

ITADICT PROJECT AND JAPANESE LANGUAGE LEARNING

Marcella MARIOTTI

Ca' Foscari University of Venice
mariotti@unive.it

Alessandro MANTELLI

Ca' Foscari University of Venice
mantrex@gmail.com

Abstract

This article aims to show how the Nuclear disaster in Fukushima (3 March 2011) affected Japanese Language teaching and learning in Italy, focusing on the ITADICT Project (Marcella Mariotti, project leader, Clemente Beghi, research fellow and Alessandro Mantelli, programmer). The project intends to develop the first Japanese-Italian online database, involving more than 60 students of the Japanese language interested in lexicographic research and online learning strategies and tools. A secondary undertaking of ITADICT is its Latin alphabet transliteration of Japanese words using the Hepburn system of romanization. ITADICT is inspired by the EDICT Japanese-English database developed by the Electronic Dictionary Research and Development Group established in 2000 within the Faculty of Information Technology at Monash University. The Japanese-Italian database is evolving within the Department of Asian and North African Studies at Ca' Foscari University of Venice, the largest in the country and one of the main teaching centres of Japanese in Europe in terms of the number of students dedicated to it (more than 1800) and number of Japanese language teaching hours (1002h at B.A. level, and 387h at M.A. level).

In this paper we describe how and why the project has been carried out and what the expectations are for its future development.

Keywords

ITADICT; Japanese-Italian database; lexicography; Japanese language; online database; collaborative editing; Japanese language learning

Izvešček

Pričujoči članek predstavlja projekt ITADICT (vodja projekta Marcella Mariotti, sodelujoči raziskovalec Clemente Beghi, programer Alessandro Mantelli) in vpliv nuklearne katastrofe v Fukushimi 3. marca 2011 na učenje japonščine v Italiji. Cilj projekta je razvoj prve spletne japonsko-italijanske baze podatkov, pri njem pa sodeluje več kot 60 študentov japonščine, ki jih zanima slovaropisje in učne strategije ter orodja na spletu. Drugi cilj projekta ITADICT je prečrkovanje japonskih besed v latinico, po sistemu Hepburn. Projekt je zastavljen po vzoru japonsko-angleške podatkovne baze EDICT, ki jo je razvila skupina Electronic Dictionary Research and Development Group (skupina za raziskovanje in razvoj elektronskih slovarjev), ki je bila ustanovljena leta 2000 na Fakulteti za informacijsko tehnologijo na Univerzi Monash. Japonsko-italijanska baza podatkov se razvija na Oddelku za azijske in severno-afriške študije na Univerzi Ca'Foscari v Benetkah, eden od glavnih centrov za učenje japonščine v Evropi in

največji v Italiji po številu študentov (1800) in številu učnih ur japonščine (1002 na prvi stopnji in 387 na drugi stopnji študija).

Članek predstavlja ozadje in način izpeljave projekta ter načrte za prihodnji razvoj.

Ključne besede

ITADICT; japonsko-italijanska baza podatkov; leksikografija; japonski jezik; spletna baza podatkov; sodelovalno urejanje; učenje japonščine

1. ITADICT Project and Japanese Language Learning

The ITADICT Project (<http://virgo.unive.it/itadict/eng/about>), is aimed at the creation of a freely accessible Japanese-Italian database, and is expressly inspired by Jim Breen's JMdict/EDICT Project that initiated in 1991 at Monash University.

The database was started separately by both Marcella Mariotti (Ca' Foscari University of Venice) and Clemente Beghi (Ca' Foscari University of Venice) between 2007 and 2008 as part of their research. At the time Beghi was a Ph.D student at Cambridge University doing research on Esoteric Buddhist Iconography, so he edited Buddhist and, for other reasons, Floral terms. In the mean time, Mariotti was a JSPS post-doc researcher at International Christian University (Tokyo), where she needed an Italian translation and transliteration in Latin alphabet of all the words present in her Hypermedia Dictionary of Japanese Grammar BunpoHyDict (Mariotti 2008), so this was her starting point for editing more than 3000 words in the database.

They are both grateful to Jim Breen (Monash University) who brought their research interests together.

ITADICT became one unified voluntary project coordinated by Marcella Mariotti, at the end of 2010, when Beghi and Mariotti were both teaching Japanese Language at the Department of East Asian Studies (now Department of Asian and North African Studies) at Ca' Foscari University of Venice. In one and a half years, they involved more than 60 of their students, who became an integral part of the project, actively translating terms from Japanese into Italian and inserting them in the ITADICT EDITOR later developed by Alessandro Mantelli.

2. Translating students: Why involving them?

The strategic role of pleasure in the long-term acquisition processes of a foreign language has been stressed by neurolinguistics and researchers such as Danesi (2003), Schumann (2006) and Balboni (2002). Moreover, the more social networks, eLearning and mLearning sites and applications spread around on PCs and smartphones, the more students are fascinated, aware of and concerned with the object of their studies,

proving that “learning is like a utility - like water or electricity - that flows in a network or a grip that we tap into when we want”. (Downes 2007)

According to a survey conducted in 2009 (Ferrari, 2010) students at Ca’ Foscari often approach Japanese Language studies not intentionally, as a conscious part of a wider life-plan they have, but more as a way to foster their curiosity, to feel closer to “the real thing”: the original language of loved novels, movies, animated movies, manga, *dorama*, inspired sutras or martial arts.

