



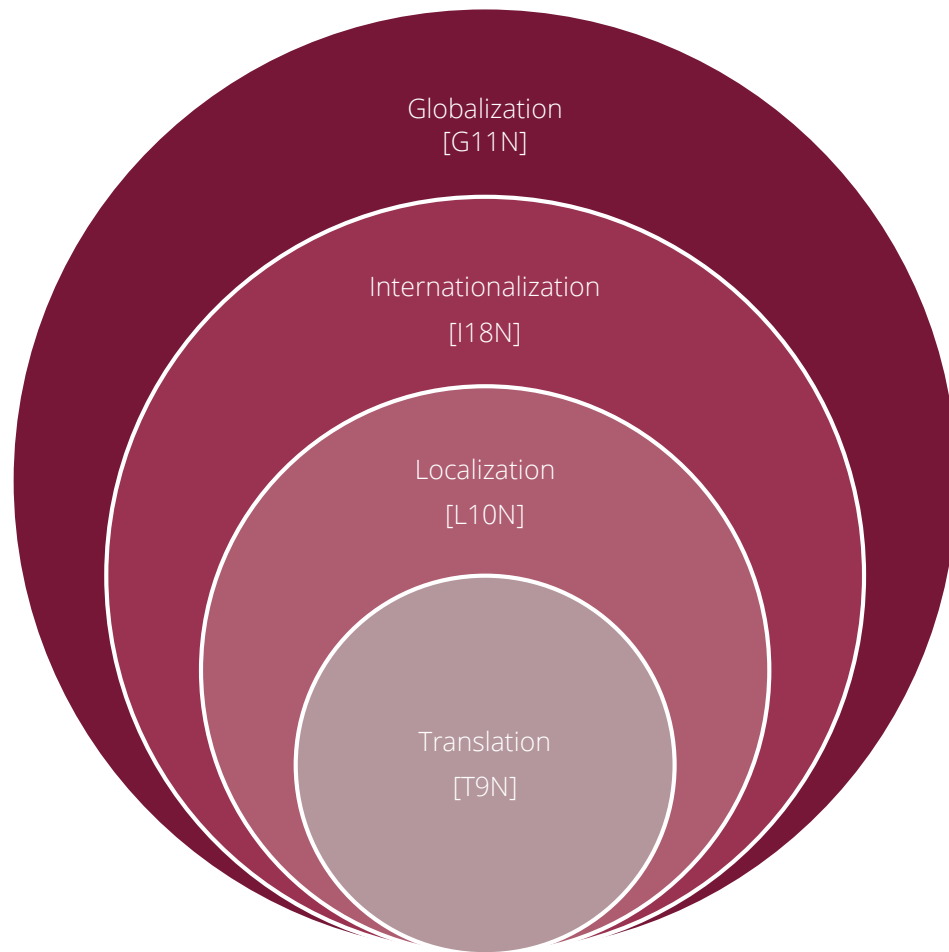
With the support of the
Erasmus+ Programme
of the European Union



Localization & CAT tools

Matúš Jožio | matjozio@gmail.com







What is typical for software / website / game translations?

- › product or service must be first prepared
 - › separation of string from code
 - › adding functionalities or features for specific locale {rtl, longer text, etc.}
- › non-linear text (segmentation)
- › dynamically changing text (placeholders)
- › out of context translation (blind translation)
- › linguistic & functional testing required (to catch possible errors)

