていたが、それも不完全であった。

3) イメージ関係

(1) クリックブルマップ

画像中の文字 (文字画像) は2つの翻訳サイトでしか出来ず、しかも不完全であった。しかしクリックブルマップは title= という形式でテキスト文字が書かれているので、座標の把握等が行えれば翻訳の可能性はある。

(2) PDF

翻訳できたのは2つに過ぎず、他は出来なかった。Web に占める PDF の多さを考えるとこれは問題である。正しく翻訳できたサイトがあることから考えて、技術的には可能であるので、対応すべきである。

(3) 文字画像

すべての翻訳サイトで翻訳不可能であった。現時点で翻訳によって文字画像の意味を把握することは難しい状態である。今後は OCR 技術の導入も含めて、ユーザとインタラクティブに翻訳を行うなどの工夫が必要である。

4) レイアウト関係

(1) マーキー

ほとんどの翻訳サイトで再現されていた。マーキーに関してはタグを見れば認識できるはずなので、十分に再現可能である。

(2) センタリング・改行，一行空けての改行，水平線
ほとんどの翻訳サイトで再現されていた。

(3) アスキーアート

すべてのサイトで再現できなかった。テキストなのかアスキーアートなのかを人間は判別できるがコンピュータが自動的に判別することが難しいためと思われる。

5) ハイパーリンク関係

ほとんどの翻訳サイトで再現されていた。ひとつのサイトでは文章自体の翻訳はされているが、ハイパーリンクが再現されていなかった。

6) リスト関係

(1) 見出し，連番，リスト記号

4つのサイトで再現・翻訳できていたが，一部のサイトでは連番が無くなってしまい再現されなかった。

7) テーブル関係

今回テストしたすべての翻訳サイトで再現・翻訳できた。

8) フォーム関係

(1) 選択リスト・ドロップダウン

ほとんどの翻訳サイトにおいて再現できていた。

(2) ボタンの文字

すべての翻訳サイトで再現できていなかった。ボタンの文字は絵文字になっているので翻訳が難しい。しかしタグの中の

```
<INPUT type = "submit" "value" = "send" >
```

の" value" 部分を認識できれば，翻訳は絶対不可能ということではないと思われる。

9) フレーム関係

ほとんどの翻訳サイトにおいて再現できていたが，フレームが一切再現されないサイトもあった。

10) タグミス関係

(1) タグの省略，タグのスペルミス

ブラウザには，タグのスペルミス・不一致がある場合に自動修正して表示するものがある。翻訳システムがブラウザと同様にタグを認識し，翻訳結果を表示できているかをテストした。全てのサイトできちんと認識し，再現・翻訳することができた。

11) JavaScript 関係

今回，大部分の機械翻訳サイトで JavaScript を正確に再現することは出来なかった。具体的には最終更新日のカウンターが正確に再現されない，ページの中やステータスバー文字のスクロールの翻訳が出来ていない。Java applet の再現が出来ていないなどである。

JavaScript 自体は非常に汎用性が高く，あらゆるサイトで利用されるようになって来

ている。そのため，原因となる要素を出来るだけ早く改善し，機械翻訳が対応できるようにしていくことが重要であると考ええる。

12) CGI 関係

CGI の一種であるアクセスカウンタを翻訳させたところ，3つのサイトで正しく再現できた。従来，CGI の再現は困難であったが，技術の発展がうかがえる。

13) 効果関係

ほとんどの翻訳サイトで再現できていたが，一部のサイトではアンカーやアドレスの翻訳がされていなかった。

14) 複合型テストページ

重複タグの表示は4つのサイトで再現翻訳されたが，残りは再現されなかった。異関係の混在タグではほとんど再現されていた。

5. 解析

前節ではそれぞれの評価項目の分析結果について検討したが，実務的には14分野33項目を縮約する方が，項目のサイト別の特徴を整理しやすい。そのためここでは次元縮約を試みる。本論文では，サイト数が6なので，主成分分析(表1, 表3)を参考として目検で行う。その結果，4成分を抽出し，それぞれフォント・イメージ成分(第1成分)，間接成分(第2成分)，高機能成分(第3成分)，レイアウト・効果成分(第4成分)と名付けた。

表1 主成分得点と各主成分の負荷量平方和

	主成分	PC1	PC2	PC3	PC4
サイト名	A	-1.2804	0.24114	1.55795	0.10642
	B	1.12802	0.36133	0.13113	0.34059
	C	-0.17488	-1.9334	-0.40888	0.47977
	D	1.18774	0.30267	0.54428	0.28769
	E	-0.05834	0.03809	-0.44974	-1.9897
	F	-0.80212	0.99013	-1.3747	0.7752
回転後の負荷量平方和	合計	6.522	5.403	5.223	5.065
	分散の%	28.356	23.493	22.707	22.022
	累積%	28.356	51.849	74.556	96.578

