Minna no Hon'yaku: A website for hosting, archiving, and promoting translations

Masao Utiyama[†], **Takeshi Abekawa**[‡], **Eiichiro Sumita**[†] **and Kyo Kageura**[§]

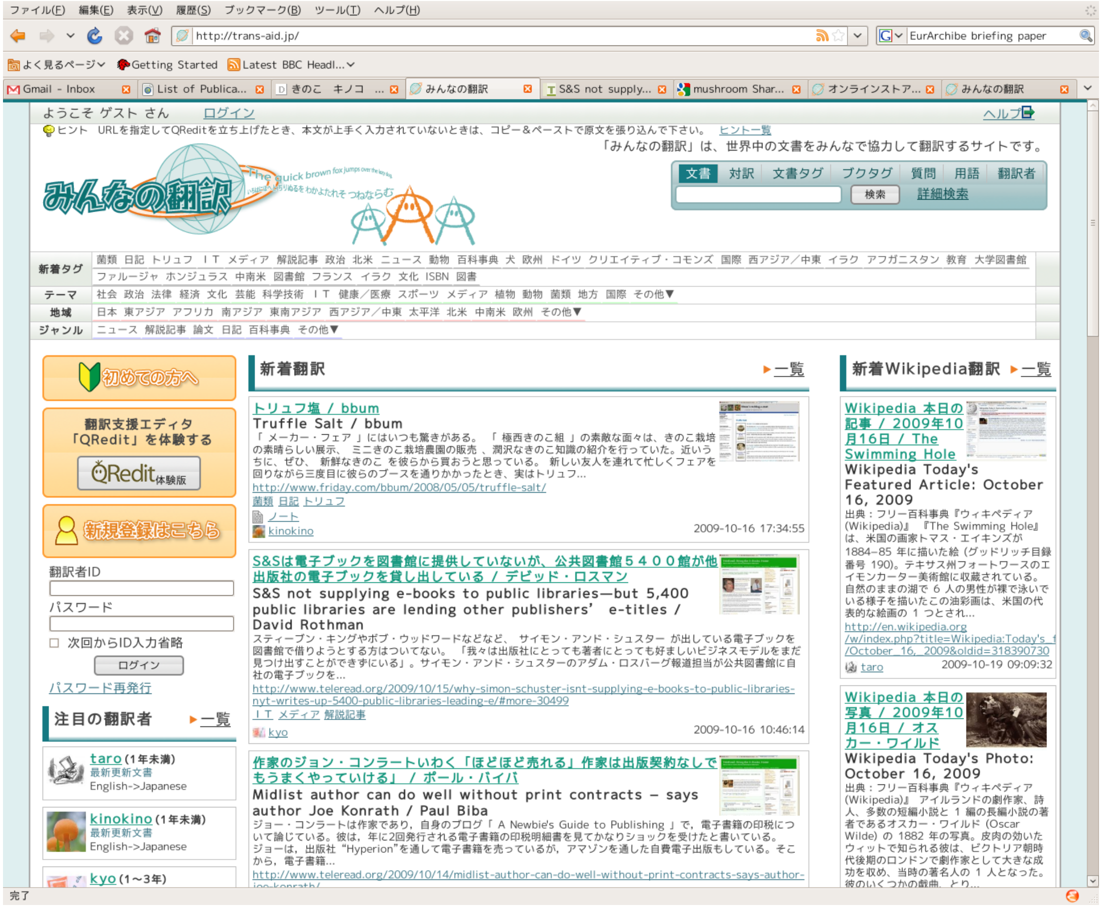
[†] MASTAR Project, National Institute of Information and Communications Technology, Japan
{mutiyama,eiichiro.sumita}@nict.go.jp
[‡] Center for Informatics of Association, National Institute of Informatics, Japan
abekawa@nii.ac.jp
[§] Graduate School of Education, University of Tokyo, Japan
kyo@p.u-tokyo.ac.jp

1 Introduction

In accordance with the rapid growth in the amount of information available in a variety of languages on the Web, translation needs and activities are sharply increasing. Many NGOs and NPOs that have been involved in translation as part of their activities since the pre-internet era have stepped up their translation activities; groups that were not previously involved in translation have started translating information related to their activities; online multilingual civic news sites have flourished; and a growing number of individuals have started translating online documents of their choice, in part or in full. Software localisation has come to constitute an important element in the distribution of software, and has become publicly visible due especially to the growth of free software.

Several notable features can be identified in the recent development of translation activities. One is the growth in the translation of online texts. This trend has gained impetus with the recent increase in the number of documents available under Creative Commons licenses, which helps promote the translation of online texts under clear licenses. A second feature is that online translation is mostly done by volunteer translators, who do not necessarily have expertise in translation. A third is that these translators rely heavily on dictionaries and Google search when translating. A fourth is that most of these translators do not use translation-aid tools. Finally, a fifth feature is that translation by volunteers has become an important source of providing unbiased or alternative information in an era of media concentration.

Against this backdrop, we have developed an integrated translation aid and hosting website, "Minna no Hon'yaku" ("Translation of/for/by everyone"; henceforth MNH), and made it publicly available since April 2009 at `http://trans-aid.jp`. The underlying concept of MNH is to make humans more intelligent by making machines more intelligent, and vice versa. MNH has two aspects: (i) a translation aid aspect that includes mechanisms for aiding online translators and facilitating the development of online translator communities by promoting inter-translator communication and collaboration; and (ii) a public media aspect that includes publishing translated information on the site and archiving translations with their originals. Section 2 of this



ヒント URLを指定してQReditを立ち上げたとき、本文が上手く入力されていないときは、コピー&ペーストで原文を張り込んで下さい。 ヒント一覧

「みんなの翻訳」は、世界中の文書をみんなで協力して翻訳するサイトです。

新着タグ: 菌類 日記 トリュフ ＩＴ メディア 解説記事 政治 北米 ニュース 動物 百科事典 犬 欧州 ドイツ クリエイティブ・コモンズ 国際 西アジア／中東 イラク アフガニスタン 教育 大学図書館 ファルージャ ホンジュラス 中南米 図書館 フランス イラク 文化 ISBN 図書

テーマ: 社会 政治 法律 経済 文化 芸能 科学技術 ＩＴ 健康／医療 スポーツ メディア 植物 動物 菌類 地方 国際 その他▼

地域: 日本 東アジア アフリカ 南アジア 東南アジア 西アジア／中東 太平洋 北米 中南米 欧州 その他▼

ジャンル: ニュース 解説記事 論文 日記 百科事典 その他▼

注目の翻訳者

taro (1年未満)
最新更新文書
English->Japanese

kinokino (1年未満)
最新更新文書
English->Japanese

kyo (1～3年)