Separating strings from developers code

```
1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4 <meta charset="utf-8">
5 <meta http-equiv="X-UA-Compatible" content="IE=edge">
6 <meta name="viewport" content="width=device-width, initial-scale=1">
7 <!-- The above 3 meta tags *must* come first in the head; any other head content must come *after* these tags
8 <title>WCG -- HTML5 Internationalization Example</title>
9
10 <link rel="stylesheet" href="css/html5.internationalization.css" type="text/css" />
11 <!-- HTML5 shim and Respond.js for IE8 support of HTML5 elements and media queries -->
12 <!-- WARNING: Respond.js doesn't work if you view the page via file:// -->
13 <!--[if lt IE 9]>
14 <script src="https://oss.maxcdn.com/html5shiv/3.7.3/html5shiv.min.js"></script>
15 <script src="https://oss.maxcdn.com/respond/1.4.2/respond.min.js"></script>
16 <![endif]-->
17 </head>
18 <body>
19 <h1 data-langkey="WCG -- HTML5 Internationalization Example">WCG -- HTML5 Internationalization Exam
20 <div id="languages">
21 
22 
23 
24 </div>
25 <fieldset>
26 <legend>
27 <span data-langkey="Application Form">Application Form</span>
28 </legend>
29 <span data-langkey="Name">Name</span><br /><input type="text" /><br />
30 <span data-langkey="Company">Company</span><br /><input type="text" /><br />
31 <span data-langkey="Mobile No.">Mobile No.</span><br /><input type="text" /><br />
32 <span data-langkey="Telephone No.">Telephone No.</span><br /><input type="text" /><br />
33 <label data-langkey="Gender">Gender</label><br />
34 <select>
35 <option data-langkey="Male">Male</option>
36 <option data-langkey="Female">Female</option>
37 <option data-langkey="Prefer not to say">Prefer not to say</option>
38 </select>
39 </fieldset>
40
41 <script src="js/html5.internationalization.js" ></script>
42 </body>
43 </html>
```

WCG -- HTML5 Internationalization Example

Application Form

Name

Company

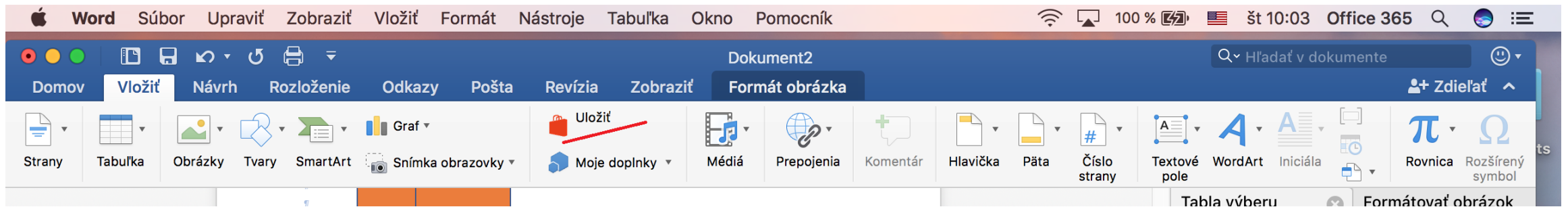
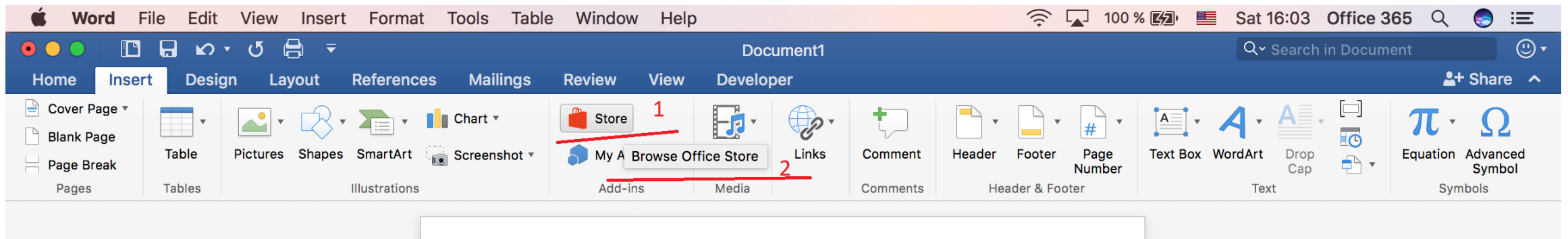
Mobile No.

Telephone No.

