

Entries/Entrées

Visas

Departures/Sorties

7

INTERNATIONALIZATION AND LOCALIZATION

The two terms most critical to the success of web globalization—internationalization and localization—are also the two most frequently misunderstood. Their odd-looking abbreviations (i18n and L10n) certainly don't help matters. Internationalization is the process of building a web site so that it can support multiple locales, while localization is the process of modifying that site for a specific locale.

At first glance, the two terms don't appear related at all.

Internationalization implies taking a global approach to web development, but localization implies just the opposite. Yet these two terms are intimately linked, so much so that it can be difficult to tell where internationalization ends and localization begins. This chapter will help you not only differentiate between the two, but also understand how to successfully use them together.

THE I18N AND L10N OF AN AUTOMOBILE

The principles behind internationalization and localization extend beyond web development, even to something much more concrete, like a car. A car, like a web site, is expensive to design and build. To minimize costs and maximize returns, a car manufacturer often develops a car model that can be easily adapted to numerous countries, instead of developing new models for each country. Doing so requires internationalization. The internationalization stage is the “behind the scenes” stage. People don’t buy internationalized cars; they buy localized cars. Internationalization mostly entails the extensive planning and testing that go into creating this global template of a car. For example, if the car will be sold in both the U.S. and the U.K., allowances must be made for placing a steering wheel on either side of the car. Some car manufacturers might decide during this process that some markets are just not worth the cost of localization efforts. It’s a cost/benefit decision that you’ll also have to make as you internationalize your web site.



After internationalization is finished, the car can be localized for each market. The more thorough the job you do of internationalization, the less time you’ll spend on localization. Localization can be as simple as moving the steering wheel to the other side or could be as complex as deciding what color palette to offer. The line between localization and personalization is not always so clearly defined. Inevitably, you want your web site to be as customized as possible to your audience, but you can’t do everything. Even car manufacturers don’t offer every option imaginable, which is why car buyers do a fair amount of customization themselves. Just as in the internationalization stage, the decisions you make during localization are heavily dictated by costs and benefits.

The key to success is striking a balance between flexibility and profitability. Similar challenges face a web team: How global do you *need* to go? How local can you *afford* to go?

Microsoft Localizes the Xbox

Before launching its new Xbox gaming console in Japan, Microsoft localized the game controllers—moving the buttons closer together—to better fit the average Japanese user’s hands.

Source: New York Times, February 18, 2002

THINKING GLOBALLY

Thinking globally requires big thinking. Put aside budgetary constraints for a moment and think of every possible country or region your company might want to target:

India?

Malaysia?

The Middle East?

Although these three markets might not be at the top of a company's globalization strategy, it's just a matter of time before they get there. There are more than 1.5 billion people in these three markets alone, and their buying power is exploding. Even if you plan to enter only a few European markets this year, you might someday end up like IBM, with 62 localized web sites. If you want to build a global company, you need to think global from the beginning.

Thinking Big

Coke has 26 localized web sites.
 Lycos has 34 localized web sites.
 IBM has 62 localized web sites.

173

What's Your Localization Timeline?

Creating a localization timeline for the next 5 to 10 years can be as simple as this example:

Sample Localization Timeline

Year 1: Spanish-Mexico, Portuguese-Brazil

Year 3: French-France, German-Germany

Year 5: Japanese-Japan, Korean-Korea

Year 8: Arabic-Middle East, Tamil-India

A localization timeline, including both language and region, helps you avoid a lot of the common technical, legal, and cultural potholes that await. For example, assume that you test your brand for the first four markets you're planning to enter. The initial launches work so well that you accelerate your rollout schedule and add a few new markets. Unfortunately, because you didn't test all potential markets initially, competitors have beaten you to them by launching similar brands and localized web sites of their own. With the Internet, anyone anywhere can spy on you. If you have global aspirations with your products or services, you need to make the investment upfront to ensure that you can launch in those markets without hitting regulatory and copyright obstacles. If you want to avoid the sort of conflict that Anheuser-Busch is involved in with Budvar (see sidebar), think big from the beginning.



When You Say Budweiser...