Maybe due to a disenchantment felt by students of Japanese in Italy in the new century, following the economical crisis of Japan, this emotional motivation is quite specific to the learners of the net-generation (Mariotti, forthcoming, Miyake 2012), while far removed from those of the Nineties.

The above may explain why, as soon as Mariotti introduced ITADICT project to her students as part of a presentation about fansubbing and language learning (Mariotti, 2011), many of them were willing to participate. Their motivations were diverse: mainly they were interested in creating a tool for translating Japanese into Italian using a mouse-over dictionary¹, and in learning strategies to use the online research tools on Japanese language sites to conduct lexicographical research (Mariotti 2012²).

A secondary motivation followed. As of 2012, Ca’ Foscari students need to complete a period of internship before they can graduate. In 2011 this internship was “warmly suggested”, and students received 5 or 6 University Credits for it. Particularly because of the nuclear incident in Fukushima following the tragic 3.11 earthquake and tsunami in North Eastern Japan, Ca’ Foscari students were discouraged from applying for an internship in Japan as they had usually done. As a result, 42 out of the 64 students collaborating on ITADICT were 2011 prospective graduates who did not know where to complete their internship and, intrigued by the ITADICT project, chose to do so by taking part in the project.

3. What is ITADICT?

As mentioned, ITADICT was born out of different needs, largely mirroring those of the original Japanese-English [EDICT](#) project, which started in 1991 from MOKE (Mark’s Own Kanji Editor), a word processor with integrated Japanese-English Dictionary (Breen 2010). ITADICT project focused on creating a file for a Japanese-Italian database that could be used by third party software to easily read and translate Japanese texts (e.g. [Rikaichan](#) – popup dictionary tool for Firefox browser-, [Japan](#)

¹ E.g. <http://www.polarcloud.com/rikaichan/>

² Online anonymous survey *Why ITADICT?*, addressed to the 66 students and collaborators who worked on ITADICT (Sept. 2012).

[Goggles](#) -iPhone app to translate words from live camera-, [Kotoba/Imiwa?](#) – iPhone dictionary-app to manage Japanese-Other Languages databases-, and more).

In 1999 the EDICT project, which had been limited by very a simple dictionary structure, evolved into the more complex [JMDict Project](#) JMDict (Japanese-Multilingual Dictionary) Project (managed by the EDRG Electronic Dictionary Research Group). (Breen 2004)

JMDict employs an XML structure to support a much richer dictionary entry format including multiple kanji surface forms and readings. The original EDICT format is generated from this project as a legacy format mainly for older software packages. An expanded “EDICT2” format is also generated which more closely follows the XML content. For our purposes we started using the simpler “Traditional” EDICT file where there is only 1 kanji form and 1 reading per entry/line in plain text, with less marking of different semantic fields than in the newer [EDICT2](#):

KANJI [KANJI] /(PoS tag) gloss/gloss/...

The file had about 160.000 entries (with one line per entry), where most common entries had a (P) mark for “priority” at the end of the line. [Breen’s online pages](#) describe the process used to determine the priority of a term, mainly marked after a) Alexandre Girardi’s (NAIST-MULTITEL) match-analysis between EDICT entries, the 1994-1998 corpus of Mainichi Shinbun, and b) the 10,000 common words in the collection *Ichimango goi bunruishū* (Senmon kyōiku Publishing 1998). Although, as Breen underlines:

While the priority markings accurately reflect the status of entries with regard to the various sources, they must be seen as only providing a crude indication of how common a word or expression actually is in Japanese. The “(P)” markings in the EDICT and EDICT2 files appear to identify a useful subset of “common” words, but there are clearly some marked entries which are not very common, and there are clearly unmarked entries which are in common use, particularly in the spoken language. (Breen, 2010)

3.1 Latin alphabet transliteration according to the Hepburn system

Since our purpose was to allow as many people as possible to approach the Japanese language and to enable, let us say, a primary school teacher to say some words to a Japanese-native-speaker child at school even without knowing *kanji* or *kana*, we added a new characteristic to Breen’s EDICT format: the Latin alphabet transliteration of the *hiragana* readings in brackets. ITADICT line, then, appeared as follows:

KANJI [KANJI, latin] /(PoS tag)/gloss/gloss/...

おい付く [おいつく, oitsuku] / (v5k,vi) raggiungere/uguagliare/arrivare al livello di/(P)/

This decision was followed by a heated debate, and since the database was developed inside university academia, and the whole process itself was part of a teaching/learning project, we adopted the Hepburn system of transliteration and did not rely on automatic transliteration tools (e.g. [Romaji Translator](#)), but rather chose to manually transcribe each entry.

