出國報告（出國類別：其他<國際會議>）

第 33 屆國際萬國碼會議 (Internationalization & Unicode Conference 33)出國報告

服務機關：行政院主計處電子處理資料中心
姓名職稱：黃柏盛　設計員
派赴國家：美國加州 聖荷西
出國期間：98 年 10 月 14 日至 98 年 10 月 16 日
報告日期：99 年 1 月 14 日

目　　錄

壹、　會議介紹 2
一、　會議名稱 2
二、　會議時間 2
三、　會議地點 2
貳、　目的 3
參、　會議過程：會議議程及紀要 5
一、　會議議程 5
二、　會議紀要 5
(一)、 10 月 14 日（星期三）導引(Tutorials)會議內容： 5
(二)、 10 月 15 日（星期四）發表會議(Presentations)內容： 8
(三)、 10 月 16 日（星期五）發表會議(Presentations)內容： 11
肆、　心得建議 16
伍、　附錄 19
一、　照片集錦 19
二、　研討會議程 25
三、　研討會議程：會議介紹 28
四、　演講者簡介 39
五、　展示區廠商列表 52

壹、 會議介紹

一、 會議名稱

第 33 屆國際萬國碼會議(Internationalization & Unicode Conference 33)

二、 會議時間

2009 年 10 月 14 至 2009 年 10 月 16 日

三、 會議地點

地點：希爾頓飯店(Hilton Hotel)

住址：美國加州聖荷西

300 Almaden Blvd., San Jose, CA 95110, U.S.A.

主辦單位：物件管理團隊(Object Management Group, OMG)

貳、 目的

國際化萬國碼會議(Internationalization & Unicode Conference, IUC)係以發表探討萬國編碼(Unicode)之標準、實作及應用等技術發展會議。該會議於第 30 屆(2006 年)以前，一年舉辦兩次；第 31 屆(2007 年)起，則改為一年舉辦一次。

Unicode 主要在於提供全球語言文字符號之輸入、傳送、處理、交換、儲存及顯示的共同編碼表示，相於各個地區自行發展各自的編碼，Unicode 用來設計適用於全球編碼(Universal Code；簡稱 Unicode)。如此，才真正落實全球統一文字交換之標準，並免除資源浪費。

本次 IUC 舉辦的目的是使與會者可以直接從國際 Unicode 領域專家們認識其基本原理，並了解國際最新潮流、工具及技術；在相關議題上另有找尋有關安全的風險及預防的評估，並增加有關雲端運算網路及整合社群網路的議題；在應用上，展示如由 Joomla 或 PHP 開發的多語言之網頁應用程式(web application)及在移動式應用程式(mobile application)中用到的表情符號及如何修飾。

IUC 邀請了 Unicode 專家、實作者、使用者及廠商，透過研討會的互動方式，促使該會議成為與專家討論交換理念的地方，並可提供

找尋潛在需求機會，或從中獲得有關新出的 Unicode-enabled 產品訊息。

本次研討會共舉辦三天；第一天全天為 Unicode 的導引(Tutorials)及第二、三天為發表會(Presentations)議程。第一天的 Unicode 導引主要介紹與 Unicode 有關的基本原理及相關應用如字元及其編碼、文字(writing)系統、Web 國際化等。第二天議程分：程式語言、字型排印、Unicode 新消息、i18n 新消息、Open Source 等方面之議題發表與討論，三天議程分：發展平台、行動式程式、國際化實務、轉移服務 API、個案研究等方面之議題發表與討論。另外，在第二天開放展示區，展示廠商之相關應用產品。

參、 會議過程：會議議程及紀要

一、 會議議程

如附件二(研討會議程)。

二、 會議紀要

(一)、10 月 14 日（星期三）導引(Tutorials)會議內容：

上午會議 1：0900 ~ 1030
Track 2：Internationalization: An Introduction, Part I: Characters and Character Encodings
主講者：Addison Phillips, *Globalization Architect, Lab126(Amazon)*
內容概要：國際化是軟體產品的設計與發展，能適用於多種不同文化、區域或語言目標使用者。
電腦對世界的看待是從位元(bit)到位元組(byte)。Glyph是指顯示於螢幕的文字單元，是一個圖像。Character是指單一的邏輯性文字單元。Coded Character Set是一個字元集其中每一個字元都有其單一的數值編碼，如Unicode、ASCII、ISO 646或JIS X 208…等。所有的文字都有一個字元編碼。Mojibake指的是亂碼，其原因通常為顯示的文字是錯誤的編碼、同一編碼轉換二次、轉換到不適當的編碼。
一個byte代表一個字元對多數拼音語言是足夠的。但僅能容納256的文字，對很多語言而是不足的，需要更多的bytes才足夠表示。不同的編碼方式可能造成同一個文字會對應到不同的編碼，在資料傳輸上可能會造成轉換的問題。Unicode/ISO-10646即為統一使一個文

字可以對到一個碼，以減少資料編碼轉換問題。Unicode目標是要使一個字元集可以支援到世界上所以的語言文字系統。其基本原則是屬世界性的屬性、邏輯性排列、有效率的、統一的、是字元而非字體及穩定的。相關資訊參考網站www.unicode.org/reports。

上午會議 2：1045 ~ 1230

Track 2：Internationalization: An Introduction, Part II: Enabling

主講者：同上午會議1

內容概要：國際化途徑：全球範圍需求蒐集(Gather requirements globally)、生效(Enable)、具體化(Externalize)、客製化(Customize)、Localize。

國際化的議題：文字處理、語言、地區表示(Locale-affected)格式、區域表示(Reginoally-affected)格式、間時相關議題、文化習慣(Cultural adaptation)、合法性的需求。生效：使軟體所顯示、處理、驗證、儲存及傳輸的資料能適合於使用者的文化、語言及區域性，包含文字、編碼、地區性的表示方式及時間區域表示格式。Local係指一種識別或資料結構容許軟體程式在系統內存取文化的或語言性的功能。具體化：從編碼中移出因語言及文化有影響的資料及元件。Localization係指將處理產品使之適合於特定的目標市場，包括訊息的轉換、適合當地取向及內容或特徵的增減。在Localization過程中應避免程式分岔(不同locale有不同的程式處理)。客製化項目包括郵政地址的驗證、郵證碼驗證、電話號碼格式、個人姓名格式等。

下午會議 1：1330 ~ 1530

Track 2：Web Internationalization – Standards and Best Practices

主講者：Tex Texin, *Xen Master XenCraft*

內容概要：描述定義I18N使用於web的架構及原則，界定markup language的範圍，提供web與國際性資料協作的實務建議，包括多語言網頁及localization考量。HTML或XML如不設定字元編碼，將資料顯示容易產生亂碼。

HTML：<META…；charset=”UTF-8”>；XML：<?xml…encoding=”UTF-8”?>；CSS2：@charset “UTF-8”；

文字版面配置支援如下：

Feature			
Lang()	y	y	N
Lang pseudo-class	N	Y	Y
Lang attr selector	N	Y	Y
Quote:qo	N	Y	½
Text-transform	Y	Y	Y
Css list-style-type	N	Y	½
Xsl number	Y	N	N
Xsl format-number	Y	Y	Y
Html bi-directional text	Y	Y	Y
Css bi-directional text	Y	Y	Y
Vertical text (SVG losing ground)	Y/N	N	N
Ruby annotation	½	N	N
Css3 combined sort	N	N	N
Xsl:sort	Y	Y	N

網頁設計在商業上的需求也許與地區特性的要求可能不相容，較簡單的方式就是不同的地區有不有的Domain名稱。另一種是為網頁加入語言的選擇。

下午會議 2：1545 ~ 1745

Track 2：Creating XHTML/HTML Pages with Right-to-Left Scripts

主講者：Richard Ishida, *Internationalization Lead, W3C*

內容概要：在本議程說明如何開發XHTML 及HTML網頁有關阿拉伯文及希伯來語內容文字書寫方式。

同時也調查對於雙向書寫方式語文，如何以最好的方式來達成正確而有效地使用appropriate markup、CSS屬性及Unicode編碼實體。在此不僅含蓋基本，並且提供地方說母語的人特殊情況處理的技術建議。在此假設與會的人熟悉阿拉伯文的雙向書寫文字且有HTML及CSS的基本知識。

Adobe集團歡迎會(Welcome Reception)：1800~1900

地點：Adobe大樓(Adobe Building, West Tower Patio, 345 Park Avenue, San Jose)

(二)、10 月 15 日（星期四）發表會議(Presentations)內容：

Keynote 發表：0915~10:00
主題：The Alphabetic Principle and its Enemies
主講者：Nicholas Ostler, *Chairman, Foundation for Endangered Languages*
內容概要：拼音字母原則是一個符號即表示一個發音。歷史性的發展有三個階段：第一階段為腓尼基人(Phoenician)的亞拉母語(Aramaic)發音符號稱為 adjad，為只有子音；第二階段為希臘語(Greek)發音符號稱為 alphabet，有子音與母音；第三階段為希伯來語(Hebrew)/希臘語(Greek)/阿拉伯語(Arabic)發音符號稱為 abugida，帶有重音及音韻。一個符號表示一個發音會隨著時間而衰微，因為發音改變比拼字快、意外屬性改變閱讀本質。實際上拼音文字的優點在於學習上的精簡，但在學習系統的技巧使用通常不如可以掌控的專家。拼音文字的挑戰有：同音異字及多符號集。