There is more than one company in the world brewing Budweiser (see Figure 7.1): the American giant Anheuser-Busch and the tiny Czechoslovakian brewer Budvar.



Figure 7.1 Budweiser Czechoslovakia and Budweiser U.S.: Similar names, different brands.

Needless to say, the big guy has won trademark battles in much of the world, but not all of it. Budvar has the right to Budweiser in Germany and Russia, so Anheuser-Busch promotes “Bud” instead. It’s a nasty fight and very much ongoing. Since 1939, Budvar has been prohibited from selling its beer under the Bud, Budweis, or Budweiser name in North America. But that hasn’t stopped it from going after A-B where it lives; it recently released a new product for the U.S. market: Czechvar.

Separating the Constants from the Variables

The main purpose of internationalization is to isolate the graphical and textual elements of a web site that change from locale to locale as well as within the locale itself. The way these elements are managed for each market—in terms of design and text—will be covered in later chapters. For now, you just want to make sure you’ve got a good understanding of what elements will change and what elements will not.



The Constants

A constant is anything that remains the same, no matter what market you localize for. It can be the design template that your site shares across all web pages, or it could be a collection of scripts and style sheets. It can also be a collection of what's called "corporate constants," such as brand names, slogans, logos, colors, and navigation menus. For example, in the Spanish banner ad for Volvo, shown in Figure 7.2, notice that the slogan "for life" remains in English. Very often, companies decide to maintain a global slogan regardless of the target market.



Figure 7.2 In the Volvo banner ad, the logo and slogan are constants.

Some companies also maintain unique colors across all locales, such as the Coca-Cola red or IBM blue or UPS brown. Despite how various cultures perceive colors, many corporations elect to err on the side of consistency.

The Variables

Variables include anything that changes from market to market or within a market. During the internationalization stage, you focus on isolating the variables and modifying your site so that they can be more easily adapted to each market. Variables include:

- Measurements and sizes
- Prices and currencies
- Dates, calendars, and time zones
- Product selection
- Contact information
- Images and icons
- Forms and input fields

The more variables your site consists of, the more challenging internationalization becomes. However, there is no rule that says you have to offer the same variables across all locales. To simplify internationalization and localization, you might decide to limit the number of variables available in each localized site. In fact, it's rare to find a company that provides the same level of functionality and support on its localized sites that it does on its source-language site.

Text Expansion and Contraction

When a block of text is translated into another language, it tends to expand or contract, depending on the target language. This phenomenon, known as *text expansion* or *text contraction*, occurs because translation is not a one-to-one process. The word *cat* in English translates to “gato” in Spanish, thus adding a character. In Chinese, a cat can be represented by just one character: 猫

Although much depends on the verbosity of the translator, general patterns are noticeable. English text often expands when translated into European languages— from approximately 15% in Spanish to as much as 35% in German (see Figure 7.3). Asian languages typically require fewer characters than their English equivalents, yet you might not see much contraction; even though you're working with fewer characters, they generally need to be displayed at a larger point size to ensure legibility. Text expansion and contraction become particularly acute when working with small amounts of text, such as with headlines or text in navigation bars. For example, *search* translated into French becomes “recherché,” a 30% expansion.

Think about how text expansion will affect your site's design and functionality. Also alert your translators to places where space is tight on your site; they can often provide more austere translations, but they can't work miracles. If you don't give your web design enough room for text expansion, you'll face a painful redesign when localizing for European markets. The best way to avoid this problem is to follow Yahoo!'s lead and simply avoid embedding text into graphics; if you must embed text, allot plenty of room for the inevitable expansion.

English:	News
Italian:	Notizie
French:	Actualités
German:	Schlagzeilen
Japanese:	ニュース
Chinese:	新闻
Korean:	뉴스

Figure 7.3 The “news,” when translated, either expands outside its allotted space or leaves room to spare, depending on the language and the style of translation.

177

Don't Forget the Back End

Some of the more complex text expansion problems are not so visible. It's not uncommon to discover, after your web site has launched, that users in Germany are having difficulty entering their complete addresses (because the fields won't allow for additional characters) or users in Japan are having problems inputting their names (because the database won't accept Japanese characters). Always make sure that you adapt the text fields to allow for longer names and other locale-specific issues.