新着翻訳

トリュフ塩 / bbum
Truffle Salt / bbum
「 メーカー・フェア 」にはいつも驚きがある。 「 極西きのこ組 」の素敵な面々は、きのこ栽培の素晴らしい展示、 ミニきのこ栽培農園の販売 、潤沢なきのこ知識の紹介を行っていた。近いうちに、ぜひ、 新鮮なきのこ を彼らから買おうと思っている。 新しい友人を連れて忙しくフェアを回りながら三度目に彼らのブースを通りかかったとき、実はトリュフ...
http://www.friday.com/bbum/2008/05/05/truffle-salt/
菌類 日記 トリュフ
ノート
kinokino 2009-10-16 17:34:55

S&Sは電子ブックを図書館に提供していないが、公共図書館５４００館が他出版社の電子ブックを貸し出している / デビッド・ロスマン
S&S not supplying e-books to public libraries—but 5,400 public libraries are lending other publishers’ e-titles / David Rothman
スティーブン・キングやボブ・ウッドワードなどなど、 サイモン・アンド・シュスター が出している電子ブックを図書館で借りようとする方はついてない。 「我々は出版社にとっても著者にとっても好ましいビジネスモデルをまだ見つけ出すことができずにいる」。サイモン・アンド・シュスターのアダム・ロスバーグ報道担当が公共図書館に自社の電子ブックを...
http://www.teleread.org/2009/10/15/why-simon-schuster-isnt-supplying-e-books-to-public-libraries-nyt-writes-up-5400-public-libraries-leading-e/#more-30499
ＩＴ メディア 解説記事
kyo 2009-10-16 10:46:14

作家のジョン・コンラートいわく「ほどほど売れる」作家は出版契約なしでもうまくやっていける」 / ポール・バイバ
Midlist author can do well without print contracts – says author Joe Konrath / Paul Biba
ジョー・コンラートは作家であり，自身のブログ「 A Newbie's Guide to Publishing 」で，電子書籍の印税について論じている。彼は，年に2回発行される電子書籍の印税明細書を見てかなりショックを受けたと書いている。ジョーは，出版社 “Hyperion”を通して電子書籍を売っているが，アマゾンを通した自費電子出版もしている。そこから，電子書籍...
http://www.teleread.org/2009/10/14/midlist-author-can-do-well-without-print-contracts-says-author-joe-konrath/

新着Wikipedia翻訳

Wikipedia 本日の記事 / 2009年10月16日 / The Swimming Hole
Wikipedia Today's Featured Article: October 16, 2009
出典：フリー百科事典『ウィキペディア(Wikipedia)』 『The Swimming Hole』は、米国の画家トマス・エイキンズが1884–85 年に描いた絵 (グッドリッチ目録番号 190)。テキサス州フォートワースのエイモンカーター美術館に収蔵されている。自然のままの湖で 6 人の男性が裸で泳いでいる様子を描いたこの油彩画は、米国の代表的な絵画の 1 つとされ...
http://en.wikipedia.org/w/index.php?title=Wikipedia:Today's_f/October_16,_2009&oldid=318390730
taro 2009-10-19 09:09:32

Wikipedia 本日の写真 / 2009年10月16日 / オスカー・ワイルド
Wikipedia Today's Photo: October 16, 2009
出典：フリー百科事典『ウィキペディア(Wikipedia)』 アイルランドの劇作家、詩人、多数の短編小説と 1 編の長編小説の著者であるオスカー・ワイルド (Oscar Wilde) の 1882 年の写真。皮肉の効いたウィットで知られる彼は、ビクトリア朝時代後期のロンドンで劇作家として大きな成功を収め、当時の著名人の 1 人となった。彼のいくつかの戯曲、とり...

Figure 1. The top page of MNH

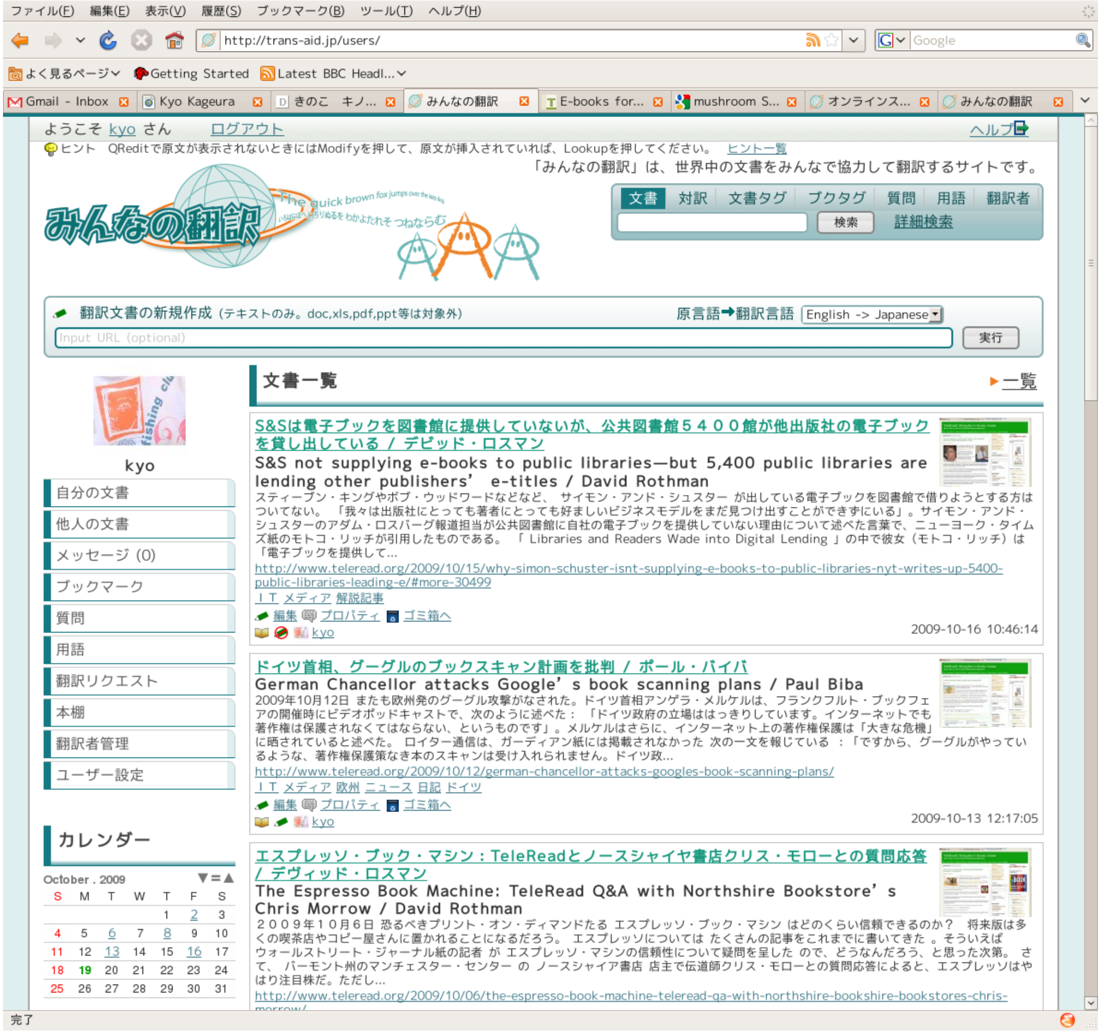
paper introduces the translation hosting aspect of MNH, and section 3 details the built-in translation aid editor QRedit. In section 4, we explain the community building functions as well as the educational functions provided by MNH. Section 5 describes the system's reference sources, and section 6 introduces its current state of usage. Section 7 concludes the paper.