展示區(Exhibit Arrea)開放時間：1000~2000
地點：Market Room, Lobby Level

會議 1：1030 ~ 1120
Track 2：Designing & Developing Pan-CJK Fonts for Today
主講者：Dr. Ken Lunde, *Senior Computer Scientist, Adobe*
內容概要：Pan-CJK字體包含多重CJK locale所適用的字形，以中國、台灣、香港、日本及韓國為主要的CJK locale，並為Unicode-based。

Pan-CJK字體的目標是對單一locale的CJK字體完全以一個字形對映到一個字碼，且多重locale的CJK字體要多個字形對映到多個字碼。CJK 統一表意文字編碼對映到一個以上(同義)字形，其中URO(Unified

Repertoire & Ordering)有20,940碼、Exension A有6,582碼、Extension B有42,711碼、Extension C有4,149碼、Extension D有223碼。
Pan-CJK實作方法：TrueType字體集、OpenType、分配64k字體限制。OpenType支援Mac OS X、Windows。

會議 2：1130 ~ 1220
Track 1：Implementing International Calendars in JavaScript
主講者：Umesh P. Nair, *Google Inc.*
內容概要：各地區對日期都有不同的表示方法。有以太陽曆法、月曆法、太陽月曆法或其他。甚至一日開始和一週開始也不同。目前Gregorian(西曆)為最普遍但可能不是最好的，而且其他文化的曆法也被延用至今，如中華文化的農民曆法。很多日曆軟體以西曆為主，但有些軟體也支援其他曆法。
目前軟體對曆法的支援如下：

Software	Hijri	Hebrew	Chinese	Indian	Coptic	Persian	Buddhist
Libraries							
ICU	✓	✓	✓		✓		✓
Jodatime	✓				✓		✓
Desktop software							
MS Outlook	✓	✓	✓	✓			
IBM Lotus notes	✓	✓					
GNU Emacs	✓	✓	✓		✓	✓	
OS/WM							
KDE	✓	✓				✓	
Online							
Google Calendar	✓		✓				

支援非西曆的的線上日曆是具有挑戰性而且複雜的。線上西曆實作先訂一固定偏移量起始日，再透過javascript的日期操作函式完成。線上非西曆實作除上述實作再參考Look-up table，但只有記錄50年的時間範圍。

會議 3：1330 ~ 1420
Track 2：The Design & Development of Fully Proportional Japanese Fonts
主講者：Ken Lunde, *Senior Computer Scientist, Adobe*
內容概要：在十數年來，字形受到抑制的日本文字，Adobe日本列印工程團隊開發了視覺上多樣化的設計新字體「kazuraki」。該字體源自日本12世紀的書法名家藤原

定家的字跡。典型的日本字體為等寬字體，使用正方空間，kazauraki字體採用非固定Adobe-Japan1-x字體，其中有兩個目標：比例字體寬度為預設及兩個直書平假名之間有連線為預設。

會議 4：1430 ~ 1520

Track 3：Language Identification and Usage

主講者：Mark Davis, *Sr. Internationalization Architect Google Inc.*
Addison Phillips**,** *Globalization ArchitectLab126 (Amazon)*

內容概要：文字處理通常與語言有很大關係，而IT系統也與語言的協調及溝通有關。軟體語言敏感區在於作業處理、資料格式及解析。。Locale意味著語言混亂，同樣的一個意思，在Locale有不同的表示方法如日期。ISO 639為國際語言編碼，分為ISO 639-1到ISO 639-6，以中文ISO 639-1為「zh」、ISO639-2為「chi」或「zho」。但ISO 639不適用於IT上，因為無法區別文字多樣性。IETF(Internet Engineering Task Force，互聯網工程工作小組)的BCP 47即可滿足上述問題。

會議 5：1600 ~ 1650

Track 2：A Systematic Approach to I18N Testing

主講者：Michael Manca, *Project Manager and Solution Quality Analyst IT Flex Services Intel Corporation*
Tomas Galicia, *Solutions Quality Analyst IT Flex Service Intel Corporation*
Loic Dufresne de Virel, *Localization Strategist IT Flex Services Intel Corporation*

內容概要：I18N問題：有那些問題會使軟體不能有效執行？如當英文版軟體開發完成後要改成其他語版時又耗時費力同時英文版會繼續改善版本，因此非英文版軟體總是落後原開發語言版本。
較好的方法是非文字(字串)程式碼及文字(字串)分開獨立處理。因此非英語版及軟體版本可同時演進。但這就產生了一些挑戰，如有時在地化特性不是開發人員所想到的或軟體開發人員相信這些編碼(程式碼)都

已國際化但並不然。

I18N testing可以manual方式測試，從需求review、使用者界面review、程式碼review及在localized作業系統執行程式，或者利用測試工具進行程式碼掃描、單元測試於localized作業系統等。

會議 6：1700 ~ 1750

Track 2：Google Internationalization Quality Control Framework

主講者：**Andrew Swerdlow,** *Internationalization Tech Program Mng Google Inc.*

Manish Bhargava, *Google Inc.*

Jens Riegelsberger, *Google Inc.*

Laura Cuozzo, *Google Inc.*

內容概要：Google的期許「組織世界資訊並使之全球都可取得及使用」，然而，這已超出英語的範圍，而Google已成功40種語言產品。Google透過i18n libraries標準化、轉譯檢核、可用性測試、使用者調查及使用者需求研究以朝向上述目標。但相關的活動有些不具全球規模，而且是使用者的片斷經驗，甚至是系統引發的問題。

Google透過一套Framework以控制品質。分成語言/轉譯、互動設計、可視性設計、特性遺失、資料品質、Bugs及其它等議題，經由工作經驗與問題回報以語言/轉譯議題最多。下步即以此Framework推到所有的產品。

Conference歡迎會(Welcome Reception)：1800~2000

地點：Market Room, Lobby Level

(三)、10 月 16 日（星期五）發表會議(Presentations)內容：

會議 7：0900 ~ 0950

Track 2：International Features of iPhone OS

主講者：Brent Ramerth, *Software Engineer Apple, Inc.*

內容概要：iPphone OS文字編碼以Unicode-based，以Bundle-driven在地化，支援超過30種語言。OS 3.0版增加雙向書寫文字支援、copy/paste及更多語言支援。

國際化的架構上提供容易使用國際化的標準APIs及工具、NSString為支援Unicode及NSLocale資料在地化，從kernel到應用程式皆支援Unicode，基本物件為字串。在地化的架構上iPhone應用程式成一個包裝，包含可執行的資援，所有在地化儲放在一個包裝為NSBundle類別。

會議 8：1000 ~ 1050

Track 2：Emoji in Unicode: Cell Phones Meet the Internet

主講者：Markus Scherer, *Unicode Software Engineer Google Inc.*
Katsuhiko Momoi, *Staff Test Engineer & I18n Consultant Google Inc.*
Mark Davis, *Sr. Internationalization Architect Google Inc.*

內容概要：Emoji為日文諧音即E (絵：picture) + moji (文字：letter/character)主要使用在日本手機等行動通訊設備，用以表示心情，多以臉部表情呈現。搭載標準set建置在手機，並以文字編碼方式排序及傳送於email或SMS(文字簡訊，Short Messaging Service)。

有三家公司(docomo, KDDI, Softbank)作出三個Emoji set，Google Emoji則整合這三個集合。利用Unicode + PUA(Private Use Area)編碼這些Emoji。

會議 9：1110 ~ 1200

Track 2：Taking Moblin to the World

主講者：Loic Dufresne de Virel, *Localization Strategist IT Flex Services Intel Corporation*
Michael Kuperstein, *Senior Localization Engineer IT Flex Services Intel Corporation*
Margie Foster, *Localization Project Manager Moblin Project Intel Corporation*

內容概要：Moblin是一個基植於Linux作業系統開放碼，用於移動網際網路設備和其他新的類別的設備如netbook或nettop。由Intel推出。從I18N的立場來看，開發碼不是完美的保證。Moblin組織在原始碼對國際化的處理是：Code reviews、Pseudo-builds及I18N-focused testing。另外也請專家協助處理建置。
Moblin相關網站：http://Moblin.org，
http://translate.moblin.org，http://bugzilla.moblin.org

會議 10：1300 ~ 1350

Track 1：Internationalization in Database Drivers for C/C++/Java/.NET Applications

主講者：Sumit Sarkar, *i18n Product Specialist DataDirect Technologies*

內容概要：資料庫資料存取對於語系有很大的關係，在資料庫資料存取層(Data Access Layer)有相對的Code Page定義，所涉及到有資料庫的協定、存取的APIs及其Driver。而資料庫元件直接影響到對Unicode的支援。不同的Data Access標準對字元集(Character Set)也有不同的式。如下圖：

Database	Protocol Encoding	Unicode Data Types	Different Charset per Column
DB2	Varies (chosen by client and server)	Graphic Strings are UTF-16	Yes
Oracle	DB's primary charset	N-types are UTF-8 or UTF-16	No
SQL Server	UTF-16	N-types are UTF-16	Yes
Sybase	DB's primary charset	Uni-types are UTF-16	No
PostgreSQL	UTF-8	None (DB must be Unicode)	No
MySQL	DB's primary charset	Defined character data types can be defined as ucs2 or utf8	Yes

JDBC的字集轉換在JVM。廠商如果需要時可以加入code page支援，但字集轉換工作會降低效率，某些資料也可能會遺失，轉換工作建議在從資料庫接收資料到driver level後並在送出前處理。

