Segmenting Global Audiences and Adapting Digital Content for Target Users

Lorcan Ryan (Centre for Next Generation Localisation)

Large enterprises invest significant resources each year developing and adapting digital content for target users all over the world. By understanding how to segment a global audience and adapt content appropriately for target users, digital content publishers can generate multilingual content that both increases international revenues and improves usability for users.

Digital content publishers traditionally authored content in one source language and distributed it to native speakers of that source language (usually, but not always, based in the same region as that of the publisher). Digital content was usually only translated for regions that displayed sufficient potential for increasing international sales. Even then, the cost of translation, combined with the expense of physically exporting software products and associated documentation, often proved prohibitive for enterprises who focused mainly on selling and marketing to domestic or "local" audiences.

The evolution of the World Wide Web, however, facilitated the growth of international trade and the development of a global economy. Enterprises were able to advertise via email with no printing costs, sell online with no physical presence necessary, store and distribute digital content electronically with no storage or shipping costs, and provide instant customer support via email and online support centers with no telephone charges. These organizations needed to adapt as their target audiences changed from local to global in the space of a few years due to globalization (Braster 2009).

A global audience typically refers to a group of target users that are based in different regions and speak different languages. The term global refers to anything which is not bound to a place, and is independent of physical and cultural contingencies (Wiegerling 2004). In an ideal world, enterprises would be able to author a single global version of digital content which would be understandable by any user (regardless of what languages they speak), acceptable in any region (regardless of its political, economic, socio-cultural, technological, environmental or legislative nuances) and functional in any locale (regardless of different hardware and software platforms). This is practically impossible to achieve however, and content usually requires some level of adaptation for users due to the linguistic, regional, cultural and technical diversity of a global audience. Indeed the very concept of *global* is disposed to an unrealistic view of the world that assumes monolingualism, homogenization of markets and removal of local identities (Wiegerling 2004).

A global audience therefore, must be segmented into different groups of users, with digital content adapted to meet the requirements of each particular group. It may be segmented by the languages spoken by target users, the locale in which they are based or their individual characteristics, and adapt digital content for them accordingly. The simplest approach is to identify which languages users speak, and adapt the source language content so that it is understandable for them. It may not be necessary to adapt content for native and non-native speakers of the source language, but non-speakers usually need the linguistic elements of the source material to be rendered into their language to be able to comprehend it. This process is called translation, and although it ensures that digital content is understandable for non-speakers of the source language, issues such as insensitive cultural references or technical bugs may still be present.

A more comprehensive level of content adaptation is localization, where a global audience is segmented according to the locales in which target users are based. Digital content is then adapted according to the linguistic, cultural and technical nuances of these local areas. Wiegerling (2004) emphasizes that local areas are characterized not only by a geographical primacy, but also a cultural primacy. Localization encompasses the process of translation, but also adapts the cultural, cosmetic and technical aspects of digital content, so that it is both acceptable and functional for target users, as well as being understandable.

The most sophisticated method of segmenting a global audience, however, is by the individual characteristics of target users. The process of segmenting global audiences by the characteristics and preferences of target users, and adapting digital content based on individual user models, is called personalization:

Strategy	Segments Global Audience By	What Is Adapted?
Translation	Language spoken by target users	Linguistic aspects of digital content
Localization	Locale of target users	Linguistic, cosmetic and technical aspects of digital content
Personalization	Individual characteristics of target users	Linguistic, cosmetic and technical aspects of digital content

Every target user has a complex and unique set of demographic and psychographic characteristics:

- Demographic Characteristics: Age, sex, occupation, location, marital status, income, education level, nationality, first language, social category
- Psychographic Characteristics: Motivation, attitudes, perceptions, knowledge, preferences

Personalization is the process of deciding what has the highest value to an individual, based on their individual characteristics and preferences (Dunwoodie 2002). It goes beyond language and locale in adapting digital content for each individual user (sometimes referred to as a "locale of one"). This is essential due to the sheer volume of online digital content, which often makes it difficult for people to locate the right information at the right time. The dynamic nature of the internet has enabled publishers to shift from delivering websites and web pages to discrete units of information organized in ways that are relevant and personal to each individual (Fromm 2009). This new method of delivering digital content differs from the traditional "static" web in that it incorporates a user model representing the characteristics of the user, and uses it to adapt the content according to individual preferences (Brusilovsky and Maybury 2002).