Gender

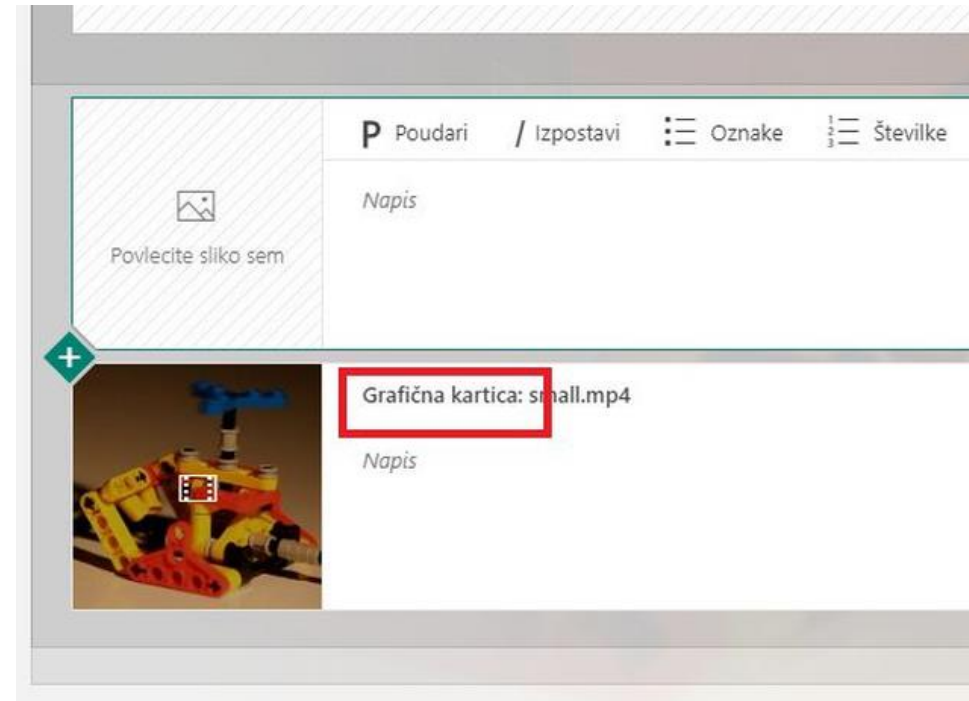
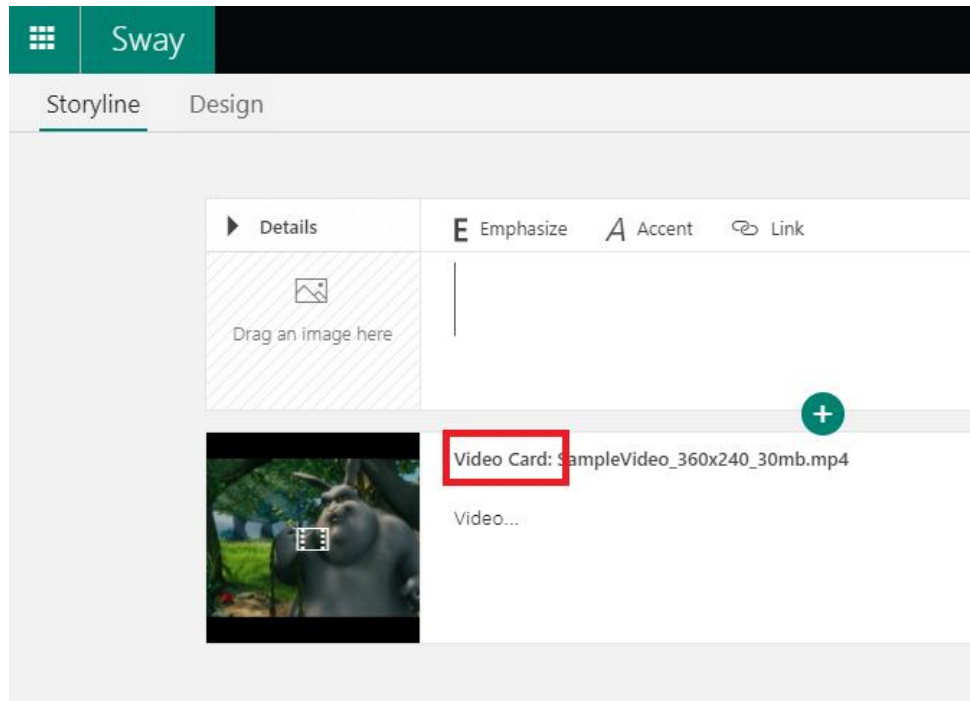
Source ambiguity

