

I18N and L10N

Mark Lin
Dec. 2019

Agenda

- What is i18n?
- What is l10n?
- i18n and l10n
- Basic Guidelines
- What needs to be localized?
- What does NOT need to be put into l10n resource?
- Q&A

What is i18n?

- i18n: internationalization

Make an application **international**; that is, make it able to support virtually any language or local setting on Earth without modify source code.

https://wiki.mageia.org/en/What_is_i18n._what_is_l10n

One single application for ALL countries.

What is i18n?

Examples:

- Software: Mobile phone camera.



What is i18n?

Examples:

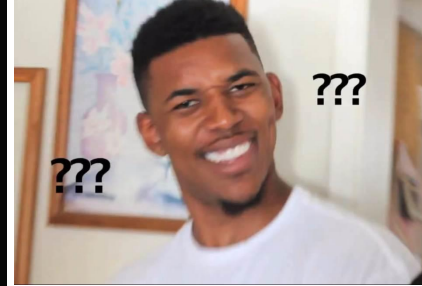
- Car: Accelerator, wheel...



What is i18n?

Examples:

- Car: Accelerator, wheel...



Men in Black II "What is a Gameboy"

<https://www.youtube.com/watch?v=IDtwTqBwqwk&t=1m5s>

What is i18n?

Examples:

- Car: Accelerator, wheel...



Men in Black II "What is a Gameboy"

<https://www.youtube.com/watch?v=IDtwTqBwqwk&t=1m5s>

What is i18n?

Examples:

- Car: Accelerator, wheel...



Men in Black II "What is a Gameboy"

<https://www.youtube.com/watch?v=IDtwTqBwqwk&t=1m5s>

What is I10n?

- I10n: localization

Make an application **localized**; that is, work for a given local and concrete context, adapted to the user.

https://wiki.mageia.org/en/What_is_i18n._what_is_I10n

Different application for different country?

What is I10n?

Examples:

- Software: Disallow to disable japan mobile phone's camera shutter sound.
<https://www.engadget.com/2016/09/30/japans-noisy-iphone-problem/>

What is l10n?

Examples:

- Car: Location of the wheel. Left or Right?



i18n and l10n

Conflict concept?

Not only do not conflict with each other, but also complement each other.

i18n and l10n

i18n makes it available to do l10n.
l10n increase its i18n.

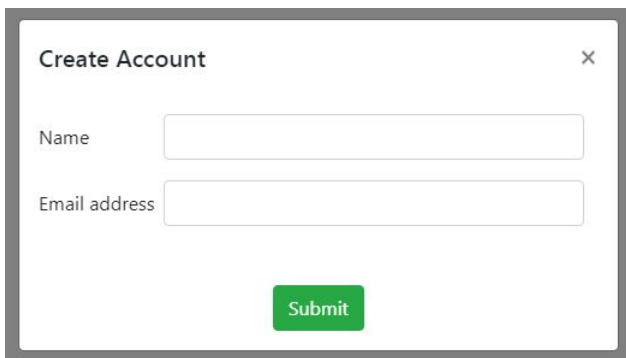


Basic Guidelines

- **Design for multi-languages.** Not only designing for English/native language but also including support to **other languages/cultures**.
 - Layout / Text String Space.
 - Color/Font-family.
 - Functionality.
 - Encoding/Character set: Unicode.
 - Structure of the program.
 - Test plane.
- **Not hardcoded. Use variables**
 - Extract all l10n target out of the source code, and put them into l10n resource folder/files.
- Keep sentence **complete** and **meaningful**.
 - Do not split or cut sentences.

Basic Guidelines - Layout / Text String Space

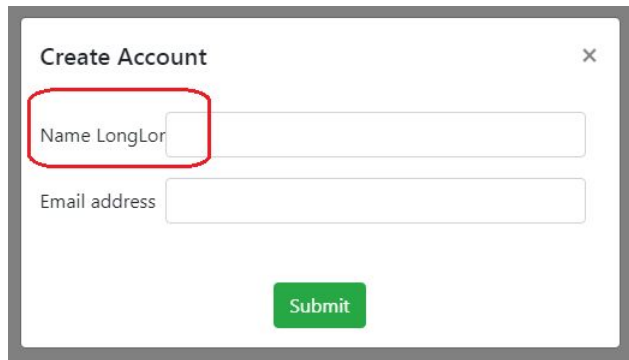
- Design for multi-languages.
 - Layout: Allow **extra space** for translations.
 - Because English is a compact size language. When the text string is translated into other languages, its **string size will increase**



