Software Localization: Modern or Messy?

by Marta Bartnicka, IBM

- Have you ever translated software?
- Have you ever translated software unknowingly?
- Have you ever translated text in Microsoft Excel?
- Have you ever used a CAT tool?

Translation vs. Software Localization

1. The subject domain is defined up front (e.g. IT, medical, legal…)
2. The source document is (almost) ready when delivered to the Translator
3. Text formats are well known (e.g. MS Word) and CAT-supported
4. The Translator knows the context and all the source text is translatable
5. The final result of the translation is visible after validation/review

1. The subject domain may be fuzzy (e.g. mixed IT, medical & legal)
2. The source text is a work in progress, can be updated multiple times
3. Text formats are either IT-specific or MS Excel (still OK for CATs)
4. The context may be hard to find and the source contains $variables
5. The final result of the translation is visible during software test

Software is everywhere

- Medical appliances, online shops, cars, and nuclear power plants
- Computers, mobile devices, and custom displays
- Translators should be notified about the subject domain up front

Software is everywhere

An agricultural computer program contains a long list of crops. There are two crop groups that are recorded as 'melon' and 'cherry'. The first group consists mainly of watermelons, but also of other melons. The second one includes Polish 'wiśnia' and 'czereśnia'. Are there any common names?

(I am actually translating this from Finnish to Russian, but the problem is exactly the same.)

Agile and Continuous Delivery

- See how apps on your smartphone update regularly
- Now imagine that someone translates these strings ☺
  - or else a machine will do it
- Customers also want business applications updated this way
  - and the gaming industry is into updates as well

Agile and Continuous Delivery

Translation: lqe\admin\nls\EditDataSourcePageMessages.js

```
'pauseIndexing': "Wstrzymaj indeksowanie",
'resumeIndexing': "Wznów indeksowanie",
'save': "Zapisz",
'accessControl':'<TWBSQ>Kontrola dostępu</TWBSQ>',
'aclUrl':'<TWBSQ>Adres URL listy kontekstów dostępu:</TWBSQ>',
'authentication': "Uwierzytelnianie",
'authType':"Typ uwierzytelniania:",
'authTypeJlipSSO': "Użyj funkcji pojedynczego logowania platformy JSA dla tego źródła danych",
'authTypeJtsPreconfigured': "Użyj wstępnie skonfigurowanego uwierzytelniania produktu Jazz Team Server dla tego źródła danych",
'authTypeOAuth':"Podaj szczegóły uwierzytelniania OAuth dla tego źródła danych.",
'authTypeNoauth':"To źródło danych nie wymaga uwierzytelniania.",
'authDiscovered':"Uwierzytelnianie tego źródła danych jest wstępnie skonfigurowane. Aby podać inne szczegóły uwierzytelniania, kliknij ${0}tutaj${1} (to nie jest
typowa czynność).",
'oauthConsumerKey':"Klucz konsumenta OAuth:",
'oauthConsumerSecret':"Hasło konsumenta OAuth:",
'oauthRequestTokenUrl':"Adres URL znacznika żądania OAuth:",
'oauthUserAuthorizationUrl':"Adres URL uwierzytelniania użytkownika OAuth:",
'oauthAccessTokenUrl':"Adres URL znacznika dostępu OAuth:",
'settings': "Ustawienia",
'refreshRateSeconds':"Częstotliwość odświeżania (w sekundach):",
'refreshRateMessage':"Częstotliwość odświeżania źródła danych.",
'baseFetcherThreads':"Number of threads for fetching during first time indexing:",
'basePgsThreads':"Number of threads for property graph work during first time indexing:",
'baseFetcherThreadsMessage':"The lifecycle query engine uses the specified number of threads for fetching during initial indexing. Zalecane jest wybranie od 2 do 4
wątków.",
'basePgsThreadsMessage':"The lifecycle query engine uses the specified number of threads for property graph work during initial indexing. Zalecane jest wybranie od
2 do 4 wątków.",
```

Translation Memory: CL606AAP001 (CL606AAP001_0_en-US_pl-PL CL606AAP001_POL_SH1_MT)

```
0 - Number of threads for fetching during first time indexing:
1[f] - Liczba wątków do indeksowania początkowego:
2[f] - Liczba wątków do indeksowania początkowego:
3[f] - Liczba wątków podczas pierwszego indeksowania
```