- › Translation does not fit the context or is generally awkward



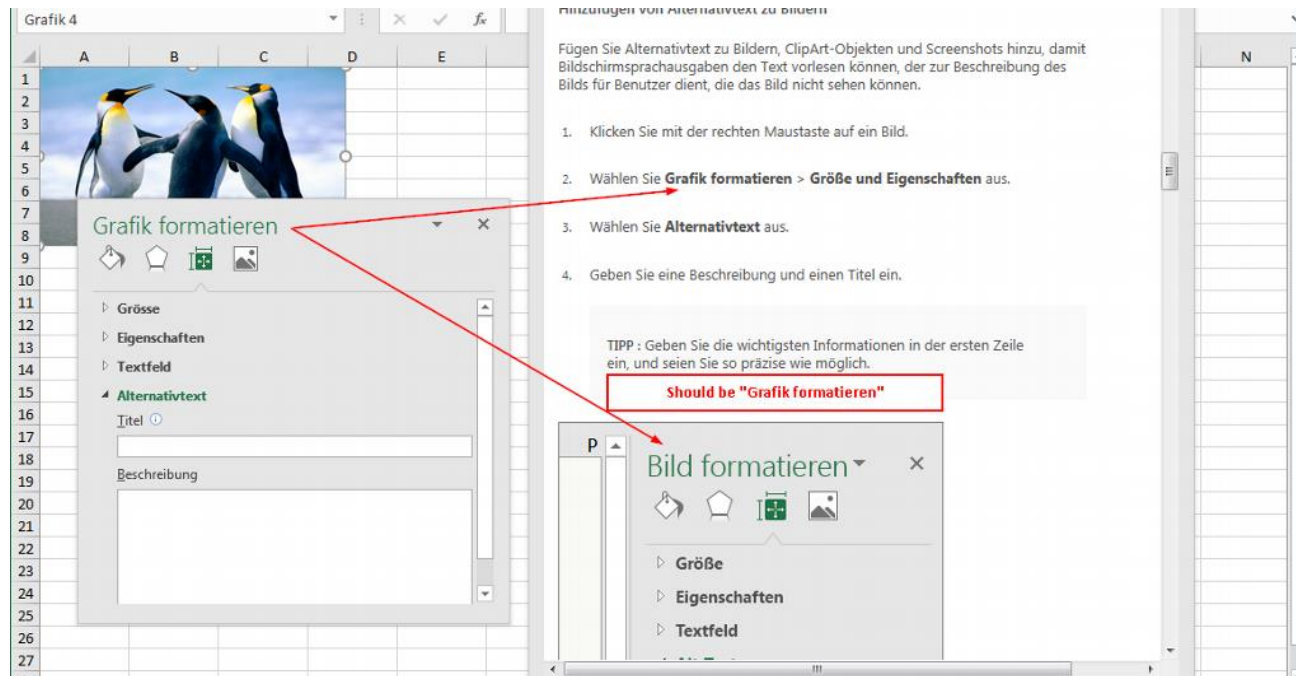
Source ambiguity

- › Translation does not fit the context or is generally awkward



Inconsistencies

- > Same strings translated inconsistently



Truncations

- > Part of the string is not visible (truncated)

Funktionsargumente

ARBEITSTAG.INTL

Ausgangsdatum = Beliebig

Tage = Beliebig

Wochenende = Zahl

Freie_Tage = Beliebig

=

Gibt die fortlaufende Nummer des Datums vor oder nach einer angegebenen Anzahl von Arbeitstagen mit benutzerdefinierten Wochenendparametern zurück.

Freie_Tage ist ein optionales Array aus einer oder mehreren fortlaufenden Datumsnummern, die aus dem Arbeitskalender ausgeschlossen werden sollen, wie etwa Bundes- oder Landesfeiertage und bewegliche

Formelergbnis =

[Hilfe für diese Funktion](#)

OK Abbrechen

should be

ist ein optionales Array aus einer oder mehreren fortlaufenden Datumsnummern, die aus dem Arbeitskalender ausgeschlossen werden sollen, wie etwa Bundes- oder Landesfeiertage und bewegliche Feiertage!

Funktionsargumente

ARBEITSTAG.INTL

Ausgangsdatum = Beliebig

Tage = Beliebig

Wochenende = Zahl

Freie_Tage = Beliebig

=

Gibt die fortlaufende Nummer des Datums vor oder nach einer angegebenen Anzahl von Arbeitstagen mit benutzerdefinierten Wochenendparametern zurück.

Freie_Tage ist ein optionales Array aus einer oder mehreren fortlaufenden Datumsnummern, die aus dem Arbeitskalender ausgeschlossen werden sollen, wie etwa Bundes-/Landes- und bewegliche Feiertage.