會議 11：1400 ~ 1450

Track 2：CLDR on the Cloud

主講者：**Benedicto Franco Jr.,** *Software Engineer Yahoo! Inc.*
Marco Aurelio Carvalho, *Senior Software Engineer Yahoo! Inc.*

內容概要：CLDR(Common Locale Data Repository)為如時間格式、文字排序及選擇語言國家名時，包容資料使用在軟體國際化和在地化。CLDR資料常為軟體固定的基礎元件，當要更新CLDR版本時可能會對新的應用程式佈署有所衝擊。

上述問題的解決方案為利用AJAX技術透過pipeline處理轉換CLDR資料，應用程式經由APIs取用在地化有效的資訊。

會議 12：1510 ~ 1600

Track 1：Extended Linguistic Services in Windows 7

主講者：Ryan Cavalcante, *Software Development Engineer Microsoft*
James Lyle, *Program Manager Microsoft*

內容概要：延伸語言服務(Extended Linguistic Services, ELS)為Windoe 7中的集中管理平台,可以使開發人員透過API取用多語系服務。在ELS的概念上分為：App Layer、Platform Layer及Service Layer三層；ELS即為Platform Layer。應用程式在App Layer，透過Platform Layer的ELS提供語言的偵測及轉譯。

會議 13：1610 ~ 1700

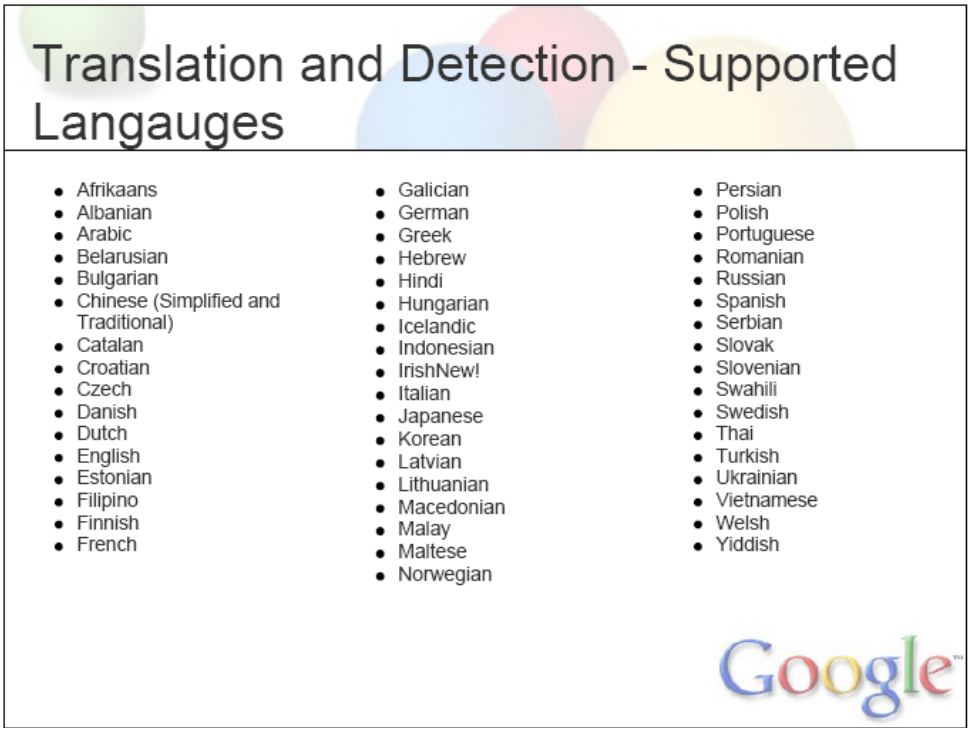
Track 1：Google APIs for Text Input and Translation

主講者：Frank Yung-Fong Tang, *Sr. Software Engineer Google Inc.*
Wenchao Tong, *Software Engineer Google Inc.*

內容概要：Google 的 AJAX API 可以利用 JavaScript 和 HTML 建置內容豐富的動態網站。API分為地圖、搜尋、資訊提供、視覺化、語言、程式庫和地球等7大類。透過「AJAX語言API」即可轉譯並偵測網頁內文字區塊的語言。

AJAX語言API有：語言轉換及偵測、虛擬鍵盤及音譯

等功能。在語言轉換及偵測功能支援的語言如下：



參考相關資訊在 http://code.google.com/intl/zh-TW/apis/ajaxlanguage/documentation/ 。

肆、 心得建議

自電腦被發明以來，即以資料為主軸，其中包含資料的表示、儲存、處理及顯示等都是很重要的課題，而電腦的最終使用者是我們人類，但全球人類存在著不同的語言及文字，因此，當電腦普及於全體人類時，即必須面對不同的語言文字。由此各語言文字的表示、儲存、處理及顯示即有大大不同。然而資料在各國之內的交換裡只要該國制訂成一協訂即不是很大的問題，但如國際間的資料要交換時就會形成很複雜的問題。此複雜的存在問題，須有統一的編碼用以解決，Unicode 編碼即應運而生。

Unicode 編碼以全域的角度，來設計全球通用的編碼，提供全球國際間的資料表示、儲存、處理及顯示等用途，以避免各國資源的浪費，並達成全球統一編碼交換的理想。目前而言Unicode為國際間所認同的編碼，並且各國大型資訊廠商也積極投入Unicode的相關應用及推動。

我國「CNS11643國家標準中文交換碼」，為解決國內不同編碼系統，如電信碼、財稅碼、戶役政碼、地政碼及Big5等之間的資訊交換問題，並持續蒐集修正國內使用的中文字形及相關屬性資料。為使我國資訊能與國際接軌，推動國際化是極為重用的工作任務。參加國際性會議(研討會)即為國際化工作之一。就參加本次IUC的建議如下：

1.積極參與國際編碼標準之相關會議：就目前的編碼系統而言，

Unicode編碼含概了各國語言文字，是為包含最多語系的編碼系統。因此將來極有可能為全球各國語言文字共同依循之編碼標準。而我國CNS11643碼所收納的文字有10萬多字，是很龐大的數量。因此，要與國際編碼相融合是亟需進行的方向。積極參與國際編標準之相關會議，不但可以提昇我國的能見度，並可以汲取外國相關經驗，對我國編碼系統與相關軟體的發展有很大的幫助。

2.加強整合國內各編碼系統，並發展CNS11643與Unicode界接相關軟體：目前我國所使用的編碼系統如大五碼(Big5)、電信碼、戶政EUC碼、財稅碼、稅務碼等種類繁多，而各編碼系統都應用於各不同的行政資訊系統，不可能一時將各編碼廢除改用統一標準編碼。因此如何整合各編碼系統即為一大課題，各編碼系統所含蓋的字集也有所差異。CNS11643碼為整合各編碼系統而成的交換碼。如此，各編碼系統既可各自持續保存發展，又可以在不同的編碼系統之間文件交換。但在各編碼間轉換過程中有可能為因為字集的差異無法轉成，如何使各編碼間轉換得順暢即可為加強各編碼系統之整合。

另外，發展Unicode與CNS11643碼的界接相關軟體，國內各編碼系統即可透過CNS1143碼與Unicode應用相銜接。

3.積極提交尚未納入Unicode編碼的CNS11643字形：目前全字庫網站所收納的字數約有10萬字，而Unicode所含中文字約8萬多字，約有1萬多字未含在Unicode內，此對我國文字在國際間文字系統

交換文件將會產生一些間隙。對於尚未納入Unicode編碼的CNS11643字形將積極提交ISO/IEC JTC1/SC2/WG2/IRG組織，納入Unicode編碼，以促使縮減國際間中文字文件的間隙。

伍、 附錄

一、 照片集錦

圖 1：開會地點(Hilton 飯店)