3.2 Use of monolingual JA-JA dictionary

With the intent to not only produce an accurate Japanese-Italian database for the general user, but also to offer our students a professionalizing experience and autonomous learning strategies, we encouraged them to refer to online and offline monolingual Japanese-Japanese dictionaries and discussion groups (e.g. [Yahoo's Chiebukuro](#) or [kotoba.ne.jp](#)), above and beyond utilizing them to only translate from English. This was intended to avoid “false friends” as well, which are quite numerous in English and Italian, such as 春分 *shunbun*, translated into English as “vernal equinox” and mistakenly translated into the Italian “equinozio d’inverno” (winter equinox), instead of “equinozio di primavera” (spring equinox). Further explanations about checking entered translations through the ITADICT Editor are given in section 5 of this article, dedicated to ITADICT Editor.

4. How was the project organized?

4.1 Repartition of the “traditional” EDICT file

The work with our students started with the extraction of 18626 priority words (P) from the “traditional” EDICT file, resulting in a 2.8 MB .txt file that was split into 18 smaller files of 1000 entries each. Students volunteering were assigned 250 entries each, while internship students were assigned 1000 entries each. The former had a very flexible deadline, while the latter had to complete the internship in 3 months.

Assigned entries were sent to the students as an .rtf file e-mail attachment, with lines numbered from 1 to 1000 for each “priority” file. (Figure 1)

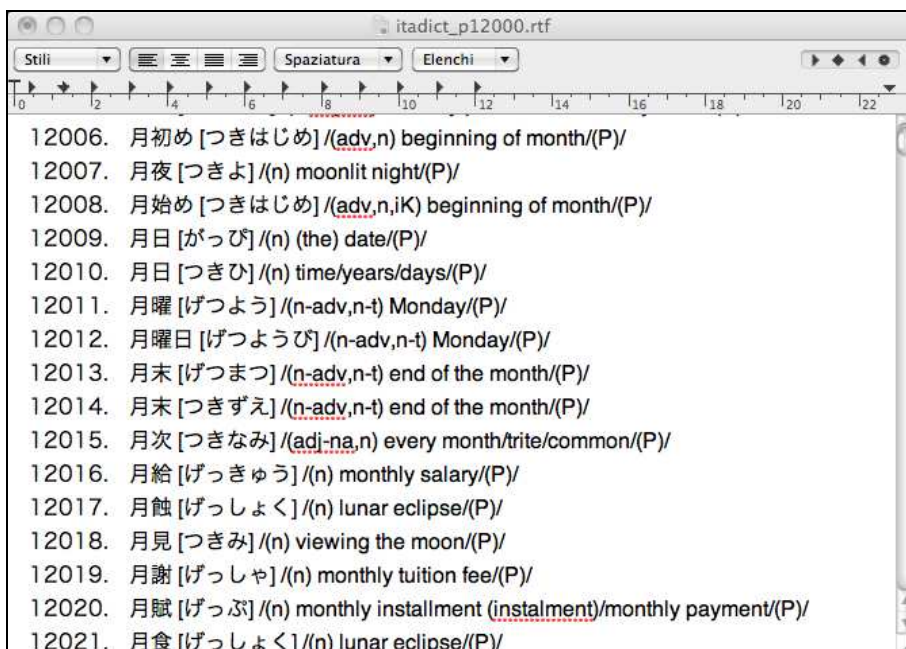


Figure 1: Partitioned .rtf file of (P)riority words with numbered lines

The overall exchange of files was managed on a shared online google spreadsheet called *Ripartizione ITADICT* created by Mariotti on November 3, 2010 (Figure 2).

The spreadsheet included the following information:

- student's name and surname,
- assigned file or portion of file,
- deadline of the work,
- first/last line to translate,
- delivered date,
- reviewed status,
- supervisor of the (later) import in the online new EDITOR developed by Mantelli,
- private e-mail (upon written agreement, so as to be able to contact the student-translator even in the future),
- grade (undergraduate or graduate),
- supervisor of the delivered entries.