PC1: 第1主成分
PC2: 第2主成分
PC3: 第3主成分
PC4: 第4主成分

既存の6サイトの分析に限っては、累積の負荷量平方和が96%を超えているので、この4成分で再現性の基本的な特徴が抽出可能であることを示した。すなわち、翻訳サイトAの強みは高機能成分にあり、弱点はフォント・イメージ成分にある。翻訳サイトBの強みはフォント・イメージ成分にあり、弱点はあまり無い。翻訳サイトCの弱点は間接成分にある。翻訳サイトDの強みはフォント・イメージ成分にあり、弱点はあまり無い。翻訳サイトEの弱点はレイアウト・効果成分にある。翻訳サイトFの強みは間接成分であり、弱点は高機能成分にある。

以上のことから翻訳サイトBと翻訳サイトDが、比較的欠点の少ない翻訳サイトであることが分かる。

6. おわりに

今回、MTの再現性を評価するテストサイトを構築した。具体的には14分野33項目からなる標準テスト項目を定義しその標準テストサイトを構築した。次に主要な6サイトを選び、標準テストサイトを用いて全項目について評価した。

その結果、テスト項目毎の再現性について分析した。どのサイトでも再現できなかった項目もあれば、ほとんど全てのサイトで再現可能な項目、サイトによって再現可能性が分かれた項目など、再現性に様々な問題点があることが分かった。

最後にテスト項目を縮約するために主成分分析を行い、主要の4成分を抽出した。翻訳サイトの各成分得点から、特定のサイトが欠点が少ないことが明らかとなった。

しかし、本論文にはいくつかの限界がある。まず、標準テストサイトでは各項目一つずつ分かれていたが、本来は同じページ内に複数の技術が複合的に含まれているので、今回のテストで翻訳できていても、一般のページで必ずしも同じ結果ができるとは限らない。

次に、今回は「日英」版についてのみの報告であり、「英日」版は行っていないので今後の課題としたい。またこの標準テストサイトはFLASHやXML等にも対応していく必要がある。

今回、現状のMTサイトの再現性の評価を試みた。Web技術の進歩は著しいので

今後も不断の評価研究が必要である。MTの品質評価研究の歴史は長いですが、MTの再現性、機能性の評価の研究は始まってまだ間がない。本研究がこの面での理論化、標準化への一歩となるよう願っている。

謝辞

最後に今回の調査研究の機会を与えていただいた Asia-Pacific Association for Machine Translation (AAMT) に感謝申し上げます。また御協力いただいた諏訪氏(電通大)、熊野氏(東芝)、柏岡氏(ATR)、小玉氏(富士通)、太田氏(JST)、研究を支えた宮澤研究室の諸君にも感謝したい。

参考文献

- 1) 宮澤信一郎, 林紘一郎: インターネット機械翻訳の機能評価に関する研究, 第17回情報通信学会大会(2000)
- 2) Miyazawa, S., Yokoyama, S., Matsudaira, M., Kumano, A., Kodama, S., Kashioka, H., Shirokizawa Y. and Nakajima Y.: Study on Evaluation of WWW MT Systems, Proceedings of MACHINE TRANSLATION SUMMIT VII '99 "MT in the Great Translation Era" (1999)
- 3) Miyazawa, S., Okada, I., Shimizu, N., Yokoyama, S. and Ohta, T.: Study on Evaluating Machine Translation in Cyber Commons, PROCEEDINGS World Multiconference on Systemics, Cybernetics and Informatics (2003)
- 4) 宮澤信一郎: 機械翻訳の評価・比較, 社団法人 情報科学技術協会『情報の科学と技術』Vol. 55, No.8, 特集 = 「機械翻訳」, p.339-344 (2005)



図 1 標準テストサイトトップ画面

MTTTS ■ MT標準テストサイト

■ 機械翻訳標準テストサイト MTTTS

機械翻訳標準テストサイトで提供するすべての評価方法は、特定の言語に依存しない汎用性のある評価方法を提供するものである。

■ 標準テストサイトの概要

現在、インターネット用の機械翻訳(以降MTと略)ソフトが数多く販売されている。この内、大部分はWWW(World Wide Web)翻訳用である。オフラインのMTはテキストを如何に正確に訳せるかという点が最大のポイントである。それに対し、WWWのMTは、テキストの翻訳精度の他に、各種のプロトコルやブラウザ等とだけ対応できているか、翻訳速度はどうかなど評価要因が多いのが特徴である。しかも新しいプロトコルの出現など、技術革新とそれらの実用化が速いのも特徴である。WWW用MTソフトの完成度が高まれば、世界の情報を共有することが可能になり、人類に多大の貢献をなすだろう。