Formelergbnis =

[Hilfe für diese Funktion](#)

OK Abbrechen

Placeholder issues

For email from [REDACTED]

String #1 Delete all messages from the Inbox folder

String #2 Delete all messages from the Inbox folder and any future messages

String #3 Always keep the latest message and Delete the rest

String #4 Always Delete messages older than 10 days

[View rules](#) [Sweep](#) [Cancel](#)

For email from [REDACTED]

String #5 Archive all messages from the Inbox folder

String #6 Archive all messages from the Inbox folder and any future messages

String #7 Always keep the latest message and Archive the rest

String #8 Always Archive messages older than 10 days

[View rules](#) [Sweep](#) [Cancel](#)

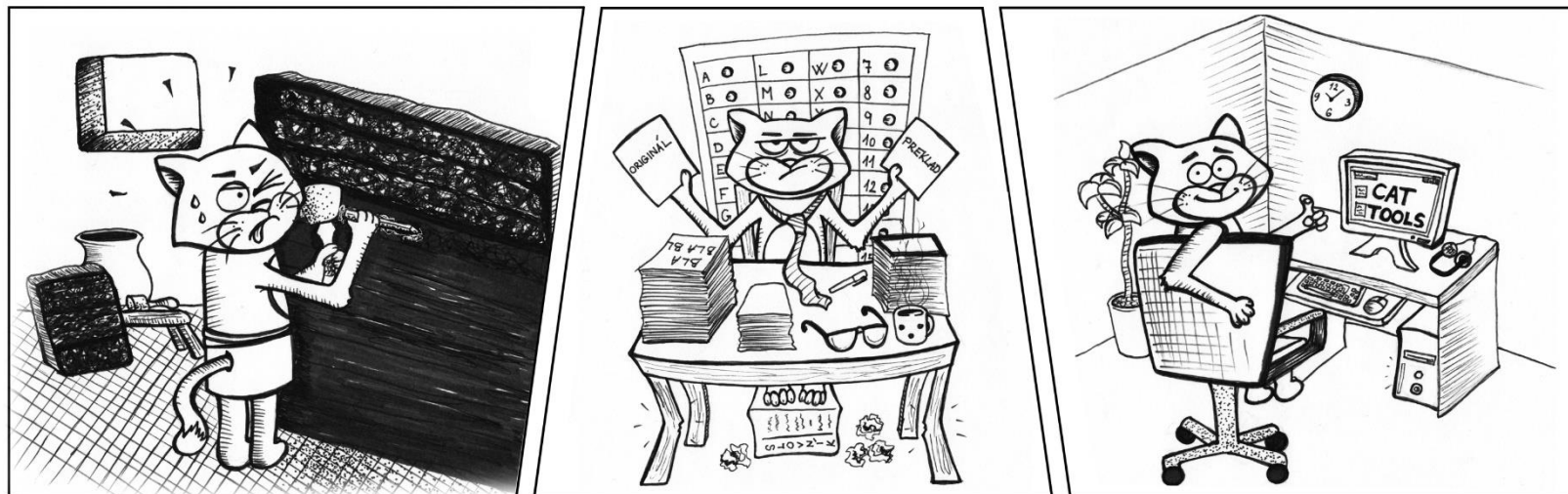
- > String #3 – Estonian – word “Delete” cannot be capitalized in the middle of the sentence
- > Placeholders should be avoided as much as possible. Norwegian has strict capitalization rules and two genders (three in Nynorsk) so placeholders for words like “document” and “presentation” often cause issues.
- > String #4 – Danish – Term “always” have to be placed after the verb (to ensure correct grammar)
*“Always delete messages...” is translated as “*Take* always *away* messages...”*
- > The Icelandic verb for “delete”, “eyða”, takes with it the dative case, while the Icelandic verb for “archive”, “safnvista”, takes with it the accusative case. Therefore the two verbs can't be interchanged without changing the whole sentence accordingly.

II. PART – CAT Tools



What are CAT tools?

- CAT is not MT
- 3 main features>
 - Translation Memories
 - Terminology Databases
 - Source & Target on one screen
- What is available on market?



Demo



DB sources

- › [IATE database](#)
- › [DGT translation memory](#)
- › [Microsoft Terminology Collection](#)



Interested about joining localization industry?

- › [Freelance Opportunities](#)
- › [Current Vacancies](#)
- › [Graduate Opportunities](#)

- › [Join Community](#)