圖 2：會議現場

圖 3：會議現場

圖 4：會議現場

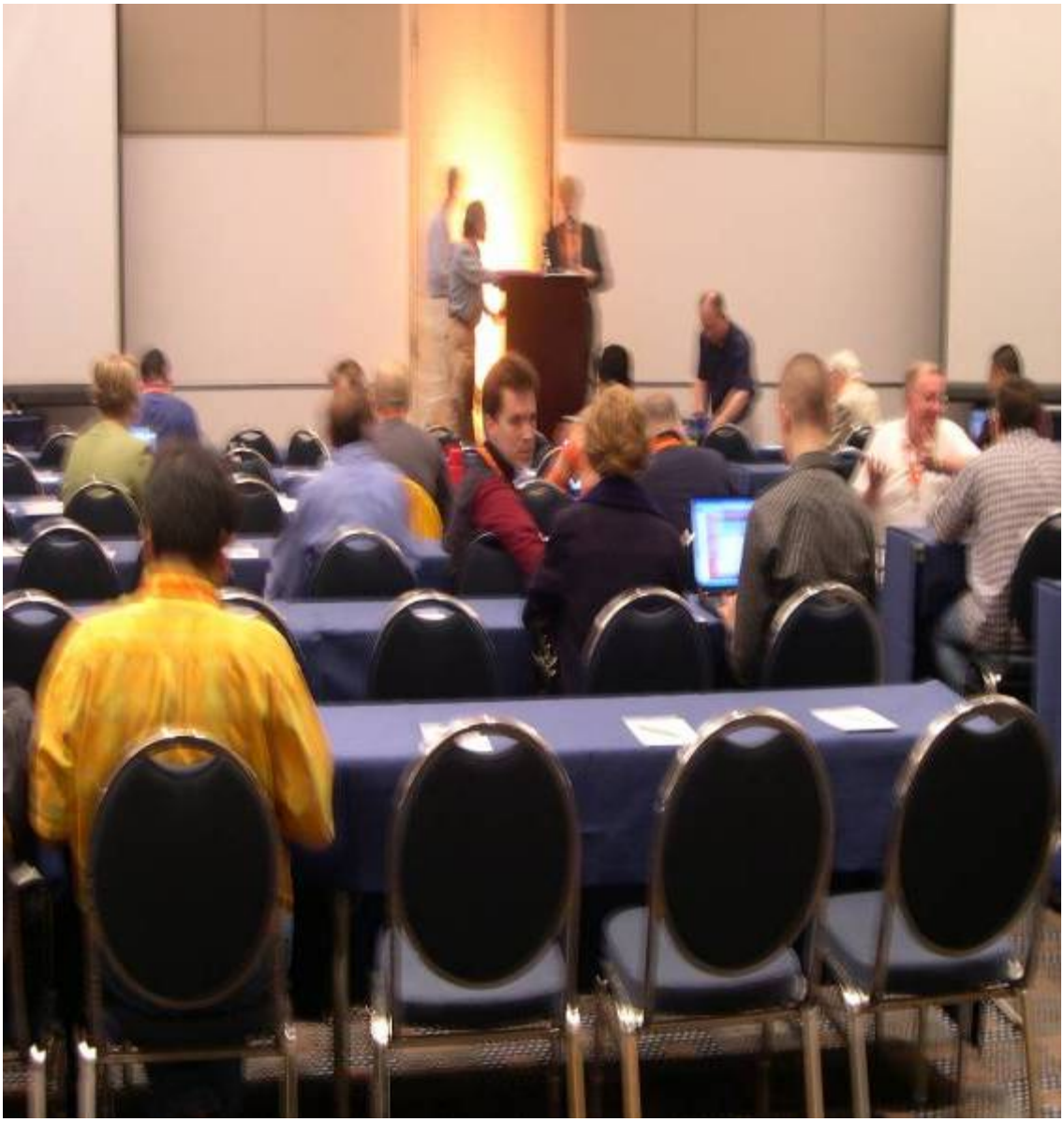


圖 5：會議現況


Your Gateway to
Global Software
WinSoft
TASMEEM
ScribeDOOR

圖 6：展示區現況

二、 研討會議程

CONFERENCE AT A GLANCE

Wednesday, October 14, 2009

TUTORIALS	Track 1 - Almaden I	Track 2 - Almaden II	Track 3 - Winchester
Morning SESSION 1 09:00 – 10:30	**An Introduction to Writing Systems & Unicode** Richard Ishida *Internationalization Lead* *W3C*	**Internationalization: An Introduction, Part I: Characters and Character Encodings** Addison Phillips *Globalization Architect* *Lab126 (Amazon)*	**Building a Custom Keyboard Layout for the Mac with Ukulele and XML** Elizabeth Pyatt *Instructional Designer* *Penn State*
10:30 - 10:45 Morning Refreshments - Ballroom Foyer, Lobby Level			
Morning SESSION 2 10:45 – 12:30	**An Introduction to Writing Systems & Unicode** (Cont'd.)	**Internationalization: An Introduction, Part II: Writing Global-Ready Code** Addison Phillips *Globalization Architect* *Lab126 (Amazon)*	**Arabic Script: Structure, Geographic and Regional Classification** Thomas Milo *President* *DecoType*
12:30 - 13:30 Lunch - Santa Clara & San Carlos, Concourse Level			
Afternoon SESSION 1 13:30 – 15:30	**Unicode - a Grand Tour** Craig R. Cummings Michael G McKenna *Internationalization Architects* *Yahoo! Inc.*	**Web Internationalization - Standards and Best Practices** Tex Texin *Xen Master* *XenCraft*	**Building Multilingual Websites in Joomla [Drupal]** Jim DeLaHunt *Principal* *Jim DeLaHunt & Associates*
15:30 – 15:45 Afternoon Refreshments - Ballroom Foyer, Lobby Level			
Afternoon SESSION 2 15:45 – 17:45	**Unicode - a Grand Tour** (cont'd.)	**Creating XHTML/HTML Pages with Right-to-Left Scripts** Richard Ishida *Internationalization Activity Lead* *W3C*	**Free Software stack for Unicode Text Rendering** Behdad Esfahbod *Software Developer* *Red Hat/GNOME*
18:00-19:00 Welcome Reception *hosted by* Adobe Systems Adobe Building, West Tower Patio, 345 Park Avenue, San Jose			

Thursday, October 15, 2009

08:00 - 09:00 Continental Breakfast - Santa Clara, Concourse Level			
09:00 - 09:15	*Welcome & Opening Remarks - Almaden Ballroom, Lobby Level*		
09:15 – 10:00	***Keynote Presentation: The Alphabetic Principle and its Enemies*** Nicholas Ostler *Chairman* *Foundation for Endangered Languages*		
10:00 – 20:00 EXHIBIT AREA OPEN - Market Room, Lobby Level			
10:00 – 10:30 Morning Refreshments - Market Room, Lobby Level			
PRESENTATIONS	Track 1 - Almaden I Programming Languages	Track 2 - Almaden II Fonts and Typography	Track 3 - Winchester Unicode News
10:30 – 11:20 SESSION 1	**Internationalization with PHP** Kirti Velankar *Senior Software Engineer* *Yahoo! Inc.*	**Designing & Developing Pan-CJK Fonts for Today** Ken Lunde *Senior Computer Scientist* *Adobe*	**Unicode Update: Unicode 5.2 and CLDR 1.7** Mark Davis *Sr. Internationalization Architect* *Google Inc.*
11:30 – 12:20 SESSION 2	**Internationalization in Ruby 1.9** Dr. Martin Dürst *Aoyama Gakuin University*	**Unicode & Fonts: a status report** Kamal Mansour *Manager of Non-Latin Products* *Monotype Imaging*	**Patching Holes in the Unicode Pipeline: A Status Report on the Unencoded Scripts of Asia and Africa** Deborah Anderson *Project Leader, Script Encoding Initiative, Department of Linguistics, UC Berkeley* Richard Cook *Post-Doctoral Researcher, Dept. of Linguistics* *UC Berkeley* Charles Riley *Catalog Librarian for African Languages* *Yale University* Anshuman Pandey *C.Phil. History* *University of Michigan*
12:30 - 13:30 Lunch - Santa Clara & San Carlos, Concourse Level			

PRESENTATIONS	Track 1 - Almaden I Programming Languages (cont'd)	Track 2 - Almaden II Fonts and Typography (cont'd)	Track 3 - Winchester I18n Standards News
13:30 – 14:20 SESSION 3	**Internationalization for JavaScript applications** Norbert Lindenberg *Internationalization Architect* *Yahoo! Inc*	**The Design & Development of Fully Proportional Japanese Fonts** Ken Lunde *Senior Computer Scientist* *Adobe*	**Update on Internationalized Domain Names and Internationalized Resource Identifiers** Dr. Martin Dürst *Aoyama Gakuin University*
14:30 – 15:20 SESSION 4	**Implementing International Calendars in JavaScript** Umesh Nair *Software Engineer* *Google Inc.*	**The Unicode-based Koran: a Conflict Between Calligraphic Tradition and Computer Typography** Thomas Milo *President* *DecoType*	**Language Identification and Usage** Mark Davis *Sr. Internationalization Architect* *Google Inc.* Addison Phillips *Globalization Architect* *Lab126 (Amazon)*
15:20 – 16:00 Afternoon Refreshments - Market Room, Lobby Level			
PRESENTATIONS	**Track 1 - Almaden I Open Source Libraries**	**Track 2 - Almaden II Assuring Quality**	**Track 3 - Winchester Scripts**
16:00 – 16:50 SESSION 5	**What's New with ICU** Steven Loomis *Software Engineer* *IBM* Markus Scherer *Unicode Software Engineer* *Google Inc.*	**A Systematic Approach to I18N Testing** Michael Manca *Project Manager and Solution Quality Analyst* *IT Flex Services* *Intel Corporation* Tomas Galicia *Solutions Quality Analyst* *IT Flex Service* *Intel Corporation* Loic Dufresne de Virel *Localization Strategist* *IT Flex Services* *Intel Corporation*	**Investigation of Opaque Glyphs Synthesized from Old Hanzi** Toshiya Suzuki *Research Assistant* *Hiroshima University*
17:00 – 17:50 SESSION 6	**HarfBuzz, the Free and Open OpenType Shaping Engine** Behdad Esfahbod *Software Developer* *Red Hat/GNOME*	**Google Internationalization Quality Control Framework** Andrew Swerdlow *Internationalization Tech Program Mng* *Google Inc.* Jens Riegelsberger Manish Bhargava Laura Cuozzo *Google Inc*	**Math Editing and Display in Microsoft Office** Murry Sargent III *Partner Software Design Engineer* *Microsoft*
18:00 - 20:00 Conference Reception - Market Room, Lobby Level			