Address fields are particularly challenging. There are more than 100 different address formats in use around the world (see sidebar “Return to Sender”). If you've developed a web order form for users in the U.S., you'll have to make some immediate changes to accommodate users outside the U.S. Countries such as the U.S., Canada, and Brazil have states, but most countries do not, so you shouldn't require that the state field be filled out. Better yet, offer a localized order form that does not include the state field. The ZIP code field is also challenging. For starters, the term *ZIP code* is unique to the U.S.; other countries call it a “postal code” or “postcode.” In the U.S., a ZIP code is either 5 or 9 digits, but in other countries, a postal code can be anywhere from 3 to 7 digits and might include letters. And just to keep you

on your toes, some countries, such as Ireland, don't even use a postal code. The technical issues of managing text in databases and web applications are complex, and beyond the scope of this book. Just be aware that you'll probably need to modify your databases to support new and longer input fields and additional character sets.

Return to Sender

With 112 different address formats in the world, odds are that you'll need to update your web site to accommodate entering and displaying addresses for different countries. And details count. For example, compare the following two addresses: the first for the U.S., the second for Germany.

**JOHN DOE
55 COOLIDGE ST
BOSTON, MA 02151-4645**

For the U.S., the standard is all uppercase letters, with the street number coming before the street name and the ZIP code on the same line as the city and state.

**Herm
Gunther Meyer
Goethestrae 25**

20002 HAMBURG

For the German address, the personal form of address, *Herm*, is written on a separate line, the house number follows the street number, and an empty line is inserted above the postal code.

Source: "A Guide to International Address Management," GRC Database Information (www.grcdi.nl)

Global Architecture

If you localize your English site for six locales, you might find yourself with six times as many web pages to manage. If you don't implement and maintain a sound structure, you may also find yourself struggling to manage all those new files. Many developers, because they begin localization with just one language, find that they've labeled the pages haphazardly, as shown in Figure 7.4.

Notice how the Spanish pages and directories are labeled inconsistently and in different languages, and are mixed together with the source-language pages. Consider how confusing this arrangement will be after the site grows to a few thousand pages. Now for an alternative approach, shown in Figure 7.5.

	A	B	C	D	E
1	Sample Web site architecture				
2					
3	/				
4		index			
5		index_espanol			
6		welcome.html			
7		bienvenidos.html			
8					
9	/about				
10			index.html		
11					
12	/about_spanish				
13			index.html		
14					
15	/products				
16			item1.html		
17			item2.html		
18					
19	/spanish_products				
20			item1.html		
21			item2.html		
22					

Figure 7.4 An architecture for trouble.

179

	A	B	C	D
1	Sample Web site architecture			
2				
3	/en/			
4		index.html		
5		welcome.html		
6				
7	/about/			
8			index.html	
9				
10	/products/			
11			item1.html	
12			item2.html	
13				
14	/es/			
15		index.html		
16		welcome.html		
17				
18	/about/			
19			index.html	
20				
21	/products/			
22			item1.html	
23			item2.html	
24				
25				

Figure 7.5 A better architecture.

By creating “es” and “en” parent directories, all locale-specific pages can be isolated from one another. Notice how filenames are mirrored in each directory. This

strategy will come in handy when it's time to update content. If, for example, your `welcome.html` page needs updating, you know exactly where all the `welcome.html` localized pages are, no matter what the language. An added benefit is that pages don't need to be renamed, and developers can switch between languages by simply changing the "es" to "fr" or "de."

Business Rules

After you've effectively organized the site, think about how to organize the content itself. You may have certain product or service categories that make sense to Americans, but not to users abroad. Or, even if the categories do make sense, they might not be needed in your target markets. Just as Wal-Mart doesn't sell snow shovels in Florida, you shouldn't build a site that sells products that people in other countries don't need.

Business rules help you automate some of the complexities of interacting with multiple locales. Rules include anything from product selection and pricing to sales tax, shipping fees, and privacy restrictions. For example, you could have a Chinese New Year promotion planned for your Chinese market while you're simultaneously planning an Easter promotion for European and Latin American markets. Later in the book, you'll examine content management systems that have business-rule capabilities built in.