Personalization may be implemented in an implicit, explicit or hybrid manner (Bowen and Filippini-Fantoni 2004). Implicit personalization is when modifications concerning the content or even the structure of a website are performed automatically by the system based on information concerning the user stored in the so-called user profile (Bonnet 2002). This user profile or model is constructed by recording the navigational behavior or preferences of each user through cookies and web server log files (dynamic profiles) (Eirinaki & Vazirgiannis 2003).

Types of Personalization	User Profile Constructed From	Personalization Implemented By
Explicit Personalization	User-generated content: - Forms - Questionnaires - Reviews	The user: The user indicates his or her own preferences
Implicit Personalization	Cookies and web server log files indicating: - Navigational behavior - User preferences	The system: The system automatically recomposes content based on the user

Explicit personalization, on the other, involves the user providing data for the user model via online registration forms, questionnaires and reviewing. Hybrid personalization utilizes a combination of implicit and explicit personalization. There are four main approaches to implementing personalization:

- Rule-based filtering
- Collaborative filtering
- Web usage mining
- Prediction based on benefit theory

Rule-based filtering enables website administrators to specify rules based on static or dynamic profiles that are then used to affect the content served to a particular user (Mobascher et al., 2000). Association rules could, for example, define that users who read Topic A and Topic B, are also likely to be interested in Topic C (Bowen and Filippini-Fantoni 2004). Collaborative filtering, on the other hand, delivers relevant material to customers by combining their own personal preferences with the preferences of like-minded others. It is achieved via recommendation engines which suggest content

that users may prefer (MacManus 2009). Recommendation engines suggest content to the user based on the individual's past behavior (personalized recommendation), the past behavior of similar users (social recommendation), or based on the current item being viewed (item recommendation). Amazon.com uses this technique to determine a user's interests from previous purchases as well as from ratings given to titles (Linden et al. 2003).

Web usage mining relies on the application of statistical and data-mining methods to the web server log data, resulting in a set of useful patterns that indicate users' navigational behaviors. User behavior is analyzed by extracting statistical information such as site activity, diagnostic, server, referrers, demographic statistics, and click stream analysis (Bowen and Filippini-Fantoni 2004). Finally, prediction based on benefit theory is based on the intuitive fact that if you find the item for a user which intrinsically benefits them, the chances are high they will like it and appreciate it.

Two current applications of personalization are adaptive hypermedia and personalized information search retrieval. Adaptive hypermedia is a method of dynamically composing digital content based on the needs of the user. It has attracted considerable attention due to its potential to provide personalized applications and services (Brusilovsky et al. 2004). Brusilovsky et al. (2004) suggest several ways in which adaptive hypermedia may function:

- An online newspaper which delivers news most relevant to reader’s interests
- A mobile tourist guide that presents information adapted to user interest and location
- An adaptive learning system (ALS) which adapts educational content and navigation based on the requirements of each individual student
- An adaptive customer support system which provides users with support documentation that is personalized to their specific needs

Personalized information retrieval involves accessing information about the user to deliver the most appropriate response for the individual’s need. This can range from search recommendations to full multilingual content search.

Personalization has some potential disadvantages (Jupiter Research 2003) such as:

- Increased costs: The cost of developing personalized websites may be up to four times those of a normal website
- Privacy concerns: Around a quarter of users may actually avoid personalized Web sites due to privacy concerns
- Usability issues: Only 8% of users are encouraged to revisit because of personalized facilities

Despite these potential disadvantages, personalization can generate numerous benefits both for digital content publishers and target users:

- Increased customer loyalty, revenue and profits for digital content publishers: More than 75% of respondents from a survey conducted by the Economist Intelligence Unit in 2007 believed that the benefits of improving customer engagement via personalization (and other methods) include increased customer loyalty, revenue and profits (Economic Intelligence Unit 2007). Personalization also enhances site “stickiness,” that is, an increased likelihood that customers will bookmark and return to your site (Dunwoodie 2002).
- Increased usability for target users: Personalization involves adapting digital content based on actual user characteristics and preferences, rather than stereotypes based on language or locale assumptions. This usually results in content that is more usable and relevant for target users.