Create Account ×

Name

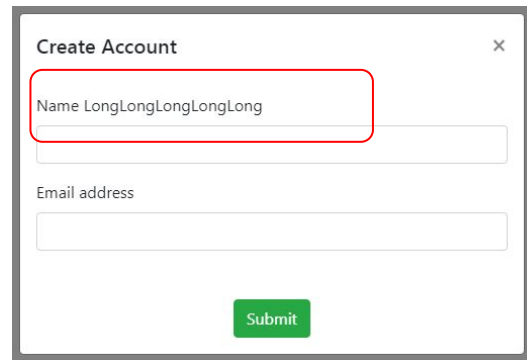
Email address



Create Account ×

Name LongLor

Email address



Create Account ×

Name LongLongLongLongLong

Email address

Basic Guidelines - Layout / Text String Space

- Design for multi-languages.
 - Layout: Allow **extra space** for translations.
 - Because English is a compact size language. When the text string is translated into other languages, its **string size will increase**

Language	Translation	Ratio	No. of characters in English source	Average expansion
Korean	조회	0.8	Up to 10	200-300%
English	views	1	11-20	180-200%
Chinese	次檢視	1.2	21-30	160-180%
Portuguese	visualizações	2.6	31-50	140-160%
French	consultations	2.6	51-70	151-170%
German	-mal angesehen	2.8	Over 70	130%
Italian	visualizzazioni	3		

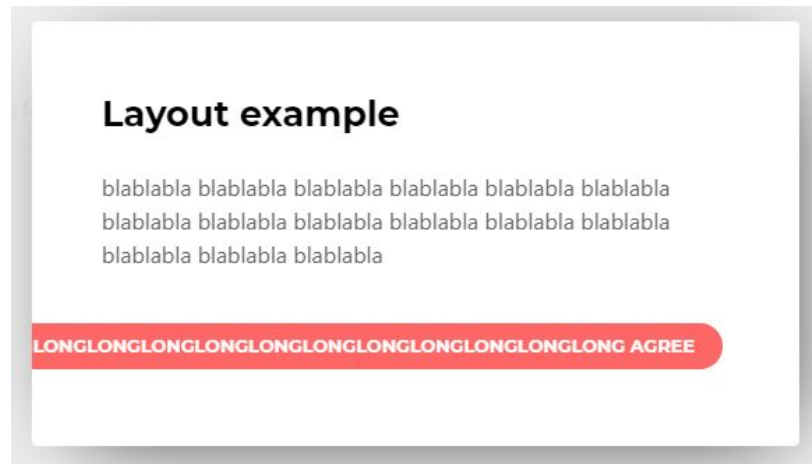
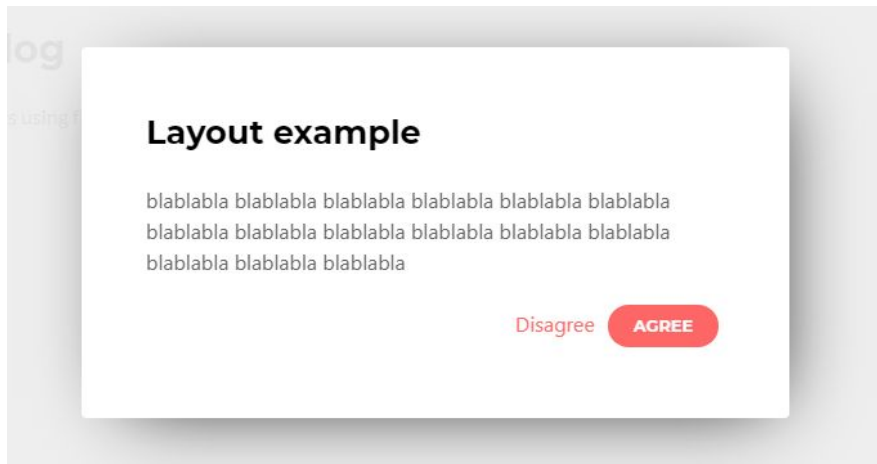
Basic Guidelines - Layout / Text String Space

- Design for multi-languages.
 - Layout: Allow **extra space** for translations.
 - **Compound nouns**
 - **Right-to-left**

Text String in English	Input processing features
in German	Eingabeverarbeitungsfunktionen
in French	Fonctions de traitement d'entrée
in Arabic	وظائف معالجة المدخلات

Basic Guidelines - Layout / Text String Space

- Design for multi-languages.
 - Layout: Allow **extra space** for translations.



Basic Guidelines - Color / Font-Family

- Design for multi-languages.
 - Color/Font-family. Be culture sensitive

A white envelope with money inside.
⇒ What does it mean to you?



The meaning in Taiwan is different from Japan

http://www.shamitsu.com/2018/10/blog-post_4.html

Basic Guidelines - Functionality

- Design for multi-languages.
 - Functionality.