2 MNH as a translation hosting site

2.1 Viewing documents

MNH provides information to those who access the MNH site looking for information. This is the same as any other news service or blog site. The only difference is that the information provided in MNH consists of translated documents. Figure 1 shows the MNH top page, which appears when internet users access the site. Anybody can read the translated documents published on the MNH site.

The main column in the centre provides contains newly uploaded translations; the right-hand column contains translations of “today's featured article” and “today's featured picture” from the English Wikipedia site [1]. The left-hand column provides a variety of information for users, i.e. how to use the MNH site, experimental use of QRedit, a link to the registration page, a



ヒント QReditで原文が表示されないときにはModifyを押して、原文が挿入されていれば、Lookupを押してください。 ヒント一覧

「みんなの翻訳」は、世界中の文書をみんなで協力して翻訳するサイトです。

翻訳文書の新規作成（テキストのみ。doc,xls,pdf,ppt等は対象外）

文書一覧

S&Sは電子ブックを図書館に提供していないが、公共図書館５４００館が他出版社の電子ブックを貸し出している / デビッド・ロスマン

S&S not supplying e-books to public libraries—but 5,400 public libraries are lending other publishers’ e-titles / David Rothman

スティーブン・キングやボブ・ウッドワードなどなど、 サイモン・アンド・シュスター が出している電子ブックを図書館で借りようとする方はついてない。 「我々は出版社にとっても著者にとっても好ましいビジネスモデルをまだ見つけ出すことができずにいる」。サイモン・アンド・シュスターのアダム・ロスバーグ報道担当が公共図書館に自社の電子ブックを提供していない理由について述べた言葉で、ニューヨーク・タイムズ紙のモトコ・リッチが引用したものである。 「 Libraries and Readers Wade into Digital Lending 」の中で彼女（モトコ・リッチ）は「電子ブックを提供して...

http://www.teleread.org/2009/10/15/why-simon-schuster-isnt-supplying-e-books-to-public-libraries-nyt-writes-up-5400-public-libraries-leading-e/#more-30499

ＩＴ メディア 解説記事

kyo 2009-10-16 10:46:14

ドイツ首相、グーグルのブックスキャン計画を批判 / ポール・バイバ

German Chancellor attacks Google’s book scanning plans / Paul Biba

2009年10月12日 またも欧州発のグーグル攻撃がなされた。ドイツ首相アンゲラ・メルケルは、フランクフルト・ブックフェアの開催時にビデオポッドキャストで、次のように述べた： 「ドイツ政府の立場ははっきりしています。インターネットでも著作権は保護されなくてはならない、というものです」。メルケルはさらに、インターネット上の著作権保護は「大きな危機」に晒されていると述べた。 ロイター通信は、ガーディアン紙には掲載されなかった 次の一文を報じている ：「ですから、グーグルがやっているような、著作権保護策なき本のスキャンは受け入れられません。ドイツ政...

http://www.teleread.org/2009/10/12/german-chancellor-attacks-googles-book-scanning-plans/

ＩＴ メディア 欧州 ニュース 日記 ドイツ

kyo 2009-10-13 12:17:05

エスプレッソ・ブック・マシン：TeleReadとノースシャイヤ書店クリス・モローとの質問応答 / デヴィッド・ロスマン

The Espresso Book Machine: TeleRead Q&A with Northshire Bookstore’s Chris Morrow / David Rothman

２００９年１０月６日 恐るべきプリント・オン・ディマンドたる エスプレッソ・ブック・マシン はどのくらい信頼できるのか？ 将来版は多くの喫茶店やコピー屋さんに置かれることになるだろう。 エスプレッソについては たくさんの記事をこれまでに書いてきた 。そういえば ウォールストリート・ジャーナル紙の記者 が エスプレッソ・マシンの信頼性について疑問を呈した ので、どうなんだろう、と思った次第。 さて、 バーモント州のマンチェスター・センター の ノースシャイア書店 店主で伝道師クリス・モローとの質問応答によると、エスプレッソはやはり注目株だ。ただし...

http://www.teleread.org/2009/10/06/the-espresso-book-machine-teleread-qa-with-northshire-bookshire-bookstores-chris-morrow/

October . 2009

S	M	T	W	T	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

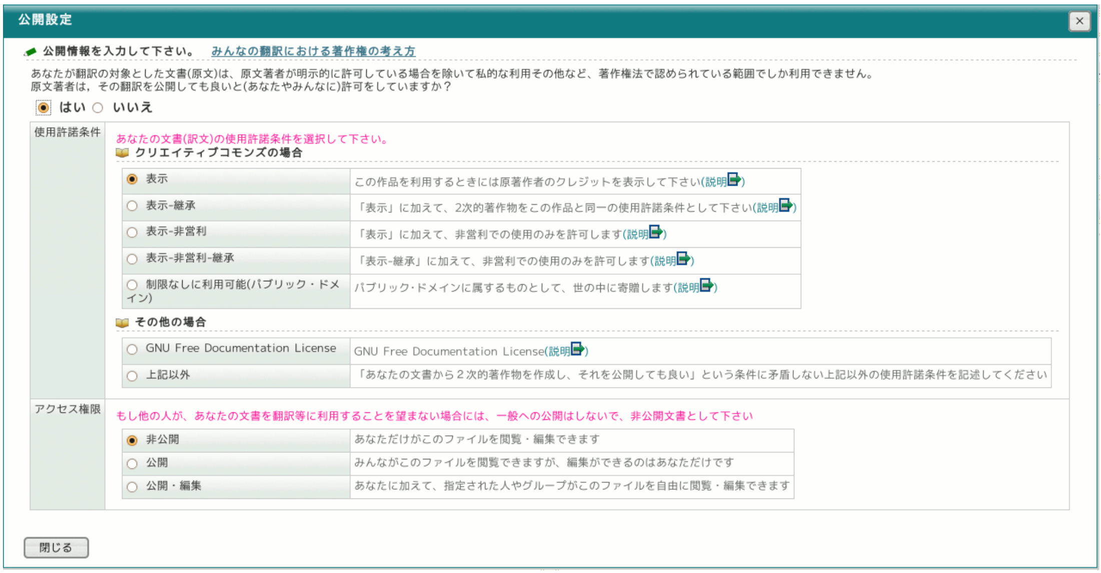
Figure 2. A personal user space in MNH

login box for registered users, and a list of translators who have recently published translations on MNH.