Friday, October 16, 2009

PRESENTATIONS	Track 1 - Almaden I Development Platforms	Track 2 - Almaden II Mobile Programming	Track 3 - Winchester Internationalization in Practice
09:00 – 09:50 SESSION 7	**International Features of Mac OS X Snow Leopard** Douglas Davidson *Software Engineer* *Apple, Inc.*	**International Features of iPhone OS** Brent Ramerth *Software Engineer* *Apple, Inc.*	**Practical "Unicode Logic" for Online Tech Courses** Elizabeth Pyatt *Instructional Designer* *Penn State*
10:00 – 10:50 SESSION 8	**Windows 7: Writing World-Ready Applications** Derek Murnam *Senior Program Manager* *Microsoft Corporation*	**Emoji in Unicode: Cell Phones Meet the Internet** Markus Scherer *Unicode Software Engineer* *Google Inc.* Katsuhiko Momoi *Staff Test Engineer & I18n Consultant* *Google Inc.* Mark Davis *Sr. Internationalization Architect* *Google Inc.*	**Creating an I18n Project Plan** Adam Asnes *President* *Lingoport, Inc.*
10:50 - 11:10 Morning Refreshments - Ballroom Foyer, Lobby Level			

11:10 – 12:00 **SESSION 9**	**Accessing Globalization Services on Multiple Operating Systems** Mihai Nita *Globalization Architect* *Adobe Systems, Inc.*	**Taking Moblin to the World** Loic Dufresne de Virel *Localization Strategist* *IT Flex Services* *Intel Corporation* Michael Kuperstein *Senior Localization Engineer* *IT Flex Services* *Intel Corporation* Margie Foster *Localization Project Manager* *Moblin Project* *Intel Corporation*	**Building a Global Names System: A Case Study** Cindy Conlin *Senior Engineer* *The Church of Jesus Christ of Latter-Day Saints*
12:00 - 13:00 Lunch - Santa Clara & San Carlos, Concourse Level			
PRESENTATIONS	***Track 1 - Almaden I*** ***Development Platforms (cont'd)***	***Track 2 - Almaden II*** ***Leveraging CLDR***	***Track 3 - Winchester*** ***Internationalization in Practice (cont'd)***
13:00 – 13:50 **SESSION 10**	**Internationalization in Database Drivers for C/C++/Java/.NET Applications** Sumit Sarkar *i18n Product Specialist* *DataDirect Technologies*	**Deploying the Common Locale Data Repository (CLDR)** Steven Loomis *Software Engineer* *IBM* Mark Davis *Sr. Internationalization Architect* *Google Inc.*	**Unicode Transformations and Security Vulnerabilities** Chris Weber *Casaba Security*
14:00 – 14:50 **SESSION 11**	**Unicode Technology and Globalization Support in IBM UNIX, AIX** Su Liu *AIX Globalization Architect* *IBM*	**CLDR on the Cloud** Benedicto Franco Jr. *Software Engineer* *Yahoo! Inc.* Marco Aurelio Carvalho *Senior Software Engineer* *Yahoo! Inc.*	**My Unicode Disk Storage Went into the Circular File** Tex Texin *Xen Master* *XenCraft*
14:50 – 15:10 Afternoon Refreshments - Ballroom Foyer, Lobby Level			

PRESENTATIONS	***Track 1 - Almaden I*** ***Translation Services API***	***Track 2 - Almaden II*** ***Bidirectional Text***	***Track 3 - Winchester*** ***Case Studies***
15:10 – 16:00 **SESSION 12**	**Extended Linguistic Services in Windows 7** Ryan Cavalcante *Software Development Engineer* *Microsoft* James Lyle *Program Manager* *Microsoft*	**Mashing-up Bi-Di** Adil Allawi *Technical Director* *Diwan Software Limited*	**Twanguages of the World: a Language Census of Twitter** Jim DeLaHunt *Principal* *Jim DeLaHunt & Associates*
16:10 – 17:00 **SESSION 13**	**Google APIs for Text Input and Translation** Frank Yung-Fong Tang *Sr. Software Engineer* *Google Inc* Wenchao Tong *Software Engineer* *Google Inc.*	**Bidirectionalization: Demystifying Bidi Enabling** Roozbeh Pournader *Internationalization Specialist* *HighTech Passport*	**Banking in the Cloud: Challenges of Internationalizing Banking Software (Case Study)** Ilya Shtein *IT Architect* *Metavante*

三、 研討會議程：會議介紹

Conference Program - Session Descriptions

Wednesday, October 14, 2009

09:00-10:30 MORNING TUTORIALS

Presenter:
Richard Ishida
Internationalization Lead, W3C

Track 1: An Introduction to Writing Systems & Unicode
The tutorial will provide you with a good understanding of the many unique characteristics of non-Latin writing systems, and illustrate the problems involved in implementing such scripts in products. It does not provide detailed coding advice, but does provide the essential background information you need to understand the fundamental issues related to Unicode deployment, across a wide range of scripts. It has also proved to be an excellent orientation for newcomers to the conference, providing the background needed to assist understanding of the other talks! The tutorial goes beyond encoding issues to discuss characteristics related to input of ideographs, combining characters, context-dependent shape variation, text direction, vowel signs, ligatures, punctuation, wrapping and editing, font issues, sorting and indexing, keyboards, and more. The concepts are introduced through the use of examples from Chinese, Japanese, Korean, Arabic, Hebrew, Thai, Hindi/Tamil, Russian and Greek. While the tutorial is perfectly accessible to beginners, it has also attracted very good reviews from people at an intermediate and advanced level, due to the breadth of scripts discussed. No prior knowledge is needed.

Presenter:
Addison Phillips
Globalization Architect
Lab126 (Amazon)

Track 2: Internationalization: An Introduction, Part I: Characters and Character Encodings
What is internationalization? What do developers, product managers, or quality engineers need to know about it? How does a software development organization incorporate internationalization into the design, implementation, and delivery of an application? This tutorial track provides an introduction to the topics of internationalization, localization and globalization. Attendees will understand the overall concepts and approach necessary to analyze a product for internationalization issues, develop a design or approach, and deliver a global-ready solution. The focus is on architectural approaches and general concepts, but will include specific examples and exercises. Part I focuses on characters, character encodings, and the basics of Unicode.

Presenter:
Elizabeth Pyatt
Instructional Designer
Penn State

Track 3: Building a Custom Keyboard Layout for the Mac with Ukulele and XML
Building custom keyboards can be a useful timesaver if you work with an unusual range of characters across a large number of documents. The tutorial will describe how to create a custom keyboard layout on the Mac OS X platform using the freeware Ukelele tool from SIL plus modifications to the XML file. Although the main example will a keyboard built for symbolic logic characters, the tutorial will cover how to create keyboards for many foreign languages.

10:45-12:30 MORNING TUTORIALS

Track 1 - An Introduction to Writing Systems & Unicode (Cont'd.)

Presenter:
Addison Phillips
Globalization Architect
Lab126 (Amazon)

Track 2: Internationalization: An Introduction, Part II: Writing Global-Ready Code
Part II focuses on preparing for the localization (translation) of user interfaces; making applications “locale-aware”, including format and display differences; as well as approaches to delivering multi-lingual and multi-locale software or content.

Presenter:
Thomas Milo
President
DecoType

Track 3: Arabic Script: Structure, Geographic and Regional Classification
A new tutorial about Arabic script (including Arabic script for dummies, structural analysis, typology, stylistic geography, technical and aesthetic aspects, language-dependant preferences within calligraphic styles, and extra attention for orthographies East of Iraq), against the background of the development of a brand-new Nastaliq typeface that covers the Unicode for all languages that require this Persian-derived style.

13:30-15:30 AFTERNOON TUTORIALS

Presenters:
Craig Cummings
Mike McKenna
Internationalization Architects
Yahoo! Inc.

Track 1 - Unicode - A Grand Tour
This tutorial will cover the next level of detail of what Unicode is, and how it is used in the real world. The modules of the tutorial will cover: The Unicode standard - what are the "Guiding Lights", or design principles behind Unicode? A tour of Unicode's structure, encoding forms, behavior, technical reports, database, and how to use the Unicode Standard. Implementation according to Unicode - a walk through the details of attributes, compatibility, non-spacing characters, directionality, normalization, graphemes, complex scripts, surrogates, collation, regular expressions and other aspects according to the Unicode Standard and associated Technical Reports. Unicode and the Real World - an overview of International Components for Unicode (ICU) and implementations supporting Unicode in web servers, application servers, browsers, C/C++, Java, PHP, SQL, and various operating systems. On-going programs – how Unicode is evolving to support more minority scripts, languages, and help solve linguistic processing issues.

Presenter:
Tex Texin
Xen Master
XenCraft

Track 2 - Web Internationalization - Standards and Best Practices
This tutorial is an introduction to internationalization on the World Wide Web. The audience will learn about the standards that provide for global interoperability and come away with an understanding of how to work with multilingual data on the Web. Character representation and the Unicode-based Reference Processing Model are described in detail. HTML, XHTML, XML (eXtensible Markup Language; for general markup), and CSS (Cascading Style Sheets; for styling information) are given particular emphasis. The tutorial addresses language identification and selection, character encoding models and negotiation, text presentation features, and more. The design and implementation of multilingual Web sites and localization considerations are also introduced.

