

Lokalizace rozsáhlých internetových projektů

18. dubna 2007, ZČU Plzeň

Ladislav Thon
Pavel Hofman

inSITE, s.r.o.
www.insite.cz
development@insite.cz

Obsah

- **základní pojmy**
- **lokalizace na platformě Java**
- **lokalizace webových aplikací**
 - **běžné postupy**
 - **jejich problémy**
- **naše řešení**

Obsah prezentace

Část I **Představení společnosti inSITE, s.r.o.**

- Klíčové projekty
- Hlavní klienti

Část II **Varianty řešení lokalizace**

- lokalizace na platformě Java
- lokalizace webových aplikací
 - běžné postupy
 - jejich problémy
- naše řešení

Část III **FAQ - odpovídáme na Vaše dotazy**

Dovolte, abychom se Vám představili

Hlavní data o inSITE, s.r.o.

- založeno v roce 1998
- pobočky Praha, Plzeň, 19 zaměstnanců

Hlavní projekty

- www.EduCity.cz
- www.JobCity.cz
- www.Meetings.cz
- InfoServis.insite.cz
- www.HRzive.cz



Hlavní klienti inSITE

- Škoda AUTO
- ČEZ, a.s.
- Česká spořitelna
- ČSOB
- Telefónica O2
- T-Mobile
- Zentiva
- Generali

Většinou vývoj interních produktů

Výhody:

- Menší stres a „šturmování“ před termínem dokončení
- Větší volnost při volbě vhodného řešení
- Kompatibilita s předchozími verzemi postačuje na úrovni uživatelského rozhraní pro smluvní klienty

Nevýhody:

- Vyšší pravděpodobnost posunutí interního termínu

Základní pojmy

- **internacionalizace**
 - vývoj či úprava softwaru tak, aby jej bylo možné používat v různých národních prostředích
 - „aby bylo možné jej lokalizovat“
- **lokalizace**
 - úprava produktu pro použití v jednom konkrétním národním prostředí

Co je předmětem lokalizace

- text (kódování, abeceda, směr psaní, font)
- čísla
- datum a čas (kalendář, letní čas, časové zóny, 12/24hod.)
- měna
- míry (anglosaský měrný systém, SI)
- názvy, jména, tituly, obrázky, symboly, barvy, rozložení klávesnice (klávesové zkratky), typografická pravidla, ...
- pravidla pro řazení

Lokalizace na platformě Java

- **Unicode**
- **locale** (`java.util.Locale`)
 - popis relevantních vlastností národního prostředí
 - **language/country/variant** (`en` vs. `en_US` vs. `en_US_MAC`)
- `java.text.*`, `java.util.Calendar`, `java.util.Currency`, ...
- **resource bundle** (`java.util.ResourceBundle`)
 - základní způsob, jak dostat jazykově závislé řetězce vně zdrojového textu aplikace
 - klíč -> řetězec

Ukázka

--- Sample.java

```
Locale locale = new Locale("cs");  
  
ResourceBundle bundle =  
    ResourceBundle.getBundle("messages", locale);  
  
String title = bundle.getString("sample.title");
```

--- messages.properties (default)

```
sample.title=Hello World!
```

--- messages_cs.properties

```
sample.title=Good night, krut\u00fd sv\u011bte...
```

Lokalizace na platformě Java

- **standardní resource bundle**
 - musí být v Latin1 (Unicode sekvence)
 - nelze upravovat za běhu
 - neumí řetězce s parametry (lze použít `MessageFormat`)
 - nízký výkon
- řešení:
 - velmi často vlastní implementace
 - příp. z různých frameworků
 - **Spring: rozhraní `MessageSource` a jeho implementace**
- **nutnost: podpora ve vývojovém prostředí**
 - **IntelliJ IDEA (editace, navigace, refaktoring)**

Lokalizace webových aplikací

- šablonování (JSP, Velocity, FreeMarker, ...)
 - oddělení vývoje HTML šablon a programového kódu
- lokalizace textů
 - speciální značky šablonovacího jazyka
 - přímé volání metod speciálního objektu
- lokalizace dalších dat
 - speciální podpora zvlášť pro čísla, data, ...
- lokalizace celých částí šablon
 - hierarchie šablon

HTML design

```
<html>
<head>
  <title>Hello World!</title>
</head>
<body>
  Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do
  eiusmod tempor incididunt ut labore et dolore magna aliqua.
</body>
</html>
```

JSP

- **téměř programový kód**

```
<%@ page contentType="text/html; charset=ISO-8859-2" language="java" %>
<%@ taglib prefix="fmt" uri="http://java.sun.com/jsp/jstl/fmt" %>
<html>
<head>
    <title><fmt:message key="sample.title" /></title>
</head>
<body>
    <fmt:message key="sample.content" />
</body>
</html>
```