For more information on managing locale-specific content, see Chapter 13, "Global Content Management."

The better you understand the various constraints of each local market, the better you can prepare for them on a global scale.

Yahoo! and France

French law prohibits the sale of racist items, such as Nazi or Ku Klux Klan paraphernalia, yet until recently, Yahoo! regularly featured such items on its auction site. Using business logic, Yahoo! kept these items out of any auction that would be seen on the Yahoo! France site. However, in late 2000 a French judge said that this solution wasn't good enough, so in early 2001, Yahoo! banned the auction of such items altogether, joining the ranks of eBay and Amazon. Other countries don't share the same laws (or lack thereof) as the U.S., so expect your site to abide by many differing laws.

Loose Strings

Thinking globally has technical requirements as well as business requirements. For example, text strings might not seem important, but they can create major problems if overlooked. Often, web developers hard-code text strings into the

scripts stored throughout a web site. For example, when you conduct a search on a site, you'll see a response similar to this:

Your search returned 15 results.

This string is actually composed of four elements, assembled dynamically:

“Your search returned”

“15” (generated automatically)

“result”

“s” (the s is added to “result” when the number of results is greater or less than 1)

This system works for English, but in some languages, such as French or Spanish, plurals aren't always formed by simply tacking an s to the end. In other languages, such as Chinese, there is no difference between plural and singular. To avoid this problem, the string should be rewritten as follows:

181

Number of results returned: 15

The process of chaining text strings together is called *concatenation*, and concatenation simply does not travel well. Every language has its own unique patterns and idiosyncrasies. Notice how the string **Results 1 - 10 of about 475,000** changes depending on whether you're using Google U.S. versus Japan (see Figure 7.6).



Figure 7.6 Google U.S. and Japan: Similar messages, differently constructed sentences.

Solving (and preferably avoiding) concatenation problems is a job for software developers. Increasingly, software developers take all these text strings and place them in a separate resource file or database. By keeping the translatable text separate from the software, the text strings can be easily translated and the developers don't have to spend their time searching for lost strings. This book won't attempt to delve any more deeply into this area, but just be aware that all those little text strings used throughout your web site—even though they might not cost a lot to translate—could cost quite a lot to internationalize.

ACTING LOCALLY

Internationalization generally focuses on those web site elements that remain invisible to the end user—architecture, databases, modular design—but localization focuses on those elements that are visible: text, images, and the manner in which they're presented.

If you consider how companies struggle to target their web sites toward various groups of users *within* the U.S., you can understand the challenge of localizing a site for new, global markets. Every locale is going to require its own degree of customization. Even within a market, there are always smaller markets that require an even finer level of customization. Some of the changes you make will be cosmetic; others will be highly technical. Given the complexity, it's not hard to imagine a time when entire books will be devoted to localizing web sites for each market.

For this book, the focus is on the macro issues. Although there is no limit to how localized a site can be, at a minimum, you should answer three questions:

Do users understand your site?

Can they find what they're looking for?

Can they purchase what they find?

Do Users Understand Your Site?

For users to understand your site, you need to understand them. Localizing your message begins with understanding how your users think. Begin by asking the following questions:

- **What are their traditions, tastes, holidays, religions?** You won't want to promote your Halloween decorations to countries that don't celebrate Halloween.
- **How many have access to the Internet, and how do they access the Internet?** In Japan, more people access the Internet through their mobile phones than their PCs.