To navigate users to relevant documents, four sets of tags are displayed above the three columns: tags assigned to newly published translations; thematic tags; area tags; and genre tags. These tags are assigned by translators when they publish their translations on MNH. MNH provides a fixed number of pre-defined thematic, area and genre tags. In addition to these tags, translators can create their own tags. Users can also assign bookmark tags to published translations, which facilitates user-oriented organisation of translations on MNH.

2.2 Hosting translations

As a translation hosting site, MNH enables users to edit, store and publish their translations on MNH through their web browsers. In order to use these functions, one needs to obtain an MNH user account, which is free and open to anybody. Users can choose to remain anonymous, as MNH allows users to use pseudonyms. Registered users are provided with their own personal space,



公開設定

公開情報を入力して下さい。 みんなの翻訳における著作権の考え方

あなたが翻訳の対象とした文書(原文)は、原文著者が明示的に許可している場合を除いて私的な利用その他など、著作権法で認められている範囲でしか利用できません。
原文著者は，その翻訳を公開しても良いと(あなたやみんなに)許可をしていますか？

● はい ○ いいえ

使用許諾条件

あなたの文書(訳文)の使用許諾条件を選択して下さい。

クリエイティブコモンズの場合

● 表示	この作品を利用するときには原著作者のクレジットを表示して下さい(説明)
○ 表示-継承	「表示」に加えて、2次的著作物をこの作品と同一の使用許諾条件として下さい(説明)
○ 表示-非営利	「表示」に加えて、非営利での使用のみを許可します(説明)
○ 表示-非営利-継承	「表示-継承」に加えて、非営利での使用のみを許可します(説明)
○ 制限なしに利用可能(パブリック・ドメイン)	パブリック･ドメインに属するものとして、世の中に寄贈します(説明)

その他の場合

○ GNU Free Documentation License	GNU Free Documentation License(説明)
○ 上記以外	「あなたの文書から２次的著作物を作成し、それを公開しても良い」という条件に矛盾しない上記以外の使用許諾条件を記述してください

アクセス権限

もし他の人が、あなたの文書を翻訳等に利用することを望まない場合には、一般への公開はしないで、非公開文書として下さい

● 非公開	あなただけがこのファイルを閲覧・編集できます
○ 公開	みんながこのファイルを閲覧できますが、編集ができるのはあなただけです
○ 公開・編集	あなたに加えて、指定された人やグループがこのファイルを自由に閲覧・編集できます

Figure 3. The popup window that asks users for the copyright status of a document

in the same manner as ordinary blog services. Users can edit their personal information; make directories to manage translations; and make use of a variety of translation-aid and community building functions, etc. Figure 2 shows an image of a user's personal space when logged in.

The main differences between MNH and blog sites are:

1. As MNH is designed to host hosting translated documents, the documents that registered users publish are translations. Users make translations using the integrated translation-aid environment QRedit, the details of which will be explained in the next section. When users save their translations with permission to publish, they are published on MNH.

2. Translations made by users are published on the MNH top page. Unlike blog sites, the personal pages of registered users are not explicitly designed for the publication of their translations, although anyone can access the personal page of a registered user by inputting their user name as follows: `http://trans-aid.jp/users/?user_id=USER_NAME`. Further, unlike most blog sites, individual users cannot change the design of their own space.

The clarification of copyright is a critical issue in hosting and publishing translated documents. Though it is users who take responsibility for copyright issues, MNH introduces three steps to make sure that users respect copyright. First, in the process of registration, users must agree to terms of use that request them to abide by copyright laws. Second, when users open QRedit to translate a document, MNH asks them to confirm that they follow copyright laws. Unless users agree, they cannot open QRedit (though they can elect to make a general agreement the first time they use QRedit so that they can open it without this step from then on). Third, when users save a translation, the system requires the clarification of the copyright status of the original document. MNH first asks whether the author of the original document permits the publication of translations. If users answer "yes", then a popup window appears. Figure 3 shows the popup window. In this window, users specify the copyright status, among the alternatives given by MNH: Creative Commons BY, BY=SA, BY=NC, BY=NC=SA, GNU Free Documentation Licence, and others. Users then need to choose whether they want to publish the translation or keep it private (there is also a third choice, which we will elaborate on in section 4).

Translations of documents whose copyright permits the publication of translations are made publicly available on the MNH top page (as shown in Figure 1) when translators choose to publish the translations, while other translations made using QRedit on MNH are restricted to personal use or to use among a limited group of users. We will come back to the mode of limited use in section 4, in relation to the community-building functions of MNH.

3 QRedit: An integrated translation-aid editor

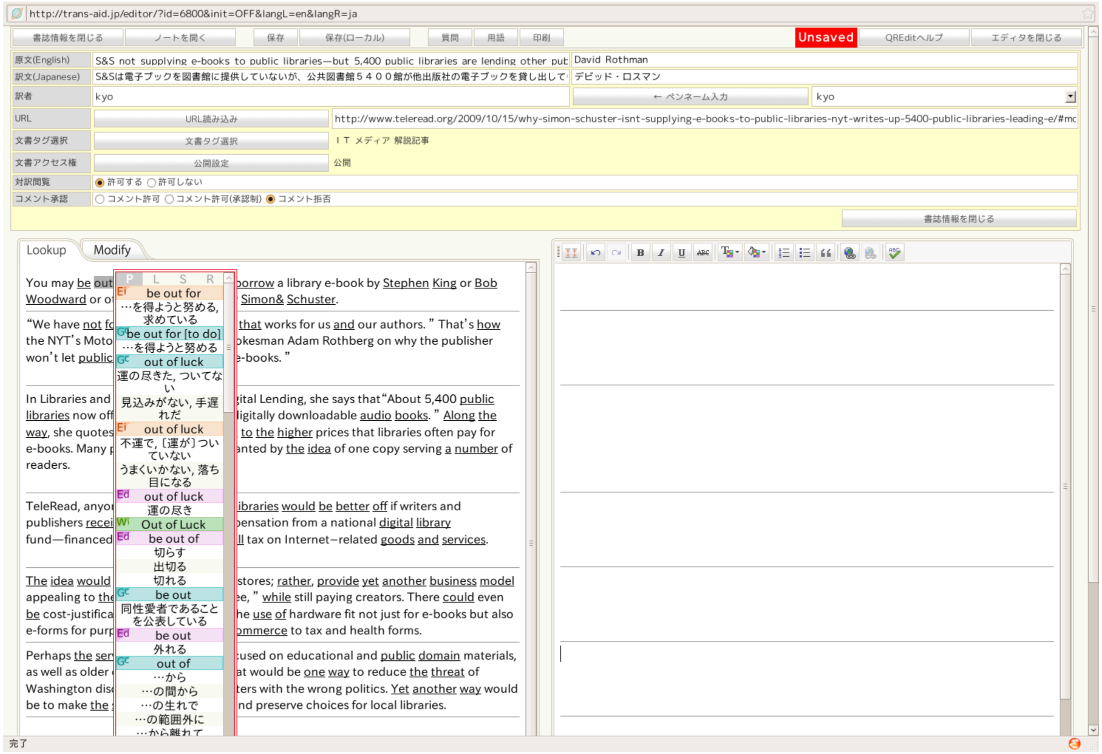
3.1 Needs of online volunteer translators