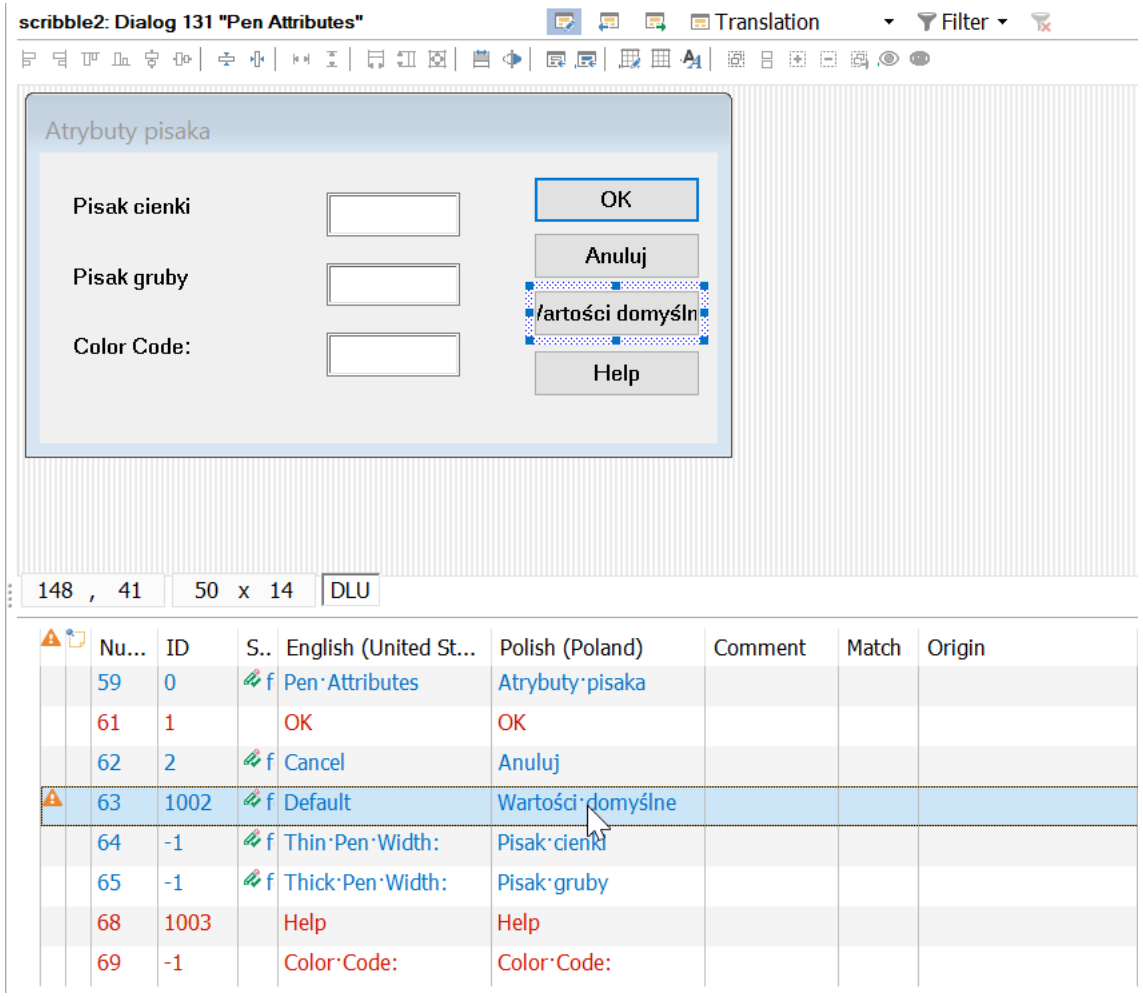
Source of Proposal(s)

```
Number of threads for fetching during first time indexing:
1[f] - Number of threads for ? first time indexing:
2[f] - Number of threads for ? first time indexing:
3[f] - Number of threads for ? first time indexing?
```

CATs, programming formats, and Excel

- The majority of CATs can support typical programming formats
- Many software houses try to develop their own CAT-like tools
- Excel is still the format of choice
  - often the worst choice
- Some CATs visualize user interface
  - yet hardly any software nowadays is written in a way that would enable this CAT function to work properly

CATs, programming formats, and Excel



	A	B	C
1	String ID	EN	DE
2	str_quest_quest_wiping_gangs_button_3	Attack	
3	str_message_not_enough_warriors_button	Recruit	
4	str_message_all	All	
5	str_quest2_header	Order	
6	str_quest_quest_firestation_upgrade_2_button	Start	
7	str_village_move	move	
8	str_quest_quest_warehouse_upgrade_2_header	Start	