Velocity

```
<html>
<head>
  <title>${bundle.getString("sample.title")} /></title>
</head>
<body>
  ${bundle.getString("sample.content")} />
</body>
</html>
```

Problémy standardních přístupů

- HTML od designéra není přímo použitelné pro nasazení a šablona od programátora je pro designéra rozsypaný čaj
- zdrojem textů je resource bundle
 - ruční práce s klíči jednotlivých položek
- obtížná práce s rozsáhlými bundly
 - duplicity
 - nepoužívané řetězce
 - synchronizace jazykových verzí
- netextová data speciálním způsobem
 - speciální značky pro data, čísla, peníze, ...
 - obrázky, ...

Problémy standardních přístupů

- **JSF+Facelets**

```
<span jsfc="h:outputText" value="#{message}">Hello World!</span>
```

- **Tapestry**

```
<span jwcid="@Insert" value="ognl:message">Hello World!</span>
```

- **Wicket**

```
<span wicket:id="message">Hello World!</span>
```

„Ideální šablona“

```
<html>
<head>
  <title>Hello World!</title>
</head>
<body>
  Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do
  eiusmod tempor incididunt ut labore et dolore magna aliqua.
</body>
</html>
```

Naše šablona

- základní tvar

```
###ISO-8859-2
```

```
<html>
```

```
<head>
```

```
  <title>§-Hello World!-§</title>
```

```
</head>
```

```
<body>
```

```
  §-Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do  
  eiusmod tempor incididunt ut labore et dolore magna aliqua.-§
```

```
</body>
```

```
</html>
```

Kompilace

- **metašablona (v daném kódování)**
 - přímo se nepoužívá
 - zdroj defaultních hodnot řetězců
- **řetězce se ukládají do speciálního resource bundlu**
- **kompilace**
 - neznámým řetězcům se přiřadí unikátní klíč
 - uloží se do bundlu
 - vloží se do šablony

Naše šablona

- úplný tvar

```
###ISO-8859-2
```

```
<html>
```

```
<head>
```

```
    <title>§-Hello World! /aaa-§</title>
```

```
</head>
```

```
<body>
```

```
    §-Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do  
    eiusmod tempor incididunt ut labore et dolore magna aliqua. /aab-§
```

```
</body>
```

```
</html>
```

Speciální resource bundle

```
<messages-bundle>  
  <msg key="aaa">  
    <text>Hello World!</text>  
  </msg>  
  <msg key="aab">  
    <text>Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed  
    do eiusmod tempor incididunt ut labore et dolore magna aliqua.</text>  
  </msg>  
</messages-bundle>
```

Speciální resource bundle

- obsahuje defaultní hodnoty i všechny jazykové verze
- je udržován zcela automaticky
 - vkládání nových položek
 - nový automaticky vygenerovaný klíč
 - odstraňování nepoužívaných položek
 - zpracování duplicit

Kompilace

- generování skutečně používaných šablon
 - pro každou jazykovou verzi + default
- údržba resource bundlu
- pouze nahrazování řetězců
 - nezávislost na šablonovacím enginu
 - stejným způsobem lze lokalizovat cokoli

```
templates/  
index.vm
```

```
templates_utf/  
index.vm  
index_cs.vm
```


Překlad

- export řetězců pro překlad z bundlu (obvykle do Excelu)
- import přeložených řetězců
- rekompilace

```
<messages-bundle>
```

```
  <msg key="aaa">
```

```
    <text>Hello World!</text>
```

```
    <text lang="cs">Good night, krutý světe...</text>
```

```
  </msg>
```

```
  ...
```

```
</messages-bundle>
```

Řešení duplicit

- základní tvar

```
$-Hello World!-$
```

```
$-Hello World!-$
```

- přiřazení klíčů

```
$-Hello World!/aaa-$
```

```
$-Hello World!/aaa1-$
```

- překlad se může lišit v závislosti na kontextu

- v řadě situací však nebude

- popis tlačítka „OK“, ...
- ruční slučování klíčů (**aaa1** -> **aaa**)
- zadání klíče už při editaci

Shrnutí

- další vrstva nad šablonovacím nástrojem
 - výborně řeší lokalizaci řetězců
 - s lokalizací ostatních dat může pomoci
 - ``
- náměty k dalšímu rozvoji
 - překladatel by měl znát kontext překládaného textu
 - přinejmenším komentář
 - vkládání značek do již hotových složitých šablon je náchylné na chyby
 - podpora v editorech, vývojových prostředích
 - možnost použití ve zdrojových textech aplikace

FAQ - závěr prezentace

Děkujeme za pozornost!