| Nome | file | termini/DATE STAGE | primo-ultimo termine: | consegnato | CORREZ | importazioni MAIL | qualifica | revisore |
|---------------|-----------|--------------------|-----------------------|--------------|--------|-------------------|-----------|--------------|
| itadictP | 751-1000 | | | consegnato | OK | km | | Lt 1 dessi-m |
| itadictP | 1-250 | | | consegnato | OK | km | LM2 | m |
| itadictP | 501-750 | | | consegnato | OK | km | | b: A.A. |
| itadictP2001 | 2751-3000 | | | consegnato | OK | km | LM2 | m |
| itadictP2001 | 1-250 | | | consegnato | OK | km | LM2 | m |
| itadictP2001 | 251-500 | | | consegnato | OK | km | LM1 | m |
| itadictP2001 | 501-750 | | | consegnato | OK | km | LM1 | m |
| itadictP1001 | 1-250 | | | consegnato | OK | ka | | Lt 3 b |
| itadictP1001 | 251-500 | | | consegnato | OK | km | LM2 | m |
| itadictP1001 | 501-750 | | | consegnato | OK | km | LM1 | m |
| itadictP1001 | 1751-2000 | | | | | | | |
| itadictP1001 | 1001-2000 | 1001:urania- | 1250:sunroof | consegnato | OK | kb | | Lt 3 b |
| itadictP3001 | 1-250 | | | consegnato | OK | km | LM1 | m |
| itadictP3001 | 251-500 | | | consegnato | OK | km | | Lt 3 m |
| itadictP3001 | 501-750 | 500: passion - | 750: fuzzy | consegnato | OK | km | | m |
| itadictP3001 | 751-1000 | | | consegnato | OK | km | | m |
| itadictP4001 | 1-250 | | | consegnato | OK | km | | |
| itadictP4001 | 251-500 | | | consegnato | OK | km | | |
| itadictP4001 | 501-750 | 501: marinba 750: | monpani | consegnato | OK | km | | LM2 |
| itadictP5001 | 1-250 | | | cons | | kb | | b |
| itadictP6001 | tutti | | | cons | OK | km | | Lt 3 |
| itadictP7000 | 1-1000 | 1 冷性, 1000 时 | | cons | | kb | | Lt 3 b |
| itadictP8000 | tutti | firmato | | cons | OK | km | | Lt 3 |
| itadictP9000 | tutti | 9000:妻子, 9999:弘い | | CONS 15-5-11 | | km | | Lt 3 |
| itadictP10000 | 1-250 | 1.弘夫(どうだい) | 250: 御方(みかた) | consegnato | | kb | laureata | magistrale b |
| itadictP10000 | 251-350 | ?? | | consegnato | | kb | laureata | magistrale |
| itadictP10000 | 350-500 | | | | | kb | laurea | |
| itadictP10000 | | 501:osorenu - 999: | | | | | | |

Figure 2: Ripartizione ITADICT's spreadsheet

Some students participated more than one time; in this case, e-mail and personal data were inserted only the first time they registered on the form.

4.2 Online discussion group

Since there were very important rules to follow (e.g. the entry syntax, according to which semantic fields had to be divided by a slash without space before or after it), and some questions (e.g. whether to add more specific information about artists' names or not, or which ISO had to be used for source language of *katakana* words) were to be decided by involving all the contributors, on October 23, 2010 we opened a forum (googlegroup Itadict), where more experienced students directly trained by Mariotti could help newcomers, and where the community could discuss how to manage problematic matters (neologisms, technical terms, fields of pertinence...), share useful links and collect/answer queries.

As of September 15, 2012, the dedicated googlegroup *ITADICT* collected 464 messages from its 86 members, who participated in 155 discussions. The first conversations, up to March 26, 2011 were mainly about:

1. how to correctly open the .rtf files,
2. how to find and fill in the documents for internship registration,

3. how to manage the repetition of similar entries with diverging pronunciation such as the example below, a problem solved by the most recent format JMDict (J apanese Multilingual Dictionary by Breen).
 - a. 付いている [ついている] /(exp,uk) to be lucky/to be in luck/(P)/
 - b. 付いている [ついてる] /(exp,uk) to be lucky/to be in luck/(P)/
 - c. 付いてる [ついている] /(exp,uk) to be lucky/to be in luck/(P)/
 - d. 付いてる [ついてる] /(exp,uk) to be lucky/to be in luck/(P)/
4. how/where to look up unknown terms
5. how to transliterate a *kanji-kana* compound word: separately or not? Adding source language or not? How?... (e.g. イベント処理 [イベントし
ょり , *ibento shori*] /(n) elaborazione degli eventi, event processing
{comp} (en: event (+ shori))
6. how to transliterate words into Latin characters following Hepburn Style.
7. ... and more.

We finally agreed to combine and summarize all collected answers into one “guideline document” (edited by Silvia Rivadossi and Marina Monego in March 2011): this enhanced and sped up the workflow. But since the number of students involved was exponentially increasing, the need for an online editor became evident. Revised instructions were posted by Piercarlo Tommasi in November 2011, after the Editor was developed.

5. The ITADICT Editor and the front-end user interface

Alessandro Mantelli, consolidated system engineer and adjunct professor of informatics at Ca’ Foscari University of Venice, volunteered to conceive and develop an appropriate web editor (ITADICT Editor) that would enable us easily to:

- a) maintain the original *kanji/kana* - English database and “Part of Speech” classification³
- b) import data from all the translation entry files produced until then,
- c) monitor collaborators’ work by login and password furnished upon request,
- d) check and where necessary modify each entered translation,
- e) monitor and coordinate the trainees showing:

³ English Part of Speech labels were not translated into Italian, but rather left in English so that third party applications already conforming to Breen’s Jmdict format would be able to process the ITADICT database as well. (E.g. Japanese Verb Conjugator Katsuyo)

- a. name of the translator of the entry
 - b. time of the entry
 - c. modification history
 - d. number of entries translated by each student
 - e. entries left to be translated
- f) filter the database by translator's name (so as to monitor the work done) and other useful fields (see. "how to search online")
 - g) further broaden the database
 - h) share the ITADICT search engine online with the world
 - i) export the database of Italian translations and of Latin transcriptions in order to merge it into third party applications (e.g. Rikaichan)

The following screenshots show each of the above features.