Meet Mr. Comment and Mrs. $variable

- Translatable text may consist of fragments or single words
- Context often hides in comments and so they should be visible
- Tags also carry important information and thus must not be hidden
- Moreover, $variables must be placed correctly in the translated text

Meet Mr. Comment and Mrs. $variable

```
'milliseconds': "ms",
'ftsString': "Full Text Search String",
// The userAtTime string will be used to specify which user ran a query and at what time.
// For example: "johndoe at 5:30 PM"
'userAtTime': "${0} at ${1}",
'node': "Node: ${0}",
'sparql': "Query",
```

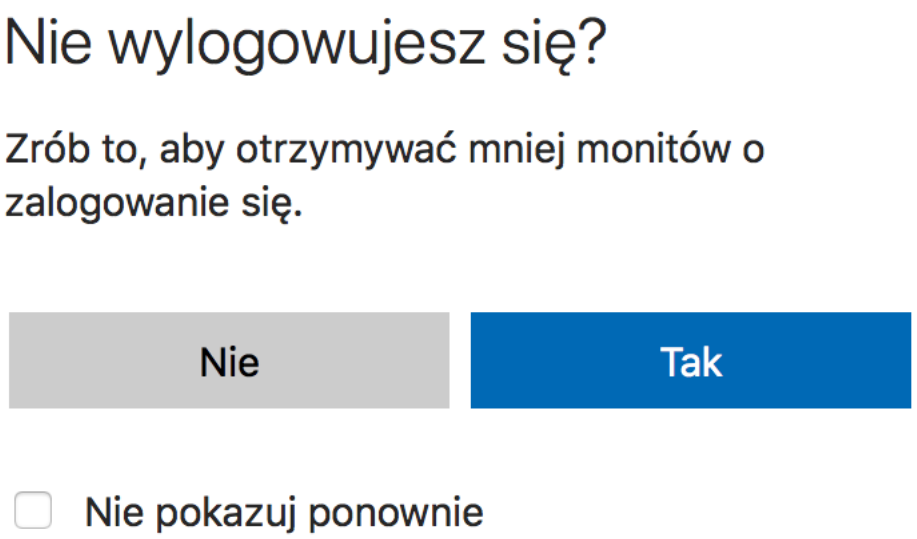
Techwriter.pl
7 maja o 08:20 ·

Kolejny wpis gościnny od Marta Bartnicka. Mamy nadzieję, że umili Wam dzisiejszy ciężki powrót do rzeczywistości po majówkowym leniuchowaniu 🙂

No matter how strange the source is…

- The goal of software localization is to save end user effort
  - not increase it
- The translated phrases should be self-inclusive
- Terminology and grammar choices should be consistent
- Word-by-word job is not an option
  - that’s already taken over by MT