Presenter:
Jim DeLaHunt
Principal
Jim DeLaHunt & Associates

Track 3 - Building Multilingual Websites in Joomla [Drupal]
A practical look at the language and locale capabilities of Joomla! and Drupal, two leading free software content management systems (CMSs). They let you build more powerful, more international websites faster. We look at: their core services for internationalization and locale support; localization of UI and content; and localization support in some leading modules. You will leave with specific tips for building your own site. We don't assume Joomla or Drupal experience, but do include material for advanced practioners. A good tutorial for web site product managers, for web designers and developers, and for managers of international web site teams.

15:45-17:45 AFTERNOON TUTORIALS

Track 1 - Unicode - A Grand Tour (Cont'd.)

Presenter:
Richard Ishida
Internationalization Lead
W3C

Track 2 - Creating XHTML/HTML Pages with Right-to-Left Scripts
This short tutorial explains how to go about creating XHTML and HTML pages containing text written in the Arabic or Hebrew scripts. The tutorial examines how best to achieve the correct effect for these bi-directional scripts using appropriate markup, CSS properties and Unicode code points or entities. It covers the basics, and goes beyond to provide recommended techniques for some of the tricky situations that even native speakers can struggle with. The tutorial assumes a basic familiarity with the bi-directional characteristics of Arabic and Hebrew, as well as a basic knowledge of HTML and CSS.

Presenter:
Behdad Esfahbod
Software Developer
Red Hat/GNOME

Track 3 - Free Software Stack for Unicode Text Rendering
The Free Software world has a lot to offer when it comes to building a stack up from the grounds. Be it building an ARM-based Linux mobile platform or cross-platform text rendering to rendering downloadable CFF fonts on Windows, the Free Software stack provides all the bits and pieces one needs to assemble a high quality OpenType-based Unicode text rendering pipeline with great flexibility. In this tutorial we will go over the building blocks involved and how to put them together.

18:00-19:00 - Welcome Reception hosted by Adobe Systems

Thursday, October 15, 2009

09:00-09:15

WELCOME & OPENING REMARKS

09:15-10:00

Nicholas Ostler
Chairman
Foundation for Endangered Languages

KEYNOTE Presentation: The Alphabetic Principle and its Enemies
The alphabetic principle for writing seems brilliantly simple, and its implementation, often subverting other options, has often caused explosive growths in literacy, with important historical consequences for cultural survival. Its great advantages are economy of effort in the learner, and ready application to new languages. However, it has drawbacks as to speed for the initiated user, and also (by being essentially mechanical and phonetic) in representing many of the cultural overtones which people like their written language to have. There is, too, a certain resistance to the role of art in writing. But as alphabetic traditions age, becoming less purely alphabetic, these disadvantages can be reduced. New structures may emerge, meaningful patterns that leave alphabets far behind. Alphabetic scripts have more recently revealed new aspects, defining a convenient order to index anything, inspiring the phonemic principle of structural linguistics, and later mapping more easily than other systems onto digital systems, and hence a whole new set of functions for written language. But the alphabet remains a rather arbitrary means of representing meanings, since its icons are parasitic on the particular sounds of particular words in particular languages, a long way from thoughts.

10:00-20:00 - EXHIBIT AREA OPEN

10:30-11:20 SESSION 1

Presenter:
Kirti Velankar
Senior Software Engineer
Yahoo! Inc.

Track 1 - Internationalization with PHP
PHP is one of the most prominent and popular platforms for modern Web development This updated session discusses PHP from the perspective of internationalization, what some of the challenges in PHP are, the features available in PHP 5, and the promise of Unicode in PHP 6.
This session also includes examples and usage in practical scenarios. You will learn how to effectively build applications for multiple languages and cultures using PHP with some of the new internationalization features such as locales, sorting, resource bundles, as well as date, number and message formatting.

Presenter:
Ken Lunde
Senior Computer Scientist
Adobe

Track 2 - Designing & Developing Pan-CJK Fonts for Today
Designing and developing Pan-CJK fonts, meaning fonts whose CJK Unified Ideographs can serve more than a single CJK locale, region, or culture, is both challenging and timeconsuming. But, like most things that require effort, there are great rewards: smaller overall font footprint, design consistency across locales, and so on. In developing such fonts, there are challenges related to the actual design of the glyphs, which transcend any font format concerns. This presentation pinpoints specific design and implementation problems that developers of such fonts will face, and then details workable solutions. A prototype Pan- CJK font will demonstrated during the presentation.

Presenter:
Mark Davis
Sr. Internationalization Architect
Google Inc.

Track 3 - Unicode Update: Unicode 5.2 and CLDR 1.7
The 5.2 version of Unicode (Fall 09) adds many new characters, new properties, and fixes to existing properties, and is being issued as a complete online book. CLDR 1.7 (Spring 09) contains over 21% more locale data than the previous release, with over 40,000 new or modified data items from over 140 different contributors, including Adobe, Apple, Google, IBM, and Sun, plus official representatives from a number of countries.
This presentation, from the president and co-founder of the Unicode consortium, covers the new features of both standards, examples of the impact on companies such as Google, and future directions for these and other globalization standards -- the new emoji characters, international domain names, Unicode security, and others.

11:30-12:20 SESSION 2

Presenter:
Martin Dürst
Aoyama Gakuin University

Track 1 - Internationalization in Ruby 1.9
Ruby is a purely object-oriented scripting language which is easy to learn for beginners and highly appreciated by experts for its productivity and depth. Internationalization of Ruby made a big leap forwards when this January, Ruby 1.9.1, the first stable release of the Ruby 1.9 series, was released. While previous versions of Ruby mostly treated text data as byte sequences, strings in Ruby 1.9 are sequences of characters. Because Ruby tags each string with encoding information internally, different applications can choose different internationalization models. The presentation will give a short overview of Ruby as a programming language, and introduce the new internationalization features in detail. We will be concentrating on how to use Ruby with Unicode, which in Ruby's case means UTF-8. We will also discuss internationalization support in Ruby on Rails, the popular Web application framework written in Ruby.

Presenter:
Kamal Mansour
Manager of Non-Latin Products
Monotype Imaging

Track 2 - Unicode & Fonts: a status report
The adoption of Unicode as the universal character code standard has profoundly changed the computing landscape. We now expect to be able to exchange multilingual text documents across platforms and software applications. Since its inception, Unicode has cautiously distanced itself from the process of displaying glyphs, delegating it to an external “rendering layer” that includes fonts. Alongside Unicode, the OpenType Standard has enabled new levels of sophistication in fonts. However, one is often disappointed by a particular font doesn’t work as it should. We will give a brief overview of what works today and what we can expect in the future.

Presenters:
Deborah Anderson
Project Leader, Script Encoding Initiative, Department of Linguistics, UC Berkeley
Richard Cook
Post-Doctoral Researcher, Dept. of Linguistics UC Berkeley
Charles Riley
Catalog Librarian for African Languages Yale University
Anshuman Pandey
C.Phil. History University of Michigan

Track 3 - Patching Holes in the Unicode Pipeline: A Status Report on the Unencoded Scripts of Asia and Africa
In 2002, 96 scripts listed on the Unicode Pipeline were unencoded. Today, the number is considerably smaller. Currently about 25 scripts from Asia and Africa remain unencoded, but they present particular challenges: many are not well-known and will involve considerable research to acquire materials and to track down experts. This session will be made up of 3 speakers who have worked on South Asian and African script proposals. They will discuss the work that remains to be done and highlight specific issues for implementers.

13:30-14:20 SESSION 3

Presenter:
Norbert Lindenberg
Internationalization Architect
Yahoo! Inc.

Track 1 - Internationalization for JavaScript Applications
JavaScript, as defined by the EcmaScript standard and implemented in browsers, is a rather weak platform for internationalized web applications. Several toolkits have attempted to fill the gap in different ways, ranging from reliance on existing server-side internationalization libraries to implementing the functionality in JavaScript itself. This presentation surveys the landscape and compares the different solutions.

Presenter:
Ken Lunde
Senior Computer Scientist
Adobe

Track 2 - The Design & Development of Fully Proportional Japanese Fonts
Japanese fonts have traditionally been designed on the principle that each glyph occupies a fixed design space. Some fonts have overcome this principle by providing alternate metrics, which really amount to pseudo proportional metrics. It is possible to develop Japanese fonts whereby each glyph has proportional metrics by default, in both horizontal and vertical writing directions. In addition to the obvious design challenges, there are also several technical hurdles related to implementing the typeface design as an OpenType font. This presentation details the unique design aspects of Kazuraki, a fully-proportional Japanese font, along with details about its OpenType implementation.

Presenter:
Martin Dürst
Aoyama Gakuin University

Track 3 - Update on Internationalized Domain Names and Internationalized Resource Identifiers
In domain names such as www.unicode.org, only a limited number of characters are allowed. This limitation also applies to Uniform Resource Identifiers (URIs) such as http://www.unicode.org. Internationalized Domain Names (IDNs) and Internationalized Resource Identifiers (IRIs) changed this a few years ago, both allowing a wide range of
characters from the Unicode repertoire. The specifications underlying these technologies are currently facing an overhaul, major for IDNs and minor for IRIs. The long-overdue and now imminent introduction of the first international top-level domain names will mean that the importance of IDNs and IRIs will significantly increase in the near future.
The presentation will give a general overview of IDNs and IRIs and discuss the current revisions of the specifications in detail. For IDNs, the set of allowed characters is defined using an inclusion-based model rather than the earlier exclusion-based model. Fixed tables are replaced by a property-based selection process to avoid fixing the specification to a single version of Unicode. The mapping step (dealing with casing and normalization, among else) is moved out of the core libraries and closer to the user to allow adaptions for special cases and reduce user surprises. The IRI specification is being extended with descriptions of widely used variants for handling characters strictly speaking not allowed in IRIs. Both specifications are affected by bug fixes to bidirectionality restrictions.