| Rif | 7ndcode | creato da | kanji | english | imp | romaji | Italiano | stato |
|-----|---------|-----------|---------------|---|-----|-----------|---|-------|
| 1 | 2197150 | marcella | × [ゼツ] | (n.uk) x-mark (used to indicate an incorrect answer in a test, etc.)/impossibility/futility/uselessness | | batsu | segno x (usato per indicare una risposta errata in un test. | |
| 2 | 0 | marcella | × [セツ] | (n.uk) x-mark (used to indicate an incorrect answer in a test, etc.)/impossibility/futility/uselessness | | peke | segno x (usato per indicare una risposta errata in un test. | |
| 3 | 0 | marcella | × [セツ] | (n.uk) x-mark (used to indicate an incorrect answer in a test, etc.)/impossibility/futility/uselessness | | peke | segno x (usato per indicare una risposta errata in un test. | |
| 4 | 1019520 | Rivadossi | A [アルファ] | (n) alpha | p | arufa | alfa (gr: alpha) | |
| 5 | 1119290 | Rivadossi | B [ベータ] | (n) beta | p | beta | beta (gr: beta) | |
| 6 | 2173670 | marcella | E [エプシロン] | (n) epsilon | | epushiron | epsilon | |
| 7 | 0 | admin | E [エプシロン] | (n) epsilon | | epushiron | epsilon | |
| 8 | 2173680 | marcella | Z [ゼータ] | (n) zeta | | zeta | zeta | |
| 9 | 2173690 | Barile | H [イータ] | (n) eta | | ita | eta | |
| 10 | 0 | Barile | H [イータ] | (n) eta | | eta | eta | |
| 11 | 2173700 | Barile | Θ [シータ] | (n) theta | | shi ta | theta | |
| 12 | 0 | Barile | Θ [シータ] | (n) theta | | theta | theta | |
| 13 | 2173710 | Barile | Θ関数 [シータかんすう] | (n) theta function | | thakansu | funzione theta | |
| 14 | 2173720 | Barile | I [イオタ] | (n) iota | | iota | iota | |
| 15 | 2173730 | Barile | K [カッパ] | (n) kappa | | kappa | kappa | |
| 16 | 0 | marcella | K [カッパ] | (n) kappa | | kappa | kappa | |
| 17 | 1139850 | marcella | Λ [ラムダ] | (n) lambda | | lamuda | lambda | |
| 18 | 2173740 | marcella | M [ミュー] | (n) mu | | myu | mu | |
| 19 | 2173750 | marcella | N [ニュー] | (n) nu | | nyu | nu | |
| 20 | 2173770 | marcella | Ξ [クサイ] | (n) xi | | cusai | xi | |

Figure 3: ITADICT Editor page

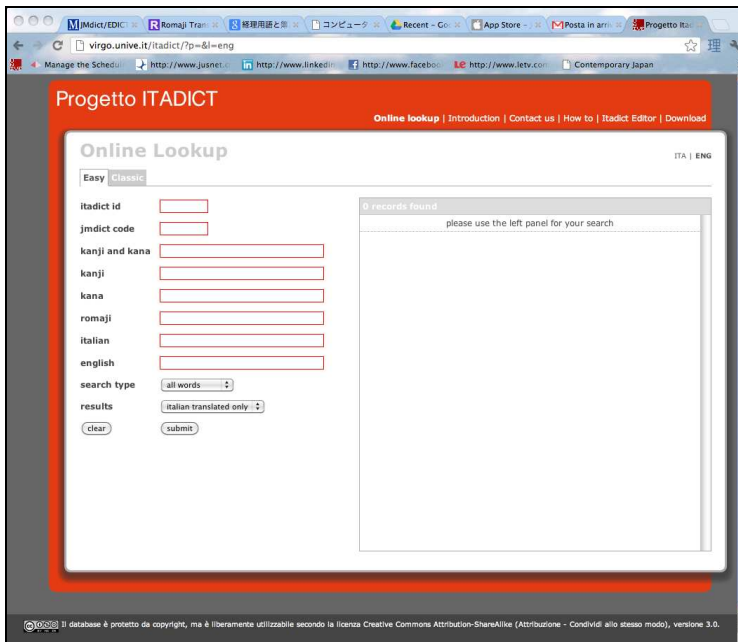


Figure 4: ITADICT online search interface (by code, by kanji and kana, by kana, by romaji, by Italian, by English, filters: only terms translated into Italian or All terms).