Volunteer translators involved in translating online documents have varied backgrounds. Some are professional translators volunteering part-time, some are interested in the topic, and some translate as a part of their NGO activities. They nonetheless share several basic characteristics:

1. They are native speakers of the target language (TL), which is mostly the case for any sort of translator;
2. Many do not have a native-level command of the source language (SL);
3. Most do not use translation aid systems or MT, because (i) most of these are commercial products that are too expensive for volunteer translators, and (ii) they are not readily accessible online;
4. They want to reduce the subjective burden involved in the process of translation. The reduction of time is also very important but perhaps less important for volunteer translators than for professional translators;
5. They spend a large amount of their translation time (about 30 to 60 per cent) looking up reference sources such as dictionaries or Google (cf. [2]);
6. The smallest basic unit of translation is perceived as a paragraph, not a sentence, and "at a glance" readability of the SL text is very important.

These traits define the basic features that any usable translation editor should provide for online volunteer translators:

1. It should be accessible online, through Web browsers;
2. It should provide high-quality dictionaries and an interface that promotes translators' awareness of dictionary entries for difficult expressions, so that they do not overlook, for example, idiomatic expressions [3];
3. It should provide a connection to Internet resources and should enable Internet searches to be initiated seamlessly from the translation editor environment;
4. It should provide a functional line of flow and interface that enables translators to maintain the rhythm of translation while using necessary functions such as dictionary lookup and Google search.



http://trans-aid.jp/editor/?id=6800&init=OFF&langL=en&langR=ja

原文(English)	S&S not supplying e-books to public libraries—but 5,400 public libraries are lending other pub	David Rothman
訳文(Japanese)	S&Sは電子ブックを図書館に提供していないが、公共図書館５４００館が他出版社の電子ブックを貸し出して	デビッド・ロスマン
訳者	kyo	kyo
URL		http://www.teleread.org/2009/10/15/why-simon-schuster-isnt-supplying-e-books-to-public-libraries-nyt-writes-up-5400-public-libraries-leading-e/#mc
文書タグ選択		ＩＴ メディア 解説記事
文書アクセス権		公開
対訳閲覧	● 許可する ○ 許可しない	
コメント承認	○ コメント許可 ○ コメント許可(承認制) ● コメント拒否	

You may be out … borrow a library e-book by Stephen King or Bob Woodward or o… Simon& Schuster.

"We have not fo… that works for us and our authors." That's how the NYT's Moto… okesman Adam Rothberg on why the publisher won't let public… e-books."

In Libraries and … gital Lending, she says that "About 5,400 public libraries now off… digitally downloadable audio books." Along the way, she quotes… to the higher prices that libraries often pay for e-books. Many p… anted by the idea of one copy serving a number of readers.

TeleRead, anyo… ibraries would be better off if writers and publishers recei… pensation from a national digital library fund—financed… ll tax on Internet–related goods and services.

The idea would… stores; rather, provide yet another business model appealing to the… ee," while still paying creators. There could even be cost-justifica… he use of hardware fit not just for e-books but also e-forms for purp… ommerce to tax and health forms.

Perhaps the ser… cused on educational and public domain materials, as well as older … at would be one way to reduce the threat of Washington dis… ters with the wrong politics. Yet another way would be to make the … nd preserve choices for local libraries.

Ei be out for
…を得ようと努める, 求めている
Gc be out for [to do]
…を得ようと努める
Gc out of luck
運の尽きた, ついてない
見込みがない, 手遅れだ
Ei out of luck
不運で, 〔運が〕ついていない
うまくいかない, 落ち目になる
Ed out of luck
運の尽き
Wi Out of Luck
Ed be out of
切らす
出切る
切れる
Gc be out
同性愛者であることを公表している
Ed be out
外れる
Gc out of
…から
…の間から
…の生れで
…の範囲外に
…から離れて

Figure 4. QRedit as it appears after start-up

3.2 Features of QRedit

QRedit has been developed as a translation-aid editor that provides various support functions to online translators based on the aforementioned requirements [3]. The user can open QRedit simply by clicking the QRedit start button, which appears immediately below the top banner of the MNH page when the user is logged in (see Figure 2). The user can either specify the URL of the document that s/he wishes to translate before clicking the start button or copy-and-paste the source document after opening QRedit. Figure 4 shows an image of QRedit after it has been opened. At this stage, (i) the bibliographic information area appears at the top of the window and (ii) dictionary lookup is activated in the source text area (which is on the left in Figure 4; users can rearrange the position of these areas). The bibliographic information area can be closed to enable the user to take full advantage of the screen for translating the document. Both the source and target text areas are divided into paragraph units. Scrolling of the source and target areas can be synchronised for both directions or from one to the other direction.

In the process of translation, translators can look up information by clicking words or phrases in the source text area. Figure 4 shows how dictionary lookup appears. When the user clicks a lexical token in the source text area, a small popup window appears, which displays translations from a high-quality dictionary provided by Sanseido [4], a few free dictionaries, and bilingual entries from Wikipedia (we will return to this point in section 5). Translation sentence pairs automatically collected from the Web and also provided by some NGO/NPO groups are also available for lookup [5], though not through the small popup window but by activating the sentence lookup session.

In addition to a simple dictionary lookup, QRedit also provides a flexible idiom/multi-word unit lookup mechanism. For instance, it can automatically look up the dictionary entry "with one's tongue in one's cheek" for the expression "He said that with his big fat tongue in his big fat cheek" or "head screwed on right" for "head screwed on wrong" [6].

QRedit provides a stratified reference lookup interface, which distinguishes three user awareness levels depending on the type and nature of the reference unit. These awareness levels are reflected in the way the reference units are displayed, for example, change in the background color or the use of an underline. The different awareness levels are assigned to reference units from a variety of reference sources, according to the criteria of "composition", "difficulty", "specialty" and "resource type" [7].

Starting from this small popup window, the user can trigger four types of action, corresponding to "P", "L", "S" and "R" shown on the top line of the popup window. In "P" mode (which is the default, though users can change their default settings), the user can paste any string in the popup window into the target text area by clicking it. If the user chooses "L" and then clicks the source language entry, a new, larger window appears, in which the full dictionary information is available. "S" mode enables the user to activate Google search by clicking a string in the small popup window. "R" activates the function of registering user defined translations.

Keyboard control remains fixed on the target text area, even when the user activates dictionary lookup by clicking source text tokens in the source text area using the mouse. This enables the user to maintain the rhythm of translation and focus on the making of target texts.