There are several challenges to overcome in the attempt to move from translation and localization to the personalization of digital content for target users. It may be difficult to gather data about every target user’s unique demographic and psychographic characteristics, especially if users prefer to remain anonymous and ignore surveys, feedback forms and so on. Even if an enterprise has access to data about users’ demographic and psychographic characteristics, it still may not have the resources to adapt its digital content to the extent that it can deliver a personalized version of it for each individual user. Enterprises can overcome these challenges by investing in personalization tools and technologies which automate data collection and user modeling tasks.

Segmenting global audiences by target users' characteristics and preferences, and adapting digital content for them accordingly, can result in multilingual content that both increases international revenues and improves usability for users. Wellman (2009) believes there currently is a shift from place-to-place to person-to-person connectivity; enterprises embracing this move from translation and localization to personalization are the front-runners in this new paradigm of digital content development. Where the 1980s saw large-scale translation of enterprise content and the mid-1990s saw the evolution of activities such as localization engineering and cultural considerations, the next important shift for both digital content publishers and localization service providers could well be a move towards generating highly personalized multilingual digital content for end users.

About the Author:

Lorcan Ryan is currently undertaking a localization PhD in the University of Limerick, in association with the Centre for Next Generation Localisation (CNGL). He has a BBS degree in Business Studies and Marketing, and an MA master's degree in Technical Communication, both also achieved through the University of Limerick.

He has worked in the biomedical, business information and internet service provider industries, and has several years experience in the localization industry working as strategic accounts manager for a tools provider. Lorcan has also authored and delivered professional localization training courses, and lectured in several localization modules for the University of Limerick.

References:

Bonnet, M. (2002). Personalization of web services: opportunities and challenges, Retrieved 03 May 2010, from http://www.ariadne.ac.uk/issue28/personalization

Bowen, J.P. and Filippini-Fantoni, S. (2004) Personalization and the web from a museum perspective, Retrieved 03 May 2010, from http://www.archimuse.com/mw2004/papers/bowen/bowen.html

Braster, B. (2009) 'Controlled language in technical writing', Multilingual Computing & Technology, 20(1), Sandpoint: Multilingual Inc.

Brusilovsky, P. and Maybury, M.T. (2002) 'From adaptive hypermedia to adaptive web', *Communications of the ACM*, Vol. 45, No. 5.

Dunwoodie, B. (2002) "Personalization overview", Retrieved 30 April 2010, from http://www.cylogy.com/library/personalization_overview-kb.pdf

Eirinaki, M. and Vazirgiannis, M. (2003) 'Web mining for Web personalization', *ACM Transactions on Internet Technology*, 3(1), 1-27.

Fromm, K. (2009) 'The real-time web: A primer, part 3', Retrieved 03 May 2010, from http://www.readwriteweb.com/archives/the_real-time_web_a_primer_part_3.php

Jupiter Research (2003) 'Beyond the personalization myth: Cost effective alternatives to influence intent', *Jupiter Media*

Linden, G., Smith, B. and York, J. (2003) 'Amazon.com recommendations: Item-to-item collaborative filtering', *IEEE Internet Computing*, 7(1), 76-88.

MacManus (2009) Top 5 web trends of 2009: Personalization, Retrieved 03 May 2010, from http://www.readwriteweb.com/archives/top_5_web_trends_of_2009_personalization.php

Mobascher, B., Cooley, R. and & Srivastava, J. (2000) 'Automatic personalization based on Web usage mining', *Communications of the ACM*, 43(8), 142-151.

The Economic Intelligence Unit (2007) ‘Beyond loyalty: Meeting the challenge of customer engagement’

Wellman, B. (2002), "Little boxes, glocalisation, and networked individualism", in: Digital Cities II, edited by Makoto Tanabe, Peter van den Besselaar, and Toru Ishida. Berlin: Springer-Verlag, 2002, 11-25.

Wiegerling, K. (2004), "Localisation versus globalisation – Claim and reality of mobile and context-aware applications of the internet", in: International Center for Information Ethics (ICIE), Volume 2, 1-7.