本サイトの目的は、各メーカー、機関がWWW特有のMT面のテストを行い、自社ソフトの問題点を見出し、改良、発展に資する事にある。



MTTTS ■ MT標準テストサイト

■ このサイトの使用方法

1. テストしたい項目を左のフレームから選ぶ。
例えばマーカーが翻訳後も正しく表示されるかどうかをテストする場合、「レイアウト関係」をクリック。
2. テストしたい翻訳ソフトを起動させる。
3. **テスト開始** ボタンをクリックしテスト対象ページを表示させる。
この例の場合新規ウィンドウが開き「右マーカー」、「左マーカー」、「中央」が表示され、それぞれ動作している。
4. 2で起動した翻訳ソフトで3のページを翻訳してみる。
5. 正しく翻訳できたか翻訳結果を見る。
この例の場合、「右マーカー」が「Right Markey」、「左マーカー」が「Left Markey」、「中央」が「Center」と表示され、正しく動作しているれば、マーカーに関しては翻訳ができていることになる。
6. **機械翻訳** のボタンを押すと複数翻訳を見ることができる。

■ テスト評価方法について

テキスト部分の翻訳品質についての評価は基本的にやっていない。従来のテキストのMTの評価と同じだからである。テキスト以外のWWW固有な部分を中心に評価を行った。結果として、この評価方法は特定の言語に依存しない汎用性のある評価方法を提供するものである。

図 2 標準テストサイトインストラクション画面

表 2 再現性評価一覧

○:再現
 △:一部再現
 ×:再現不能

結果と比較		検証項目	A	B	C	D	E	F
1	ページ設定関係	① タイトルの表示	○	×	×	×	×	○
		② ヘッダの翻訳	○	○	○	○	×	○
2	フォント設定関係	① フォントの再現	△	△	×	○	○	×
		② 訂正線の表示	×	○	×	○	×	×
		③ タグを使ったフォントの再現	×	△	×	△	△	×
3	イメージ設定関係	① クリック画像中の文字	×	△	×	△	×	×
		② PDF	○	×	×	△	×	×
		③ gif	×	×	×	×	×	×
4	レイアウト設定関係	① マーキー (Marquee) の再現	○	○	×	○	○	○
		② 中央寄せ・改行・スペース	○	△	△	△	△	○
		③ 一行空けての改行	○	○	○	○	×	○
		④ <p>タグを使った文章の表示	○	○	○	○	○	○
		⑤ 水平線<hr>の表示	○	○	○	○	○	○
		⑥ アスキーアートの表示	×	×	×	×	×	×
5	ハイパーリンク関係	①	○	○	×	○	△	○
6	リスト関係	① 見出し・連番などの再現	○	○	×	○	○	×
		② リストのマークの有無	○	○	×	○	○	×
		③ リストのタイトルの有無	○	○	×	○	○	×
7	テーブル関係	① 体言止めの可・不可	○	○	○	○	○	○
		② 複数行のテキストボックス	○	○	○	△	○	○
8	フォーム関係	①	△	△	×	△	△	△
9	フレーム関係	①	△	○	△	○	△	×
10	タグミス関係	① タグの省略	○	○	○	○	○	○
		② タグのスペルミス	○	○	○	○	○	○
11	JavaScript関係	① 最終更新日の表示	△	△	×	△	×	×
		② ステータスバー	×	×	×	×	×	×
		③ ページ中のスクロール文字	×	×	×	×	×	×
		④ Javaアプレット (applet)	×	×	×	×	×	×
12	CGI関係	①	○	○	×	○	×	×
		②	○	○	×	○	×	×
13	効果関係	① altアンカーの翻訳	○	○	△	○	×	×
		② アドレスの翻訳と再現	△	△	○	△	×	△
14	複合型テストページ	① 重複タグの表示	×	○	×	○	○	×
		② 異関係の混在タグ	○	○	△	○	○	○

表 3 バリマックス回転後の主成分分析結果

		成分			
		1	2	3	4
フォント・イメージ成分	訂正線の表示	.897	.257	.262	.243
	クリック画像中の文字	.897	.257	.262	.243
	フレーム関係	.828	-.087	.545	-.039
	タグを使ったフォントの再現	.824	.256	.082	-.497
	重複タグの表示	.824	.256	.082	-.497
	中央寄せ・改行・スペース	-.807	.477	.071	.341
間接成分	タイトル表示	-.807	.477	.071	.341
	フォーム関係	.086	.947	.200	-.235
	マーキー (Marquee) の再現	.086	.947	.200	-.235
	異関係の混在タグ	.086	.947	.200	-.235
ハイパーリンク関係	ハイパーリンク関係	.098	.915	.303	.246
	PDF	-.328	.188	.875	.120
	alt アンカーの翻訳	.386	-.025	.825	.396
最終更新日の表示	最終更新日の表示	.378	.331	.816	.268
	CGI 関係	.378	.331	.816	.268
レイアウト・効果成分	一行空けての改行	.029	-.019	.220	.975
	ヘッダの翻訳	.029	-.019	.220	.975
	アドレスの翻訳と再現	-.037	-.623	.013	.781
複数行のテキストボックス	複数行のテキストボックス	-.582	-.148	-.267	-.141
	見出し・連番などの真理	.378	.365	.691	-.486
	リストのマークの有無	.378	.365	.691	-.486
	リストのタイトルの有無	.378	.365	.691	-.486
	フォントの再現	.471	.287	.420	-.661