In an informal preliminary experiment, we observed two major effects of using QRedit:

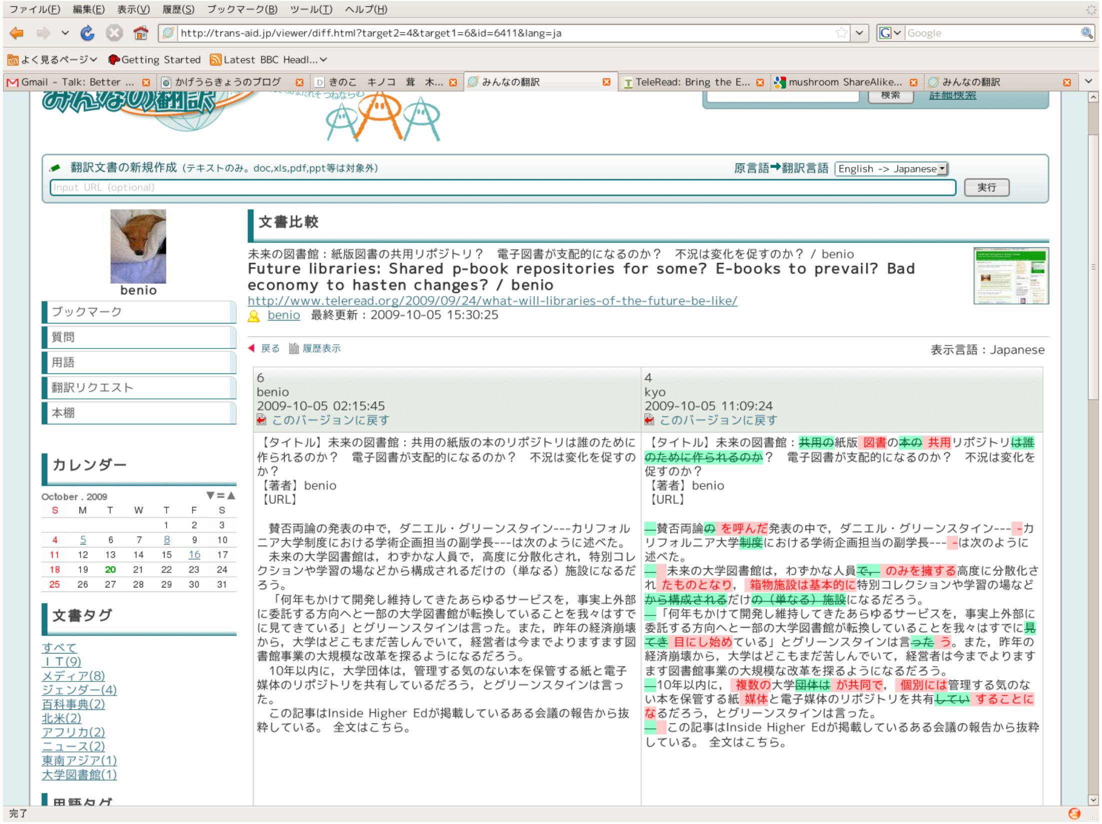
1. Translation time is reduced by 20 to 30 per cent, mainly because QRedit dramatically reduces the time necessary for dictionary lookup and web search; and
2. The quality of draft translations is improved, because paragraph and textual reading is not hampered by a cumbersome process of reference lookup.

Though further detailed and systematic research is needed on this point, QRedit provides most of the translation-aid functions required by online volunteer translators.

4 Community building and implicit translation training

Although the main target of MNH is individual translators and groups of translators who already know one another (such as a team of volunteer translators working in an NGO), MNH has several Web 2.0 features to facilitate the community building among those who register at MNH. It has a social tagging system that allows users to tag translations with their favorite keywords; it allows users to submit questions about translation to MNH, so that users can help each other; and it also allows users to issue request for translation. Message exchange function is also provided in MNH, so that users can communicate with each other within MNH, even if they do not know each other's external e-mail addresses or contacts information.

Users can issue permission for other translators to edit their original translation. Such permission can be open-ended, or restricted to a particular group of users. This makes it easy for groups of translators sharing the same topics of interest to form loosely networked communities.



翻訳文書の新規作成（テキストのみ。doc,xls,pdf,ppt等は対象外）　原言語➡翻訳言語　English -> Japanese

文書比較

未来の図書館：紙版図書の共用リポジトリ？　電子図書が支配的になるのか？　不況は変化を促すのか？ / benio
Future libraries: Shared p-book repositories for some? E-books to prevail? Bad economy to hasten changes? / benio
http://www.teleread.org/2009/09/24/what-will-libraries-of-the-future-be-like/
benio　最終更新：2009-10-05 15:30:25

表示言語：Japanese

6 benio 2009-10-05 02:15:45 このバージョンに戻す	4 kyo 2009-10-05 11:09:24 このバージョンに戻す
【タイトル】未来の図書館：共用の紙版の本のリポジトリは誰のために作られるのか？　電子図書が支配的になるのか？　不況は変化を促すのか？ 【著者】benio 【URL】 賛否両論の発表の中で，ダニエル・グリーンスタイン---カリフォルニア大学制度における学術企画担当の副学長---は次のように述べた。 未来の大学図書館は，わずかな人員で，高度に分散化され，特別コレクションや学習の場などから構成されるだけの（単なる）施設になるだろう。 「何年もかけて開発し維持してきたあらゆるサービスを，事実上外部に委託する方向へと一部の大学図書館が転換していることを我々はすでに見てきている」とグリーンスタインは言った。また，昨年の経済崩壊から，大学はどこもまだ苦しんでいて，経営者は今までよりますます図書館事業の大規模な改革を探るようになるだろう。 10年以内に，大学団体は，管理する気のない本を保管する紙と電子媒体のリポジトリを共有しているだろう，とグリーンスタインは言った。 この記事はInside Higher Edが掲載しているある会議の報告から抜粋している。 全文はこちら。	【タイトル】未来の図書館：~~共用の~~紙版 図書の~~本の~~ 共用リポジトリ~~は誰のために作られるのか~~？　電子図書が支配的になるのか？　不況は変化を促すのか？ 【著者】benio 【URL】 ~~—~~賛否両論~~の~~ を呼んだ発表の中で，ダニエル・グリーンスタイン--- -カリフォルニア大学~~制度~~における学術企画担当の副学長--- -は次のように述べた。 ~~—~~ 未来の大学図書館は，わずかな人員~~で，~~ のみを擁する高度に分散化され たものとなり， 箱物施設は基本的に特別コレクションや学習の場など~~から構成される~~だけ~~の（単なる）施設~~になるだろう。 ~~—~~「何年もかけて開発し維持してきたあらゆるサービスを，事実上外部に委託する方向へと一部の大学図書館が転換していることを我々はすでに~~見てき~~ 目にし始めている」とグリーンスタインは言~~った~~ う。また，昨年の経済崩壊から，大学はどこもまだ苦しんでいて，経営者は今までよりますます図書館事業の大規模な改革を探るようになるだろう。 ~~—~~10年以内に， 複数の大学~~団体は~~ が共同で， 個別には管理する気のない本を保管する紙 媒体と電子媒体のリポジトリを共有~~してい~~ することになるだろう，とグリーンスタインは言った。 ~~—~~ この記事はInside Higher Edが掲載しているある会議の報告から抜粋している。 全文はこちら。