The better you understand your audience, the more focused your message will be. Along the way, however, be sensitive to anything that might confuse or offend your audience:

- **Colors:** Pay close attention to colors and their meanings. Just because black signifies death in the U.S. does not guarantee a similar meaning abroad. In China, white symbolizes death.
- **Flags:** Flags are best avoided because they are more political than cultural. The Canadian flag doesn't indicate whether a given web site is written in French or English and a Brazilian flag doesn't indicate Spanish or Portuguese. Unless you have a clear reason for using flags, you probably shouldn't.
- **Icons:** Web sites rely on icons to assist the user, such as the ubiquitous shopping cart icon. Yet imagine what a shopping cart means in a culture where people rarely use shopping carts. The famous garbage can icon on the Macintosh was not universally recognized because garbage cans around the world don't all look alike. The same goes for mailboxes. And in France, the house icon doesn't signify "home page" because they usually call their home pages the "welcome page."
- **People and their body language:** Some cultures are much less diverse than the U.S. and are keenly sensitive to photographs of people who do not reflect the general population. Also be sensitive to how the models are dressed and how they pose. An open palm may mean "stop" in the U.S., but is offensive to other cultures. *Body language is just as important as the written language.*
- **Writing style:** In the U.S., writers are expected to communicate messages in the first two sentences, yet this style might be considered rude in Asia, where the writing style is more subtle and the point of a message is typically located near the end of a paragraph, not the beginning.
- **Spelling and grammar:** The rules of language are rarely simple or straightforward. Hyphenation and spelling vary within the language itself (such as in English and Spanish), depending on where they're used. The use of accent marks can also be confusing. In Canadian French, the uppercase letters generally retain accent marks, but not so in France.

For more information about body language and culture-specific design, see Chapter 11, "World Wide Design."

To learn more about writing for a global audience, see Chapter 10, "Writing for the World."

Mind the (Language) Gap

A shared language does not save you from translation:

U.S.	U.K.
Sneakers	Trainers
Restroom	Loo
911	999
Cookie	Biscuit
Wrench	Spanner
Truck	Lorry
Elevator	Lift
Parking Lot	Car Park

Can They Find What They're Looking For?

Your web site needs to be usable to people around the world, yet not all people have the same Internet connections, browsers, web savvy, and preconceptions. To ensure that your site remains as usable in Norway as it is in Nebraska, pay close attention to the factors in the following sections.

Bandwidth

The broadband revolution is taking a little longer than expected. According to most analysts, the U.S. won't reach a critical mass in broadband usage until 2005 or later. The rest of the world, with the exception of parts of Asia and Europe, is much further behind in high-speed Internet connections. So if most of the world is still accessing the Internet at 56Kbps or slower, why do companies build sites that can be easily viewed only with broadband connections? Even if you're targeting businesses, which are more likely to use broadband connections, the percentage of broadband users outside the U.S. rarely justifies designing a bandwidth-hogging web site.

Note

Only 10% of households in France, Germany, and the UK will have broadband Internet access by 2005.

Source: Gartner (www.gartner.com), February 2002.

As shown in Figure 7.7, the weight of your web site in kilobytes directly affects the amount of time your audience must wait to view your site. If you want your site to be popular outside the U.S., keep its total weight at 70KB or lower. Given that the average web page weight for Fortune 500 sites is more than 90KB, odds are you have some graphics to remove. But the result will be worth it, not just for your foreign visitors, but even for a good portion of your domestic audience.

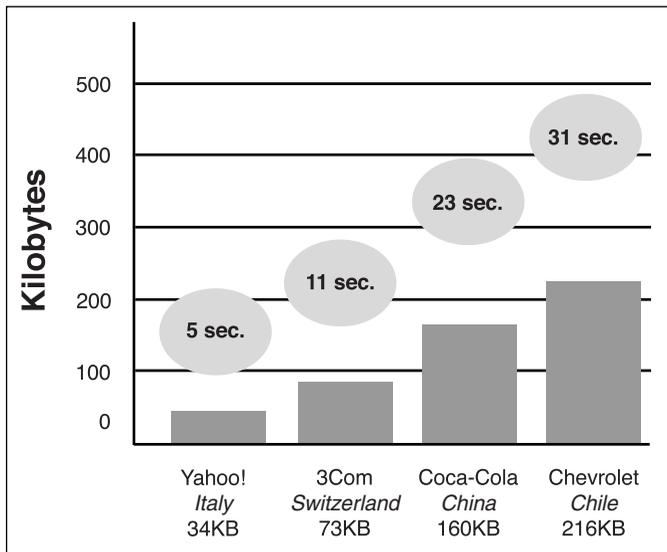


Figure 7.7 The weight and the waiting: Home page weight and download time when using a 56Kbps modem.