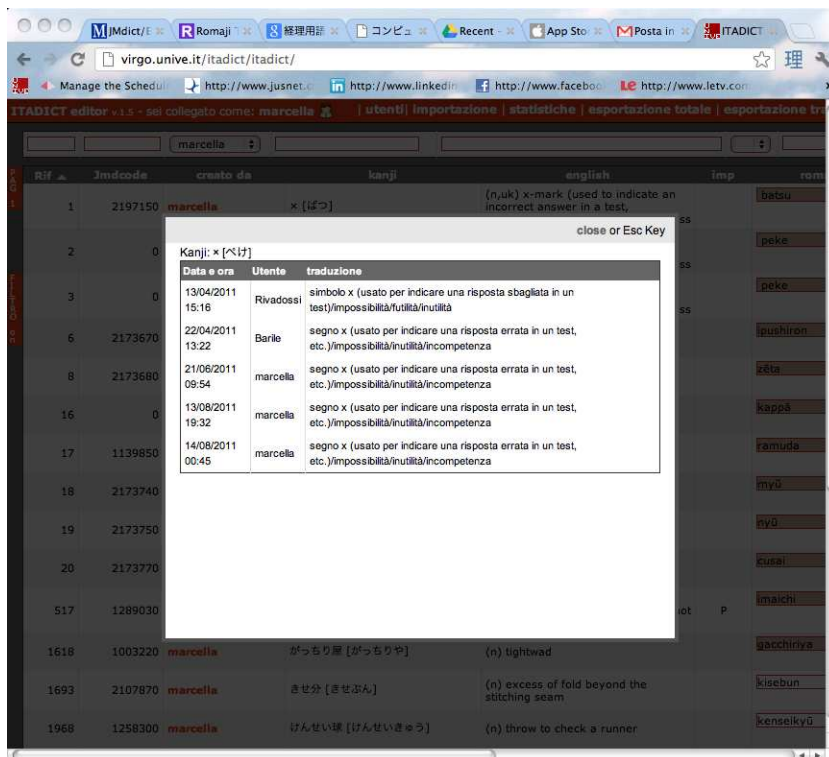


Figure 5: History of each inserted term (date, time/ user/ translation history)

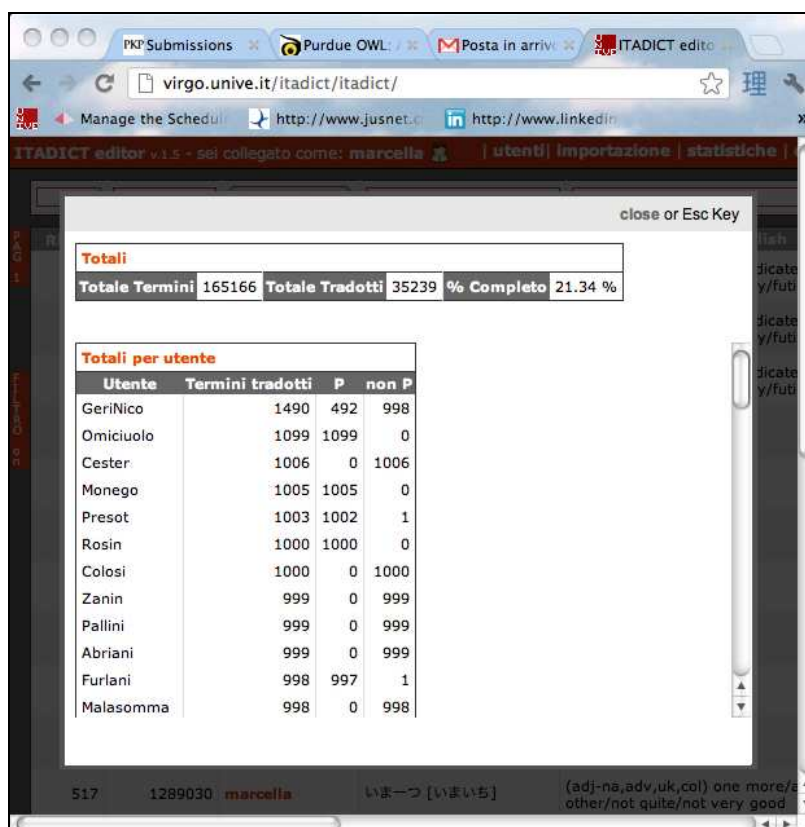


Figure 6: Statistics (Total terms, translated, % completed, terms by translator -priority/non priority)

6. Technical considerations

6.1 Php language to override static html

From a data management point of view, one of the main ITADICT project goal is to provide a scalable development architecture as well as an easy way to enter and edit dictionary data.

Considering the limitations of a traditional html-based approach that lacks a database structure and advanced conditional procedures, ITADICT was developed using a client server model architecture.

A client server application is a piece of software that runs on a client computer and makes requests to a remote server. In ITADICT, the client is the browser in which the web url is invoked such as Firefox or Chrome. From a client perspective, nothing changes compared with a traditional html page: the client reads pages from a server (a remote computer where pages are stored) and displays them. If there are javascript procedures, they are executed as well. From a server point of view, however, there are two components that cause significant changes in the process of creating the page: an

advanced and widely used server side language called *php*, and a database (data organizer system) called *mysql*. These two components are normally invoked before the http service (a piece of software installed in the server that provides the web page⁴). In ITADICT the http service used is called *apache*⁵ and is the most widely used web service⁶.

Thanks to a server side language such as *php*, *static* limits of *html* can be overridden. Programmers can use a high level language to produce software and call on the data in the database through direct *sql* invocation. The page is parsed by the *http* server and displayed in the browser.

ITADICT database contains all the dictionary data organized in a sql table schema. Vocabulary data can be entered using a dedicated editor whose access is granted by login and password. Kanji data is based on the original EDICT dictionary and registered in the database in binary and base 64 format⁷.