カレンダー

October . 2009

S	M	T	W	T	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

文書タグ

すべて
ＩＴ(9)
メディア(8)
ジェンダー(4)
百科事典(2)
北米(2)
アフリカ(2)
ニュース(2)
東南アジア(1)
大学図書館(1)

用語タグ

Figure 5. A comparative view of different versions of a translation

The most important contribution of this function to volunteer translators, however, is its implicit translation training function for inexperienced translators. Many NGOs and NPOs constantly face the problem of a lack of good volunteer translators as well as high turnover among those they do have. Due to the low retention rate of volunteer translators, the core, experienced translators become busier, leaving them no time to give advice to inexperienced translators, thus creating a vicious circle.

The implicit translation training in MNH works as follows:

1. An inexperienced translator, after translating a document, gives permission for experienced translators to edit the draft translation. Experienced translators edit the draft, and finalise the translation. Note that this is the normal task flow adopted in many NGOs/NPOs.
2. MNH keeps a log of these finalised translations (a maximum of 10 versions for any one document). Thus, a record of how experienced translators modified the draft translation is retained in the MNH database.
3. Translators who have the permission to edit can see the difference between any pair of translation versions kept in the MNH database. Inexperienced translators can therefore check where and how their translations were modified by experienced translators. Figure 5 shows the comparative view of two different versions of a translation.

Although this mechanism still requires that inexperienced translators make the effort to learn from the modification logs, it can potentially provide great assistance in improving their skills, as they can take advantage of information that has not been accessible in the traditional working environment.

5 Reference resources

One of MNH's main features is that it provides rich – though limited in terms of language pairs – reference resources necessary for translation in QRedit. One of the most frequent requests made by volunteer translators is easy lookup of high-quality dictionaries and Wikipedia. Terminological management and TM are considered by many volunteer translators as desirable but not essential.

Bilingual dictionaries: The supply of a high-quality dictionary is essential to attract volunteer translators to MNH. We negotiated with one of the most prestigious dictionary publishers in Japan, Sanseido, and obtained permission to use *Sanseido's Grand Concise English-Japanese Dictionary*, which contains 360,000 entries. This is one of the most highly regarded English-Japanese dictionaries for translators. In addition, we provide the free Japanese-English dictionary EDICT [8] for both English to Japanese and Japanese to English translations. English-Chinese and Chinese dictionaries are to be supplied soon, which will extend MNH to Chinese.

Encyclopedic information: QRedit provides entries from Wikipedia. When there are corresponding Japanese and English entries, users can look up English to Japanese or Japanese to English translation information extracted from Wikipedia entries. When there are only English or Japanese entries, users can look up monolingual information.

Terminology and proper names: When MNH was made public, QRedit did not provide rich terminological resources. Since then, large terminologies of medicine and law have been uploaded by developers and users and made public. We have also developed a Web crawler to collect bilingual terminologies from the Web. This resource, consisting of about 2 million terms, will be made public as soon as information cleaning up is complete and copyright status is clarified [9]. An automatically collected compilation of proper-name translation pairs, containing around 400,000 proper-name translations, is also to be made available [10]. In addition, individual users can register terms either manually or using the term extraction and management mechanism, in which users can extract terms from their own or specified text pairs, check the result of extraction, and register them as their own terminological data.

Translation memory: The translations published on MNH are analysed and recycled as a TM resource. Major NGO users, such as Amnesty International and Democracy Now! Japan, have also provided their translations as TM for MNH. We also provide copyright free translation pairs collected from the Web (approximately 1.4 million English words) [11]. These are not currently retrievable from the QRedit small lookup window; another window must be used to retrieve TM. We will integrate TM lookup into QRedit as soon as possible.

6 Current state of usage

We made MNH publicly available on 7th April, 2009. Figure 6 shows statistics on MNH as of 20th October, 2009. The upper panel shows the number of registered users; the middle panel the number of registered documents (the top line shows the total number of documents, including both published and unpublished documents, and the lower line the number of published documents); and the bottom panel the number of registered bilingual term pairs (the top line shows the total number of terms, including both published and unpublished terms, and the lower line the number of published terms).

In the six months since MNH has become publicly available, it has attracted more than 900 registered users, of which 39 publish their translations through MNH. Amnesty International Japan and Democracy Now! Japan, as well as the Japanese translation team of GlobalVoices Online and some members of Translators United for Peace (TUP) are using MNH. A few universities use MNH in their lecture or seminar courses. One publisher, Sakuhinsha, has used MNH on an experimental basis to translate books into Japanese. One title has already been published [12], and others are due to follow.

Though the growth of registration has slowed down recently, we are expecting more users to register as we have just made the site fully compatible with Safari on the Mac and IE8 on Windows (previously, the full range of functions was only available with Firefox or Google Chrome) and are planning to launch a second-stage campaign for attracting NGOs/NPOs as well as publishers involved in translation work.

As for the number of translations, MNH currently has 2,340 translated documents, among which 1,235 have been published on the site. The other documents remain unpublished, or the translators may use them on their own sites, outside MNH. The total of 2,340 documents amounts to 17,066 translation pairs as counted by sentences. Since the end of May, the growth of translated documents has been linear.

Nearly 50,000 bilingual term pairs have been registered to MNH so far. Looking at the bottom panel of Figure 6, we can observe that most of these pairs were registered in April and May, 2009. This is because large terminological lists were uploaded during that period. Although the growth rate of bilingual terms registered to MNH has slowed down, it is still increasing steadily.

7 Conclusions

In this paper, we have introduced the basic features of the translation hosting site Minna no Hon'yaku (MNH) and QRedit, a translation-aid editor provided as part of MNH. The goal of MNH is:

1. to provide a large-scale archive of translations and their originals available to all under clear licenses, which constitutes a translation news site, and
2. to facilitate translation and nurture online translation communities

by providing online volunteer translators with a hosting service for translation activities and an easy-to-use translation-aid editor.