185

For more information on building web pages that don't keep international users waiting, see Chapter 11.

Measurements

While the world largely embraced the metric system, the U.S. largely ignored it. Today, Americans still measure their driving in miles per hour, their gas tanks in gallons, and the temperature in Fahrenheit degrees. Although these forms of measurement work fine in the U.S., they range from troublesome to meaningless in most other countries.

Paper sizes are also a source of frustration. The standard letter size in the U.S. is $8\frac{1}{2} \times 11$ inches, but the rest of the world has a different idea of "standard." The A4 standard is actually much more common. Remember the International Standards Organization (ISO)? It has a standard for paper, too: A4 (specified by ISO 216). A4 is metric-based and, at $21\text{cm} \times 29.7\text{cm}$ (approximately $8\frac{3}{4} \times 11\frac{3}{4}$ inches), is narrower and longer than the American letter.

FYI

For more information about international page sizes, go to www.cl.cam.ac.uk/~mgk25/iso-paper.html.

Why do paper sizes matter to web developers? If you want users to be able to print your web pages, you have to make sure the margins are narrower than usual. Also, many companies provide marketing brochures and white papers in PDF files for download. All too often, they are formatted in letter size, which only frustrates users who find that the pages print with the edges chopped off.

Clothing sizes could also use a bit of standardization, not just within the U.S., but globally. When Victoria’s Secret first ventured into localization, it focused on developing a customer service page in each of the target languages. A good percentage of these pages was devoted to converting sizing information across locales, as shown in the conversion chart in Figure 7.8. In the future, Victoria’s Secret plans to make this function dynamic so that users in a given locale are instantly presented with the local size, without having to convert it themselves.

Bra Size Conversions				
US	UK	France	Australia	International
32AA	32A	85A	10AA	70A
32A	32B	85B	10A	70B
32B	32C	85C	10B	70C
32C	32D	85D	10C	70D
32D	32DD	85DD	10D	70DD
34AA	34A	90A	12AA	75A
34A	34B	90B	12A	75B
34B	34C	90C	12B	75C
34C	34D	90D	12C	75D
34D	34DD	90DD	12D	75DD
34DD	34E	90E	12DD	75E

Figure 7.8 Victoria’s Secret conversion chart.

Numerical Notation

The number 2.455 means different amounts in different countries. In the U.S., 2.455 is less than 3; in Germany, it is more than 2,000. Blame numerical notation for the confusion. Not all countries use periods and commas in the same fashion.

In the U.S., the period indicates the decimal point; in Germany and other countries, the period is the thousands separator.

All too often, American companies expect users around the world to understand their numerical formats, yet numerical notation is not universal, as shown in the following chart.

Country	Notation
U.S.	123,456,789.00
France	123 456 789,00
Spain	123.456.789,00
Germany	123.456.789,00
Sweden	123 456 789,00
Russia	123 456 789,00

For a more comprehensive list of notation standards, see Appendix F, "International Notation Standards."

Phone numbers also have country-specific notation standards. Unfortunately, there's no international standard that all countries should follow. In fact, you'll often see notation differences *within* countries. In the U.S., for example, the number (555) 555-5555 can also be written as 555.555.5555 for purely stylistic reasons. When you feature phone numbers on your web site, try not to be stylistic. Your localized web sites should present phone and fax numbers in the formats that are most commonly used in each locale. Even more challenging than displaying phone numbers is making sure you can properly accept phone numbers. Make sure that input forms on your site don't force users to adhere to the American 10-digit phone number format, as the form in Figure 7.9 does.

187

Figure 7.9 Phone number entry fields need to be localized. This one is not globally friendly.