6.2 ITADICT basic architecture

Client – Server data communication follows two approaches:

1. Traditional or synchronous method: Data is sent to the server. Page content is rebuilt and presented to the browser.
2. Ajax or Asynchronous Method: Data is sent to the server. Only the needed portion of the page is recreated and passed asynchronously (in the background) to the page, to substitute for old content in real time without interfering with the display and behavior of the existing page.

The client-server approach enables us to separate the interface from the data layer, allowing us to create different interfaces to be connected to the same database in the future.

The asynchronous method is especially used in the front-end (publicly available) user interface where the interaction with the server is smooth and dynamic. In place of the traditional browser loading status, an animated icon is displayed when data communication between client and server occurs.

As data management is separated from the client interface, it is possible to create different ITADICT interfaces such as mobile applications that connect to the same database.

⁴ Without an http service or web server no page can be displayed on a browser. Every page of the World Wide Web is provided by an http service.

⁵ http://projects.apache.org/projects/http_server.html

⁶ Netcraft survey July 2012: <http://news.netcraft.com/archives/2012/07/03/july-2012-web-server-survey.html>

⁷ Traditional varchar format has been avoided to avoid misleading utf8 conversions.

The following diagram represents ITADICT's basic architecture.

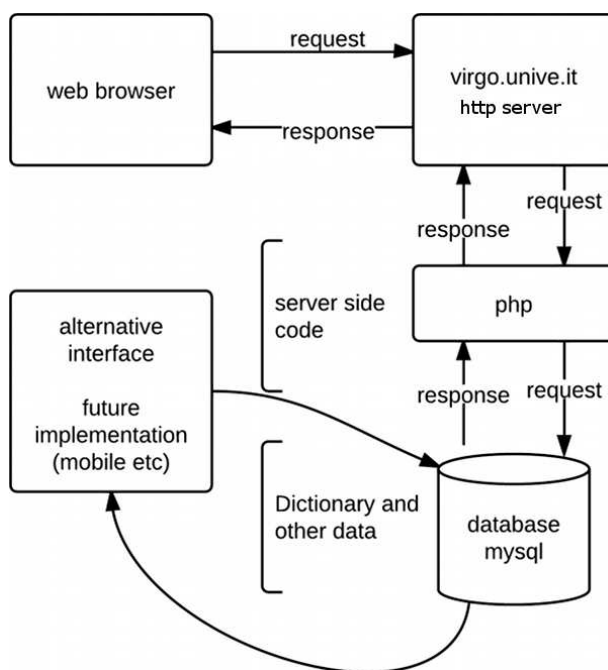


Figure 7: ITADICT's basic architecture

6.3 Future system implementation

ITADICT is a system still in development. Currently we are moving from the original EDICT data source to the more advanced JMdict data format. The database structure itself is going to change and new features - such as adding custom terms - are going to be implemented shortly in the editor.

From a front-end point of view, we are planning to expand search functions and include an svg render engine for the displayed kanji.

7. What did ITADICT internship offer to the students?

Working on ITADICT (.rtf file or Editor) helped students acquire several skills, beside learning new English and Japanese words, such as:

- how to use shared documents,
- how to use online Japanese searching engines,
- how to use a monolingual Japanese-Japanese dictionary,

- how to autonomously translate terms not present in Japanese-Italian dictionaries,
- how to thoroughly adhere to strict editing rules,
- how to ask for help from colleagues and via online translation forums (e.g. Biblit, Langit, WordReference and more),
- how to use online discussion forums,
- how to open a new e-mail account,
- how to correctly transliterate Japanese into Latin alphabet according to the Hepburn system.

Besides these acquired skills, one of the major satisfactions offered by the ITADICT internship is the opportunity to actively participate in the creation of a language tool the students themselves will use in the future.

The database of the translated terms can be used online as well as downloaded and merged into third party applications through a [dedicated link](#)⁸.

8. Conclusions and future work

Up to September 2012 about 35239 entries have been transliterated into Latin characters (*romaji*) translated and (most but not all) revised. The most common 18.681 entries (marked with a “priority” P in the database) have been completed and, as mentioned above, we are now proceeding to match the “traditional Edict” format we started from, with the newer JMdict format. This is meant primarily to avoid misleading readings such as the pair mentioned above, where ついてる (*tsuiteru*) is actually not the transcription of the entry 付いている (*tsuite iru*):

付いている [ついてる, tsuite ru] /(exp,uk) to be lucky/to be in luck/(P)/

We are planning to reach 50.000 terms by the beginning of 2013 and then reach our final goal of 180.000 terms hopefully within two more years. Special projects are underway, such as the development of lists of specialized words (nanotechnology, leather processing industry, textile industry, furniture industry⁹) and a more user friendly layout of the open search engine reachable at virgo.unive.it/itadict.

The exported database can be used with open source dictionary software and reading aid software the same way as Edict files can. The file is in Unicode/ISO-10646 coding with UTF-8 encapsulation.

⁸ <http://virgo.unive.it/itadict/?p=download&l=eng>

⁹ As part of a European Social Fund project selected from Ca' Foscari University to be presented to Veneto Region. Title: “Language tools to support the internationalization of Veneto Region's companies: terminology in the textile, leather, chemical-environmental, securities and furniture industries” (Project Leader: Marcella Mariotti).