Basic Guidelines - Unicode

- Design for multi-languages.
 - Encoding/Character set: Unicode.

But...


<https://en.wikipedia.org/wiki/N%C3%BCshu>

Nüshu in Unicode [\[edit \]](#)

Main article: Nushu (Unicode block)

Nüshu is included in the [Unicode Standard](#) under the name "Nushu" (because Unicode character names, block names and version numbers are all in English). It was added to Unicode version 10.0 which was released in June 2017. An [iteration mark](#) for Nüshu, U+16FE1 | NUSHU ITERATION

The Unicode block for Nüshu is U+1B170–U+1B2FF:

Nushu ^{[1][2]}																
Official Unicode Consortium code chart  (PDF)																
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+1B17x																
U+1B18x																21

Basic Guidelines - Complete and Meaningful String

- Keep string and sentence **complete** and **meaningful**. Do not split or cut sentence.

Bad: Split the sentence.

```
"Hi " + "Wade" + ", are you ready to race?"
```

L10N resource

```
{
  "001": "Hi ",
  "002": "Internationalization (i18n) and localization (L10n)",
  "003": "Internationalization is the process of designing a software application",
  "004": "Not only picking i18n framework & apply it to front-end codebase",
  "005": "technical requirements",
  "006": ", are you ready to race?",
}
```

Basic Guidelines - Complete and Meaningful String

- Make sentence **complete** and **meaningful**. Do not split or cut sentence. Use **variables**.

Better: Keep the sentence complete, and apply variables.

```
"Hi ${playerName}, are you ready to race?"
```

L10N resource

```
{  
  "001": "Hi ${playerName}, are you ready to race?",  
  "002": "Internationalization (i18n) and localization (L10n)",  
  "003": "Internationalization is the process of designing a software application",  
  "004": "Not only picking i18n framework & apply it to front-end codebase",  
  "005": "technical requirements",  
}
```

Basic Guidelines - Complete and Meaningful String

- Keep sentence **complete** and **meaningful**.

Question	Answer
What kinds of games does Alice play?	<input type="text"/>
What kinds of drinks does Alice want?	<input type="text"/>
What kinds of movies does Alice like?	<input type="text"/>
What kinds of dogs does Alice have?	<input type="text"/>

Basic Guidelines - Complete and Meaningful String

- Make it as template?

Option A	Option B
<pre>UserName = Alice; Data = [{ 0: game, 1: play }, { 0: drinks, 1: want }, { 0: movies, 1: like }, { 0: dogs, 1: have },]; Loop in Data: output: What kinds of {0} does \${UserName} {1}?</pre>	<pre>UserName = Alice; Data = [{ 0: What kinds of games does \${UserName} play? }, { 0: What kinds of drinks does \${UserName} want? }, { 0: What kinds of movies does \${UserName} like? }, { 0: What kinds of dogs does \${UserName} have? },]; Loop in Data: output {0}</pre>

Basic Guidelines - Complete and Meaningful String

- Make it as template?

English	Japanese
What kinds of games does Alice play?	アリス はどのような ゲーム をプレイしますか？
What kinds of drinks does Alice want?	アリス は何を 飲みたい ですか？
What kinds of movies does Alice like?	アリス はどんな 映画 が好きですか？
What kinds of dogs does Alice have?	アリス にはどんな 犬 がいますか？

Basic Guidelines - Grammar Matters

- I10n resource is **NOT** template between different sentences.
- **Grammar matters**
 - Leave grammar in I10n resource, and let translator handle it.



Basic Guidelines - Grammar Matters

- I10n resource is **NOT** template between different sentences.
- **Grammar matters**
 - Leave grammar in I10n resource, and let translator handle it.

English: Invalid license

Adjective + Object

French: Licence non valide

Object + Adverb + Adjective

Basic Guidelines

- **Design for multi-languages.** Not only design for English/native language but also include **other languages/cultures**.
 - Layout.
 - Color/Font-family
 - Functionality.
 - Encoding/Character set: Unicode.
 - Structure of the program.
 - Test plane.
- **No hardcoded. Use variables**
 - Extract all l10n target out of the source code, and put them into l10n resource folder/files.
- Keep sentence **complete** and **meaningful**.
 - Do not split sentence.

What needs to be localized?

- All displayable words.

What needs to be localized?

- Punctuation.
 - Comma, Period, Ellipsis, Parentheses...

English: I have a pen, and an apple.

Chinese: 我有一枝筆, 和一顆蘋果。



- Symbols and Signs.
 - Percent sign, ...