As you can see, some countries don't even have 10-digit phone numbers:

Country	Notation
Hong Kong	1234 5678
Singapore	123 4567
Sweden	12 345 67 89
France	12-34-45-67-89
Poland	(12) 345.67.89
U.K.	1234 567899
U.S.	(123) 456 7890
Brazil	5102-1800
Germany	1234 567-8

Dates and Times

The Gregorian calendar that the Western world relies on is hardly the only calendar in use around the world. There are also the Islamic, Hebrew, Buddhist, Ethiopian, Hindu, and Japanese “Genko” calendars. Each calendar has its own unique holidays, which often vary from year to year. The way dates are represented is rarely consistent, even between countries that share a common calendar. For example, when is the Fourth of July not the Fourth of July? It all depends on the locale, as shown in the following chart:

Country	Notation
U.S.	7/4/2002
France	4/7/02
Germany	4.7.02
Japan	2002 年 7 月 4 日
ISO 8601	2002-07-04

For a more comprehensive list of notation standards, see Appendix F.

There has been a global push toward adopting the ISO 8601 standard for date formats, which follows the formula **YYYY-MM-DD** (February 1, 2002, is represented as 2002-02-01). But until everyone follows this standard, there is bound to be confusion. For the time being, a more practical solution is to simply spell out the name of the month and use the full four digits for the year, such as February 1, 2002. This notation prevents any ambiguity, even between differing calendar systems.

There's also global ambiguity about how times are represented. The U.S. is the only major country to use the 12-hour a.m./p.m. notation. The 12-hour notation is

prone to error; for example, is 12:00 midnight or midday? With 24-hour notation, known in the U.S. as “military time,” 12:00 is always midday and 00:00 is midnight. This system is widely used around the world, although the exact notation varies regionally. ISO 8601 also recommends a standard time format: **HH:MM:SS**. As with the date format, the larger units start on the left.

Finally, there’s the pesky challenge of working with time zones. There are 24 global time zones, but many countries set their own time zones. For example, Canada has one time zone that spans three American time zones. Other countries have half-hour time zones. Do not assume that people in other countries will understand that “EST” means Eastern Standard Time. You could refer to GMT (Greenwich Mean Time), yet this is hardly universally understood and has since been replaced by the Coordinated Universal Time (UTC). In other words, if you must display times on your site, make sure they are fully localized for the end user.

Rules and Regulations

Every country has its own unique regulations. Although there are efforts underway to “harmonize” regulations globally, for the time being, you need to rely on legal experts in each local market to help you play by the rules. Here are some regulatory issues to consider:

189

- **Privacy laws:** Europe has much stricter rules than the U.S on collecting and sharing customer list information.
- **International advertising laws:** In the U.S., head-to-head comparisons are commonplace, but Germany and Japan don’t allow comparative advertising. Also, many countries prohibit price competitions and lotteries, and place tight restrictions on direct mail marketing.
- **Liability:** What if a translator makes a mistake that leads to an injury in another country? What are your responsibilities and liabilities, and how should you be prepared?
- **Labeling:** Should your packaging include any regional or country-specific markings or terminology? For example, the European Union closely regulates packaging labels for the health care industry. If you don’t abide by these rules for even one country, you are prohibited from entering *any* of the EU countries.

Can They Purchase What They Find?

Now that a web user has found a product, you need to ensure that he or she can easily purchase it (and receive it). Currency conversion is the first challenge.

The user naturally wants to know what the product or service costs in his or her currency. At the least, you can supply a link on your site to one of the many free currency conversion sites on the Web, such as xe.com (www.xe.com/ucc), shown in Figure 7.10. Ideally, you save your customers the added work and provide the conversion dynamically so that all they see are prices in their currency.



Figure 7.10 Do-it-yourself currency conversion.

A lot of companies mistakenly assume that because credit cards are universally used, they can just take credit card orders. However, credit cards are not universally used or trusted. In Germany, money orders are a popular method of payment. Wire transfers are also common, particularly in B2B transactions.

If you want to collect payment in Japan, you had better consider doing a deal with the convenience store chain 7-Eleven. Only about 10% of Japanese use credit cards; they prefer bank transfers, COD, and, increasingly, payment at their local 7-Eleven stores. Since 1999, web users in Japan have been able to order goods online and pay for them at 7-Eleven stores. With nearly 9,000 locations throughout Japan (compared with 400 in the U.S.), 7-Eleven is a major force in Japanese payment collection (see Figure 7.11).



Figure 7.11 7-Eleven Japan: More than just Big Gulps.