The authors welcome new collaborators: anyone interested in the project can contact us using the online format: we will be happy to further improve our project.

List of collaborators from whom we received the disclosure agreement.

| | |
|------------------------------|-------------------|
| Nicola Angaran | Andrea Giolai |
| Nicolò Anesa | Alessandra Grillo |
| Alice Aniello | Alberto La Spada |
| Andrea Belluomini | Martina Malasomma |
| Antonio Benasaglio Berlucchi | Lara Marinozzi |
| Alice Berto | Elisa Martini |
| Giulia Bianco | Marina Monego |
| Stefano Boggia | Alessia Omiciuolo |
| Rita Bovina | Elena Ominetti |
| Paola Celentano | Chiara Pallini |
| Marco Colosi | Giulia Perin |
| Matteo Contrini | Giovanni Presot |
| Elia Dal Corso | Silvia Rivadossi |
| Massimo Dalla Pria | Elena Tessari |
| Mauro Dalle Prane | Jacopo Tiezzi |
| Giorgia Dessì | Federico Tombari |
| Giacomo Orseolo Ferro | Aurora Torchia |
| Valentina Gastaldello | Cinzia Zanin |
| Niccolò Geri | |

References

- Balboni, P. E. (2002). *Le sfide di Babele*. UTET.
- Breen, J. (2004). JMDict: a Japanese-multilingual dictionary. In: *Coling 2004 workshop on multilingual linguistic resources*, Geneva, Switzerland, pp. 71-78.
- Breen, J. (2010). JMDict/EDICT Japanese/English Dictionary Project
http://www.csse.monash.edu.au/~jwb/edict_doc.html#IREF05
- Danesi, M. (2003). *Second Language Teaching, A View from the Right Side of the Brain*. Dordrecht: Kluwer.
- Downes, S. (2007). *Trends and Impacts of E-Learning 2.0*. Keynote presentation delivered at the International Conference on Open Course Ware and e-Learning, Taipei, Taiwan.
- Ferrari, M. (2010). *La didattica del giapponese a studenti universitari italiani attraverso la canzone. Analisi di un ciclo di tre lezioni sperimentali sulla canzone all'Università Ca' Foscari di Venezia*. (Japanese language teaching to Italian university students. Analysis of

three experimental classes at Ca' Foscari University of Venice) Master Thesis (tutors: Mariotti, M. & Balboni, P.). Ca' Foscari University of Venice.

Mariotti, M. (2011). "Fansubbing e didattica: vecchie e nuove sfide e Tecnologie" (Fansubbing and language learning/teaching: old and new challenges and technologies", paper presented at the workshop *La traduzione dalle lingue orientali: limiti, specificità e prospettive*, (Translating from East-Asian Languages: limits, specificities and perspective, Nicoletta Pesaro coordinator), Ca' Foscari University of Venice, 15 February 2011.

Mariotti, M. (Forthcoming). *Introduzione alla lingua giapponese* (Introduction to Japanese Language), Carocci.

Mariotti, M. (2011). BunpoHyDict: a Hypermedia Dictionary of Japanese Grammar and its Development. In: Association of Japanese Language Teachers in Europe (AJE), *Japanese Language Education in Europe (Yoropa Nihongo Kyoiku)*, 15, pp. 179-188.

Miyake, T. (2012). "Japanese Studies as 'subculture'", paper presented at the conference "Manga Worlds": *Subcultures, Japan, Japanology*, Kyoto International Manga Museum, 3 June 2012.

Schumann, J. H., et al. (2006). *The Neurobiology of Learning*. Routledge.

Quoted links

Breen's *Lexicographical* Details:

http://www.csse.monash.edu.au/~jwb/edict_doc.html#IREF05

BunpoHyDict (under construction):

www.bunpohydict.com

Chiebukuro:

<http://chiebukuro.yahoo.co.jp/>

Goo Japan:

goo.ne.jp

Introduction to EDICT (2005):

http://www.csse.monash.edu.au/~jwb/edict_doc_old.html

Introduction to EDICT within the JMdict Project (2012):

<http://www.csse.monash.edu.au/~jwb/edict.html>

ITADICT database download:

<http://virgo.unive.it/itadict/?p=download&l=eng>

ITADICT:

<http://virgo.unive.it/itadict>

Japan Goggles:

<http://japangoggles.lucsens.com/>

Kotoba.ne.jp:

<http://www.kotoba.jp/>

Kotoba/Imiwa?:

<http://imiwa.pierrephi.net/>

Online survey *Why ITADICT?*":

<https://docs.google.com/spreadsheet/ccc?key=0Anfm7ZNoCqUCdHhxVVJLU3FjZXNBW19iU3RQRUNHREE#gid=0>

Rikaichan:

<http://www.polarcloud.com/rikaichan/>

Romaji Translator:

<http://www.romaji.org/>

The “traditional” EDICT:

http://www.csse.monash.edu.au/~jwb/edict_doc.html

The Jmdict Project (2012):

<http://www.csse.monash.edu.au/~jwb/jmdict.html>