No matter how strange the source is…



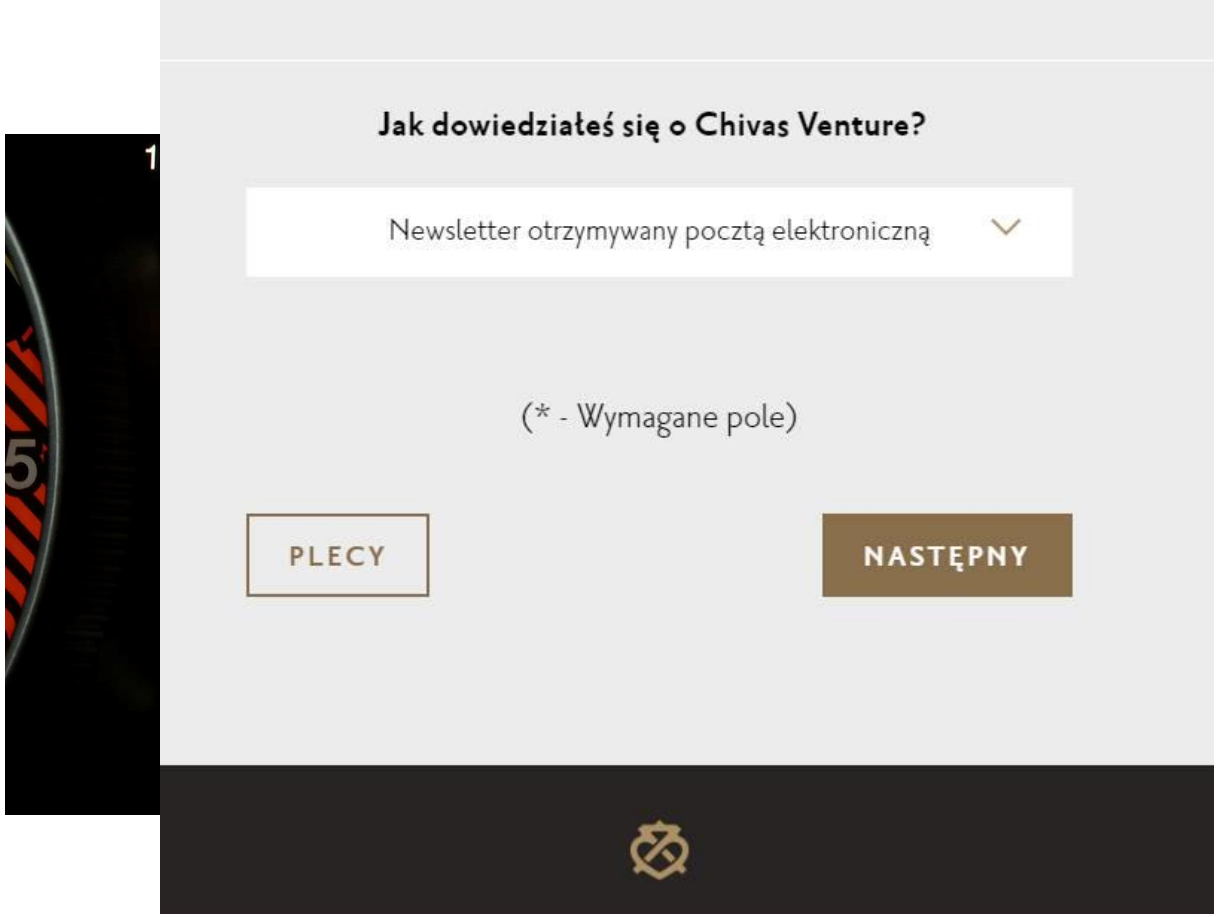
You don’t sign off?

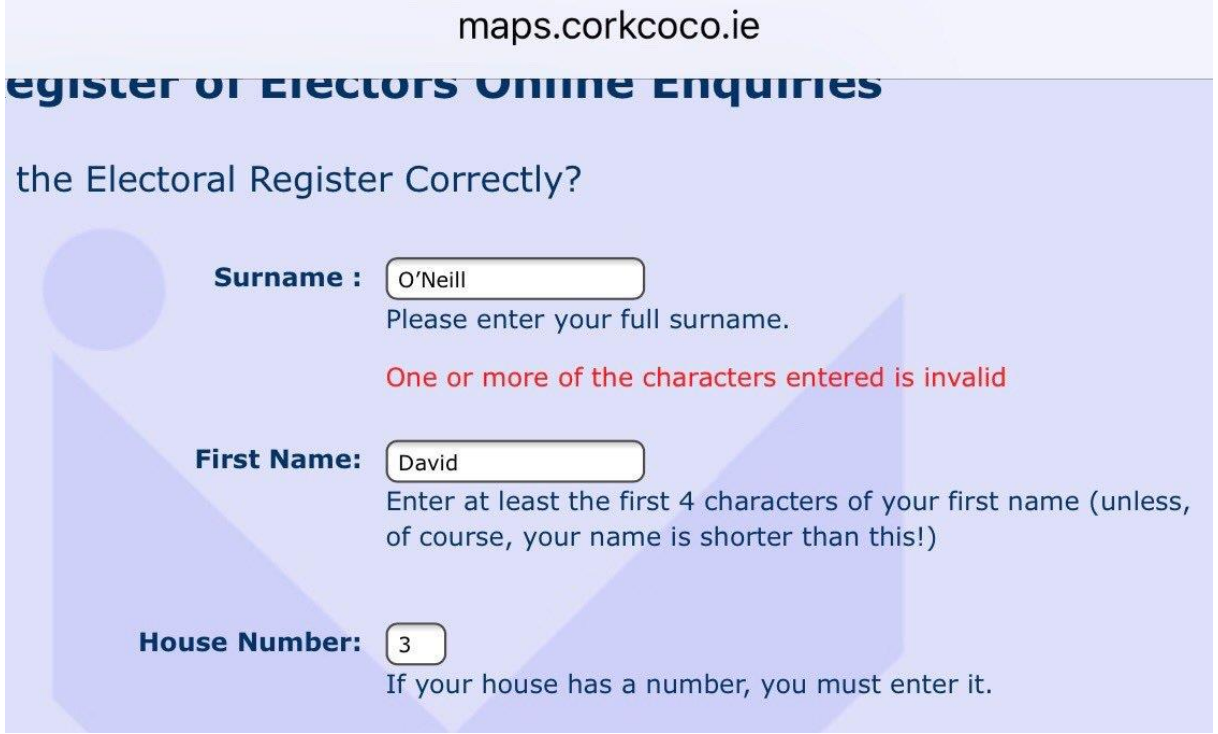
Do **this** to receive less requests to sign on.