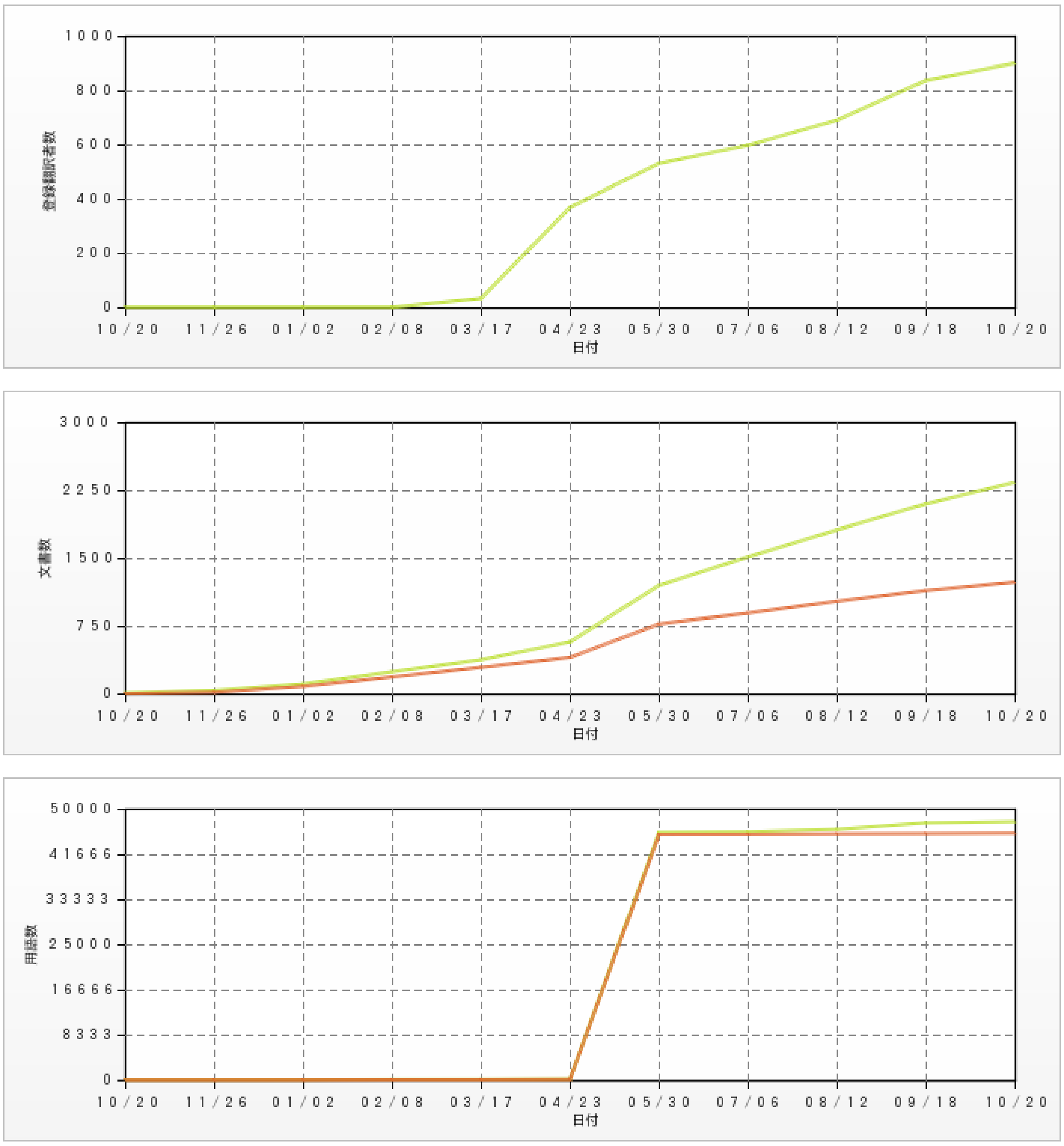


Figure 6. Some statistics on MNH

As mentioned above, according to an informal interview with three MNH users, translation time was reduced on average by 25 per cent, and the quality of the first, draft translation was improved by using MNH. Although we need to carry out systematic evaluation in order to confirm these reports, the fact that the number of users as well as the number of translated documents has steadily increased shows the practical usefulness of MNH. Though seemingly similar services have become available, such as Google Translator Toolkit [13], the translation-aid functions and hosting functions provided by MNH are unique and of higher quality. The basic concept of MNH is, as mentioned in the introduction, to make humans more intelligent by making computers more intelligent, and vice versa. Our emphasis at present is on human translators, but the translations accumulated in MNH are to be used to improve MT and the translation-aid functions of MNH further.

The biggest limitation of MNH for now is that it can only deal with English to Japanese and Japanese to English translations. This limitation of language pairs is inevitable in the short term, because to materialise fully the design concept of MNH, we need high-quality dictionary resources, such as the Sanseido dictionary we are currently using [4]. In the second stage, we will extend MNH to English to Chinese and Chinese to English translation. We hope that by extending the system to cover other languages with the cooperation of dictionary publishers, MNH will provide a social infrastructure for truly universal multilingual communication in the not-so-distant future.

Acknowledgements

This work is partly supported by the Japan Society for the Promotion of Sciences (JSPS) grant-in-aid (A) 21240021 “Developing an integrated translation-aid site which provides comprehensive reference sources for translators” and by the 2009 NII research cooperation project “Construction and use of practical terminological resources from a variety of information sources”.

References

[1] Wikipedia. `http://en.wikipedia.org/`

[2] Alain Désilet, Louise Brunette, Christiane Melanon and Geneviève Patenaude (2008). “Reliable innovation: A tecchie’s travels in the land of translators,” *8th AMTA Conference*, pp. 339–345.

[3] Takeshi Abekawa and Kyo Kageura (2007). “QRedit: An integrated editor system to support online volunteer translators,” *Digital Humanities 2007*, pp. 3–5.

[4] Sanseido (2006). *Grand Concise English-Japanese Dictionary.* Tokyo: Sanseido.

[5] Masao Utiyama and Mayumi Takahashi (2003). *English-Japanese translation alignment data.* `http://www2.nict.go.jp/x/x161/members/mutiyama/align/index.html`

[6] Koichi Takeuchi, Takashi Kanehila, Kazuki Hilao, Takeshi Abekawa and Kyo Kageura (2007). “Flexible automatic look-up of English idiom entries in dictionaries,” *Machine Translation Summit XI Proceedings*, pp. 451–458.

[7] Takeshi Abekawa and Kyo Kageura (2007). “A translation aid system with a stratified lookup interface,” *Proceedings of the 45th ACL Annual Meeting Demos and Poster Sessions*, pp. 5–8.

[8] Jim Breen. EDICT. `http://www.csse.monash.edu.au/~jwb/edict.html`

[9] Takeshi Abekawa and Kyo Kageura (2009). “QRpotato: An exhaustive collector of bilingual terms from the web,” *Proceedings of the 15th Annual Meeting of the Japan Society for Natural Language Processing* (in Japanese).

[10] Satoshi Sato (2009). “Crawling English-Japanese person name transliterations from the Web,” *WWW 2009*, p. 1151–1152.

[11] Masao Utiyama and Mayumi Takahashi (2003). “English-Japanese translation alignment data,” `http://www2.nict.go.jp/x/x161/members/mutiyama/align/index.html`

[12] Sakuhinsha. `http://tssplaza.co.jp/sakuhinsha/`

[13] `http://translate.google.com/toolkit/`