English: 40% chance of rain for Ohio State-Nebraska game

Turkish: Ohio State-Nebraska oyunu için% 40 yağmur ihtimali

What needs to be localized?

- Date/Time.

09/06/2019

What date?

USA (mm/dd/yyyy): September 6th

France (dd/mm/yyyy): June 9th

Format matters

What needs to be localized?

- Date/Time.

ISO-8601: 2018-04-05

Taiwan: 民國 107 年 4 月 5 日

Japan: 平成 30 年 4 月 5 日

USA: April 5, 2018

What needs to be localized?

- Date/Time.

Timestamp: 1569574800

Taiwan (UTC+8): 17:00

Japan (UTC+9): 18:00

Time zone matters

What needs to be localized?

- Numbers with symbols.

12,345

How much?

USA : 12 thousands

France: 12

Format matters

https://en.wikipedia.org/wiki/Decimal_separator#Examples_of_use

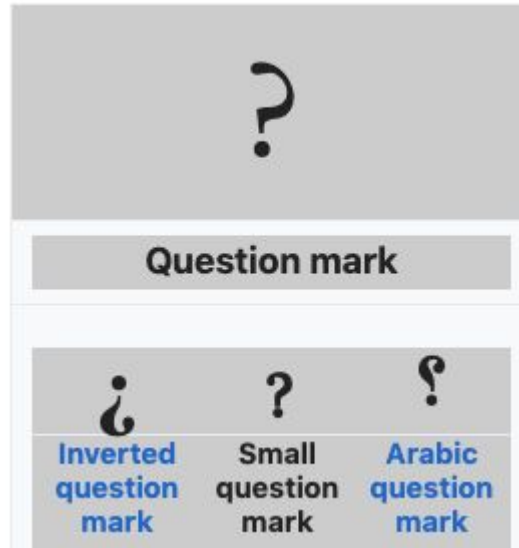
What needs to be localized?

- Images with wording in it.



What needs to be localized?

- Icon (Sometimes).
 - Question mark



What needs to be localized?

- All displayable sentences.
- Punctuation, Symbols and Signs.
- Date/Time.
 - Format (yyyy/mm/dd)
 - Timezone (UTC+8)
- Numbers with symbols.
 - Decimal point/comma
 - Thousands Separators
- Images with wording in it.
- Icons (Sometimes).
 - Question mark

What does NOT need to be put into l10n resource?

- Source code.
 - Original: To learn more, please check `Online help`.
 - Suggest: To learn more, please check `${link__OnlineHelp}`.

What does NOT need to be put into l10n resource?

- Spical token used by our program.
 - Original: Use * to represent any series of characters.

We don't want it become: 使用 * 符號代表任意字元。

(Note: * is not *)

- Suggest:
Use \${token__Asterisk} to represent any series of characters.

What does NOT need to be put into l10n resource?

- Dynamic content.
 - Original: Disk usage: 80%

We cannot list them all.

- Suggest: Disk usage: \${number}%

What does NOT need to be put into l10n resource?

- Source code.
 - To learn more, please check [here='https://example.com/'](https://example.com/)Online help.
- Spical token.
 - Use * to represent any series of characters.
- Dynamic content.
 - Disk usage: 80%

Basic Guidelines

- Design with i18n and l10n.
- All **displayable** Text, Date/Time, Numbers, Punctuation, Symbols, Signs, Images and Icons (Some cases).
- Keep sentence **complete** and **meaningful**.
- L10N resource is **NOT** template between different sentences.
- Do not include **no-need-to-translate** element into the l10n resource.
 - Source code.
 - Spical token.

Q&A

Thank You

Appendix

Reference

- [1] https://wiki.mageia.org/en/What_is_i18n,_what_is_l10n
- [2] <https://blog.mozilla.org/l10n/2011/12/14/i18n-vs-l10n-whats-the-diff/>
- [3] <https://www.w3.org/International/questions/qa-i18n>
- [4] https://developer.mozilla.org/en-US/docs/Mozilla/Localization/Localization_content_best_practices
- [5] <https://www.infragistics.com/community/blogs/b/devtoolsguy/posts/l10n-and-i18n-best-practices>
- [6] https://en.wikipedia.org/wiki/Percent_sign
- [7] https://en.wikipedia.org/wiki/Date_format_by_country
- [8] https://en.wikipedia.org/wiki/Decimal_separator#Examples_of_use
- [9] <https://www.engadget.com/2016/09/30/japans-noisy-iphone-problem/>
- [10] <https://www.infragistics.com/community/blogs/b/devtoolsguy/posts/introduction-to-localization-and-internationalization>
- [11] <https://www.infragistics.com/community/blogs/b/devtoolsguy/posts/l10n-and-i18n-best-practices>