14:30-15:20 SESSION 4

Presenter:
Umesh Nair
Software Engineer
Google Inc.

Track 1 - Implementing International Calendars in JavaScript
Conversion routines between the Gregorian calendar and non-Gregorian calendars involve complex floating point computations, large lookup tables and calendar-specific computations. Floating point operations impact performance and accuracy, while lookup tables impact memory footprint and download time. Calendar-specific computations require special algorithms and data structures. Implementing such algorithms efficiently with compact data structures is essential for the successful deployment of online calendars for the international audience. This presentation discusses several such techniques for calendrical calculations in client-side JavaScript. The techniques described here are applicable to a number of

other areas in internationalization as well as general software usage with JavaScript.

Presenter:
Thomas Milo
President
DecoType

Track 2 - The Unicode-based Koran: a Conflict Between Calligraphic Tradition and Computer Typography
A technical talk about the practical problems encountered in the project to produce a Unicode-based Koran on the behest of the Omani Ministry of Awqaf and Religious Affairs. The focus is on the discrepancies discovered between the age-old calligraphic tradition and the 1924 revision of the Koran. The pivotal issues will be identified and explained. A workable solution will be presented.

Presenters:
Mark Davis
Sr. Internationalization Architect
Google Inc.
Addison Phillips
Globalization Architect
Lab126 (Amazon)

Track 3 - Language Identification and Usage
In 2006, the IETF issued an updated version of BCP 47 "Tags for Identifying Languages", which updated the way languages are identified in most computer programs and protocols. The latest version of BCP 47 (2009) incorporates over 7,000 new languages and many other improvements. This presentation, from the authors of the updated and previous RFCs, covers:

- the format of language tags and the language subtag registry
- the matching algorithms for comparing language tags to user preferences
- plus distance-based algorithms
- the new features in BCP 47 and their impact on developers and how BCP 47 is being used in:
- Unicode locales (CLDR)
- prominent open-source libraries such as ICU
- companies such as Google and Amazon

16:00-16:50 SESSION 5

Presenters:
Steven Loomis
Software Engineer
IBM
Markus Scherer
Unicode Software Engineer
Gooale Inc.

Track 1 - What's New with ICU
The International Components for Unicode library, or ICU, provides a full range of services for Unicode enablement, and is the globalization foundation used by many software packages and operating systems. Freely available as open-source, it provides cross-platform C, C++ and Java APIs, with a thread-safe programming model. This presentation will provide a brief overview of ICU, with emphasis on the current status of ICU (4.2), including the latest support for Unicode 5.1 and CLDR 1.7, and an update on ICU’s planned direction for 4.4 and future releases.

Presenters:
Michael Manca
Project Manager and Solution
Quality Analyst
IT Flex Services
Intel Corporation
Tomas Galicia
Solutions Quality Analyst
IT Flex Service
Intel Corporation
Loic Dufresne de Virel
Localization Strategist
IT Flex Services
Intel Corporation

Track 2 - A Systematic Approach to I18N Testing
Building on last year's presentation "We're World-Ready, What Does This Really Mean?", Intel's localization experts will present and discuss the steps they follow, the tools they use, and their overall I18N testing philosophy. They will explain in details how they proceed when working with development teams to ensure applications are properly internationalized before they're released or localized. Based on recent I18N testing efforts conducted by Intel, this interactive session will provide a solid framework of reference for I18N testing, as well as valuable pointers that can be easily and directly applied to your own localization projects or reused within your organization.

Presenter:
Toshiya Suzuki
Research Assistant
Hiroshima University

Track 3 - Investigation of Opaque Glyphs Synthesized from Old Hanzi
After the long efforts during 7 years, finally ISO/IEC 10646:2008 have included CJK Unified Ideographs Extension C. It has 366 glyphs taken from "Index to Collections of the Inscriptions in Yin-Zhou period" (I2CIYZ) proposed by PRC, and more glyphs are scheduled for future Extension E project. They are suspected to be the glyphs

invented only for the specification of Old Hanzi. In this report, the source is investigated and compared with existing dictionaries for Bronze scripts. The requirements of some glyph shapes are questionable, the expected procedure to standardize these opaque glyphs is discussed.

17:00-17:50 SESSION 6

Presenter:
Behdad Esfahbod
Software Developer
Red Hat/GNOME

Track 1 - HarfBuzz, the Free and Open OpenType Shaping Engine
In this session we will introduce HarfBuzz, the unified Free Software and Open Source, OpenType-based, text shaping engine. We will discuss design considerations, technical decisions made, and performance and other features that make HarfBuzz an attractive alternative to the existing OpenType engines. HarfBuzz is already being used by both GNOME and KDE desktop environments and is at the heart of the GTK+ and Qt desktop and mobile platforms, with others planning to use it in the coming months, including Mozilla Firefox, OpenOffice.org, and ICU Layout.

Presenters:
Andrew Swerdlow
Internationalization Tech Program Mng
Google Inc.
Manish Bhargava
Google Inc.
Jens Riegelsberger
Google Inc.
Laura Cuozzo
Google Inc.

Track 2 - Google Internationalization Quality Control Framework
There are many obstacles to a great international user experience. There is a range of issues that cut across organizational boundaries, such as localization, internationalization, visual design, interaction design, business analysis, usability analysis, and market research. Against this backdrop we at Google started experimenting with a standardized review framework that relies on a global network of external evaluators. These evaluators live in market and thus are familiar with local standards and practices. This framework allows us to identify themes that may point to requirements that are common across multiple regions aiding in prioritizing features or giving resources to projects.

Presenter:
Murry Sargent III
Partner Software Design Engineer
Microsoft

Track 3 - Math Editing and Display in Microsoft Office
Math editing is described that uses math context menus, a math ribbon, keyboard navigation, and formula autobuildup in Microsoft Office 2010. The math typography is similar to TeX's, the input methods are state of the art, the math character set is Unicode's, and the environment is Office's, which comes with the many features one expects from a leading office suite. Demonstrations will be given using Office 2010.

18:00-20:00 - IUC33 CONFERENCE RECEPTION *(IN EXHIBIT AREA)*

Friday, October 16, 2009

09:00-09:50 *SESSION 7*

Presenter:
Douglas Davidson
Software Engineer
Apple, Inc.

Track 1 - International Features of Mac OS X Snow Leopard
From its inception, Mac OS X has been designed with top-to-bottom international and multilingual support. The latest version, Mac OS X 10.6 Snow Leopard, expands on that with new bidirectional input support, multilingual spellchecking, and many other new features. This session covers the international capabilities of Mac OS X from both a user and a developer perspective, with a particular emphasis on new features in Snow Leopard. Topics covered include localization, locale data, text input, text display, proofing tools, and user customization.

Presenter:
Brent Ramerth
Software Engineer
Apple, Inc.

Track 2 - International Features of iPhone OS
The iPhone OS platform starts with the internationalization architecture fundamental to Mac OS X, and adds a unique virtual keyboard and text input system that handles a wide array of languages. This session covers the international capabilities of the platform from both a user and a developer perspective, with particular attention to iPhone-specific features. Topics covered include localization, text display, and text input.

Presenter:
Elizabeth Pyatt
Instructional Designer
Penn State

Track 3 - Practical "Unicode Logic" for Online Tech Courses
This session describes some of the challenges and workarounds for implementing Unicode content in two online courses in symbolic logic and thermodynamics.

Topics include development utilities, templates and guidance for students, issues with multiple applications and font selection across platforms. The presentation will also discuss some differences between implementing Unicode for math courses and Unicode for foreign language courses.

10:00-10:50 SESSION 8

Presenter:
Elizabeth Pyatt
Instructional Designer
Penn State

Track 1 - Windows 7: Writing World-Ready Applications
This session centers on the new globalization features for Windows 7, including sorting and string comparison, locale support, and coverage for new languages, with an eye to helping developers extend their applications to a global user base. In addition to introducing the Extended Linguistic Services API, this session will also cover the Multilingual User Interface (MUI) resource technology available in Windows 7. This session will provide an end-to-end look at how to make your application world-ready so that you can easily take your application worldwide and extend your customer base into
new language markets.

Presenters:
Markus Scherer
Unicode Software Engineer
Google Inc.
Katsuhiko Momoi
Staff Test Engineer & I18n Consultant
Google Inc.
Mark Davis
Sr. Internationalization Architect
Google Inc.

Track 2 - Emoji in Unicode: Cell Phones Meet the Internet
Emoji" symbols or "picture characters" are used in email by more than 80 million Japanese cell phone users. They are treated as characters, via vendor-specific extensions of the Japanese character sets. Other email providers have to be able to exchange emails with the Japanese cell phone companies without losing or corrupting data. Most email providers use Unicode, requiring conversion of mail data to/from Unicode. Unicode Private Use characters are used for this purpose. However, they do not provide for reliable public interchange. For a permanent solution, the Unicode Consortium has approved the addition of the Emoji symbols to Unicode 6.0, and is working with ISO to ensure inclusion in the corresponding version of ISO 10646. This paper presents the state and progress of the Unicode encoding proposal with an overview of the Emoji symbols.

Presenter:
Adam Asnes
President
Lingoport, Inc.