After you've offered the necessary payment options, you'll need to collect applicable taxes, clear customs for products that need to be delivered, and make allowances for the occasional return. With so many issues to be resolved, there's also a growing list of application service providers that provide full-service global payment collection, tax collection, customs clearance, and fulfillment.

Searching and Sorting

Enhancements need to be made to your search engine so that users entering a local term will get the relevant response. For example, "jumper" is another name for "sweater" in the U.K., but if you enter **jumper** in Lands' End's U.K. search engine, it returns no results, even though Lands' End offers a wide array of sweaters. Your translators and editors should play an active role in highlighting such terminology issues, which you can add to your terminology glossary as the project progresses. There are a number of technical solutions to making search engines locally friendly, but first you need to do the groundwork of building the terminology glossary.

Accented characters can also pose problems for search engines. The characters *e* and *é* are not considered equal to a search engine, but a web user might not make that distinction. For example, an American web user may want to search for the company Vésper, but type **vesper** into the search engine. How will the search engine know what the user is looking for? It all depends on the search engine and how well it has been prepared for handling different language and characters. For example, when entering **vesper** into Google, the top result just happens to be the company Vésper (see Figure 7.12).

191



Figure 7.12 If you want to find the company Vésper using Google, just enter **vesper**. Google doesn't get confused by the missing accent mark. Would your search engine?

Sorting might seem trivial, yet you'd be surprised how it affects the way people find things: Online dictionaries and site maps rely on sorting, for example. Suppose you have a dictionary of terms on your site. After you translate that dictionary, many of your terms might need to be re-sorted. When sorting, characters with accents typically follow their base characters, as shown here:

cable

câble

câblé

Not all characters sort in ways that make sense in English, however. For example, in Danish, the character æ follows z, and in Icelandic, the character ð falls between *d* and *e*. Sorting Asian languages is particularly challenging, as rules are based on a complex combination of phonetics, radical order, and number of pen strokes.

Different Languages, Different Characters

As you work with other Latin-based languages, prepare to use a variety of characters not often used in English.

Spanish: á, ch, ll, ñ, ó, ú, ü

French: à, â, ä, ç, é, è, ê, ë, î, ï, ô, œ

Danish: ñ, æ, ø, å

Swedish: å, ä, ö

Czech: á, c, e, š, r, ž, é, u, ú, í, ý

And don't forget punctuation marks, such as:

- Guillemets (French quotation marks, also used in Spanish): « and ».
- French uses the colon (:), but requires a space to be inserted on both sides.
- Spanish uses upside-down question marks and exclamation points to preface sentences, as in ¿ ... ? and ¡ ... !

Uppercase, Lowercase, and Neither

Different languages have different rules for lowercase and uppercase characters. Some languages, such as Arabic, Chinese, and Hindi, do not even have separate cases. Many American web designers make frequent use of all uppercase or all lowercase headlines for stylistic effect. However, a lowercase headline appears sloppy in a country such as Germany, where capitalization rules are more rigidly observed (see sidebar "Capitalization Case Study: German").

Capitalization Case Study: German

German follows slightly different capitalization rules than English. The beginnings of sentences, proper names, and *all* nouns (including adjectives and verbs used as nouns) are capitalized. The second person formal pronoun *Sie*, the polite form of address, is capitalized in all contexts. However, the second person familiar pronouns *du*, *dich*, *dir*, *ihr*, and *euch* are not capitalized. The first person pronoun *ich* as well as adjectives that refer to nationality, such as *russisch* (Russian) or *amerikanisch* (American), are also not capitalized.

English has its many quirks, and so do other languages. The way to be prepared for these many quirks is to work with people who understand the target languages thoroughly and understand how the target languages will influence the functionality of your source-language web site.

THE GLOBAL/LOCAL INTERNET

With this chapter, you now have a wide-angle view of the key issues in adapting a web site for the world and for a locale. There will always be a natural tension between internationalization and localization, a tug-of-war between global efficiency and local customization. As long as you don't sacrifice one for the other, however, you'll be well on your way to creating a successful, globalized web site. The following chapter delves into the most important, and most noticeable, component of localization: translation.