Localization testing

- The Translator may or may not test, but should know the test results
- Does the translation fit in the right place?
- Does it make sense there?
- Has the Translator handled all the $variables well?
- …and there’s much more to test: dates, numbers, sorting…
  not only translation

Localization testing





Thank you!

martab@pl.ibm.com

References

- IBM Globalization Design Guidelines

https://www-01.ibm.com/software/globalization/guidelines/

- Carme Mangiron LUDUS EX MACHINA: THE QUEST FOR TECHNOLOGY IN GAME LOCALIZATION

Translation and Localization Conference 2018

- SDL Passolo 2018
- Fellow Translators ☺

How MT overperformed a (bad) translator

Translation: icm-admin-client\src\main\java\com\ibm\ecm\icm\accm\plugin\WebContent\accm\nls\labels.js

box_connection_delete_success_heading: "Połączenie z Box zostało usunięte",

box_connx_client_id_standard_hover_help: "Enter the client ID that is associated with the Case application on Box that uses standard authentication. You can get this value from Box after you create the Box application for your IBM Business Automation Workflow server.",
box_connx_client_secret_standard_hover_help: "Enter the client secret that is associated with the Case application on Box that uses standard authentication. You can get this value from Box after you create the Box application for your IBM Business Automation Workflow server.",
box_connx_client_id_server_hover_help: "Enter the client ID that is associated with the Case application on Box that uses server authentication. You can get this value from Box after you create the Box application for your IBM Business Automation Workflow server.",
box_connx_client_secret_server_hover_help: "Enter the client secret that is associated with the Case application on Box that uses server authentication. You can get this value from Box after you create the Box application for your IBM Business Automation Workflow server.",
box_connx_private_key_hover_help: "Select the private key file that Case Client will use to authenticate the server application with Box. To use Box collaboration with Case Client, you must generate an RSA private-public key pair and associate your Case application on Box with your public key. Case Client will use the private key that is associated with the public key that you have registered with Box. Plik klucza prywatnego powinien mieć format DER.",
box_connx_replace_private_key_hover_help: "Klucz prywatny połączenia z Box już istnieje. Możesz określić zastępczy plik klucza prywatnego. Select the private key file that Case Client will use to authenticate with Box. To use Box collaboration with Case Client, you must generate an RSA private-public key pair and associate your Case application on Box with your public key. Case Client will use the private key that is associated with the public key that you have registered with Box. Plik klucza prywatnego powinien mieć format DER.",
box_connx_public_key_hover_help: "Enter the ID for the public key that is associated with the Case application on Box that uses server authentication. To use Box collaboration with Case Client, you must generate an RSA private-public key pair and associate your Case application on Box with your public key.",
box_connx_automation_app_id_hover_help: "Enter the Enterprise ID that is associated with the Case application on Box.",
box_connx_icn_desktop_hover_help: "Wybierz istniejący pulpit programu IBM Content Navigator, aby powiązać go z połączeniem z Box.",

box_connx_client_id_constraint_invalid_msg: "Wprowadź wartość liczbową.",
box_connx_client_id_constraint_range_msg: "Niepoprawny odstęp publikowania. Wprowadź wartość z przedziału od 5 do 30.",

create_box_connection_running_panel_heading: "Tworzenie połączenia z Box...",
box_connection_running_heading_update: "Modyfikowanie połączenia z Box...",

Translation Memory: WFL18AGP018 (WFL18AGP018_POL_SH1_MT)

0 - You can get this value from Box after you create the Box application for your IBM Business Automation Workflow server.
1[m] - Tę wartość można uzyskać z serwisu Box po utworzeniu aplikacji Box dla serwera IBM Business Automation Workflow.
2[rf] - Wartość tę można uzyskać od Box po utworzeniu aplikacji Box dla swojego serwera IBM Business Automation.