Track 3 - Creating an I18n Project Plan
Many initial internationalization scoping efforts focus on creating findings documents. But often the real trick is gathering accurate metrics and turning them into realistic, budge table and actionable project plans. In this presentation we will demonstrate how we assess source code and architecture, and then review a detailed project plan and how we arrived at tasks, durations and staffing.

11:10-12:00 SESSION 9

Presenter:
Mihai Nita
Globalization Architect
Adobe Systems, Inc.

Track 1 - Accessing Globalization Services on Multiple Operating Systems
This presentation will cover the experience gained by implementing a cross platform C library that makes use of the operating system dependent services for the following language and region specific functionality. In contrast to ICU which carries its own set of locale data, this solution provides a cross platform set of APIs but uses the facilities provided by the operating system. This presentation will explore the pros and cons of such an approach, trade-offs, implementation issues, major traps, and some of the surprises we encountered.

Presenters:
Loic Dufresne de Virel
Localization Strategist
IT Flex Services
Intel Corporation
Michael Kuperstein
Senior Localization Engineer
IT Flex Services
Intel Corporation
Margie Foster
Localization Project Manager
Moblin Project
Intel Corporation

Track 2 - Taking Moblin to the World
When the Moblin project asked for our help to localize their application, our initial reaction was enthusiastic! "Finally a cool open-source project to work on", we thought! After getting back to our senses, we realized that localizing Moblin (Moblin stands for Mobile Linux) was not our typical localization project... Far from it! In this session, we will review the thought process we followed to define and limit the scope of this significant undertaking, give an update on the current status of this on-going project, explain how we addressed the first major challenges of this amazing journey, and provide an overview of the first-ever attempt at community-based translation by Intel's localization team.

Presenter:
Cindy Conlin
Senior Engineer
The Church of Jesus Christ of Latter-Day Saints

Track 3 - Building a Global Names System: A Case Study
This case study discusses our experience building a global names application containing records for all members of the LDS Church worldwide. We'll discuss the interesting challenges and requirements we face, such as building a data structure flexible enough to accommodate names from multiple cultures simultaneously. We'll talk about using ICU's transliteration functionality to generate romanizations of non-Latin names, and about our experience supporting private-use characters in Chinese names. We'll also discuss how we've created a user interface that allows users from multiple locales to work with data that originated in many other locales

13:00-13:50 SESSION 10

Presenter:
Sumit Sarkar
i18n Product Specialist
DataDirect Technologies

Track 1 - Internationalization in Database Drivers for C/C++/Java/.NET Applications
Everything you want to know about i18n and database drivers across C/C++/Java/.NET programming languages. Discussion starts by asking what Unicode support encompasses at the Database Access API level, and what components affect Unicode Support. Take a closer look under the covers at the low level data access across major RDBMS including DB2, SQL Server, Oracle, and Sybase. This includes identifying who is doing the conversions at each component of the data access application layer. To summarize and apply the learned concepts, host will answer key questions about your globalized application's data access: Why should conversions be avoided when possible; and what high level features of a database driver are recommended?

Moderators:
Steven Loomis
Software Engineer
IBM
Mark Davis
Sr. Internationalization Architect
Google Inc.

Track 2 - Deploying the Common Locale Data Repository (CLDR)
The Common Locale Data Repository is a project for the exchange of language and locale information used in application development, and to gather, store, and make such data publicly available. By pooling resources, the time and expense of collecting good data is minimized, and language groups have an avenue to get their data into implementations. This session will discuss implementation of CLDR, the latest project status, and how the process is being improved to produce higher-quality data. Ample time will be given for comments and questions from the audience.

Presenter:
Chris Weber
Casaba Security

Track 3 - Unicode Transformations and Security Vulnerabilities
Web-applications are being exploited every day as attackers find new vectors for performing cross-site scripting attacks. This talk will cover ways which latent character and string handling can transform clever inputs into malicious outputs. Many application frameworks such as .NET and ICU enable these behaviors without the developer's knowledge. String transformations through best-fit mappings, casing operations, normalization, over-consumption and other means will be discussed, with inputs useful for testing. A testing tool is also planned for release. The current state of visual spoofing attacks will also be discussed. Phishing attacks are prevalent on the Web, and well-designed URL's can increase an attack's chance of success. It's eye-opening to see demonstrations of just how vulnerable modern Web browsers still are to many forms of visual spoofing attacks

14:00-14:50 SESSION 11

Presenter:
Su Liu
AIX Globalization Architect
IBM

Track 1 - Unicode Technology and Globalization Support in IBM UNIX, AIX
AIX, an IBM UNIX, supports more than 60 languages and about 250 locales. Unicode is a key technology to support globalization features to meet different national language requirements. This presentation discusses Unicode impacts on globalization strategy and mechanism in UNIX operating system level. It focuses on how Unicode technologies are used to simplify globalization configurations first. Then, topics are covered on Unicode impacts on system performance, locale data test, and national language support procedure. Examples are given to explain show Unicode support on complex texts, CJK input methods, Unicode conversions, and automated tests. A further looking into Unicode highlights customization subjects on

user-defined locale settings and user-defined Unicode conversion tables. Finally, issues in implementations, market requirements and solutions for future Unicode support in UNIX are assessed.

Presenters:
Benedicto Franco Jr.
Software Engineer
Yahoo! Inc.
Marco Aurelio Carvalho
Senior Software Engineer
Yahoo! Inc.

Track 2 - CLDR on the Cloud
The value of CLDR (Common Locale Data Repository) for global applications is undeniable. But how do you update time zone and daylight saving rules, or a new currency, or geo-political changes that might be relevant for the application without taking the inherent risks and costs of a release deployment process? In this presentation, we are going to talk about a solution that exposes CLDR as a service and how CLDR on the Cloud can be used to help create robust internationalized JavaScript and Ajax applications fed by CLDR data published in JSON format ubiquitously.

Presenter:
Tex Texin
Xen Master
XenCraft

Track 3 - My Unicode Disk Storage Went into the Circular File
This session will present some of the difficulties of providing a common international interface to file services on different operating systems. Although Unicode supports all the necessary characters, identifying the set of characters that are legitimate on any OS can be difficult, and rules for case-insensitivity, normalization, etc. vary, and may even vary by user. The presentation will describe the problem space. It may offer possible solutions.

15:10 - 16:00 SESSION 12

Presenters:
Ryan Cavalcante
Software Development Engineer
Microsoft
James Lyle
Program Manager
Microsoft

Track 1 - Extended Linguistic Services in Windows 7
In this presentation we will discuss the Extended Linguistic Services (ELS) platform, new to Windows 7, which provides diverse linguistic services to developers through a common API. We will discuss the linguistic services now available to developers through the ELS platform in Windows 7—Language Detection, Script Detection, and various Transliteration services—as well as the future vision for the platform.

Presenter:
Adil Allawi
Technical Director
Diwan Software Limited

Track 2 - Mashing-up Bi-Di
Mash-ups is a relatively new fashionable word on the Web - taking bits of other web sites to build up your own web page. It is not new or special - any search engine showing a snippet of a web site that it has found is a form of mash-up. Integrating a news or microblogging feed is another. And it seems that every company and their mother has its own mash-up API. But what happens when you have an Arabic web-site integrate content that may be Arabic or English or both? The Unicode Bi-Di Algorithm can render text and numbers unreadable. URL's may become unusable or, in the worst case, direct to fraudulent sites. It can be hard to predict how to mark-up the integrated content for the right result. This presentation will cover real world issues and attempt to suggest practical solutions.

Presenter:
Jim DeLaHunt
Principal
Jim DeLaHunt & Associates

Track 3 - Twanguages of the World: a Language Census of Twitter
What "twanguage" do you "tweet"? Twitter, the buzzing conversation of brief web and SMS messages, exploded into wide use in 2009. But just how wide? To how many countries has it spread? And into which languages? We aimed to find out. Our "Twanguage" project is a language census on a sample of Twitter's global traffic. Come hear our findings. Which are the top languages? Are #hashtags localized? How does language correlate with location? And which Unicode character is the most rarely used? Accessible to everyone, this talk is especially interesting to students of social media and of quantative language analysis.

16:10 - 17:00 SESSION 13

Presenters:
Frank Yung-Fong Tang
Sr. Software Engineer
Google Inc.
Wenchao Tong
Software Engineer
Google Inc.

Track 1 - Google APIs for Text Input and Translation
In this talk, we introduce several Google public APIs to empower web developer build more powerful internationalized web site, including, but not limited to:
- Use Google AJAX Language API and element API to perform Machine Translation

- Use Google AJAX Language API and element API to empower user to input text of different language by transliteration
- Use Maps in Google Chart API and Geomap in Google Visualization API to represent information divided by geographical distribution

For each of these topics, we will first introduce the issues, following by the brief description of the API, and demonstrate with some real Google or non Google products which utilize these APIs. Short sample codes will also be walk though.

Presenter:

Roozbeh Pournader

Internationalization Specialist

HighTech Passport

Track 2 - Bidirectionalization: Demystifying Bidi Enabling

Bidirectionalization, or enabling software to be usable to people who write in bidirectional languages like Arabic and Hebrew, has sometimes been discarded as a superfluous and strenuous endeavor. This presentation will explain why bidi enabling is a must for every application and website intended for bidirectional users of the Middle East, as well as for other parts of Asia and Africa. It will also include suggestions on how to plan for, design, code, and test the bidirectionalization of such applications and sites. Last but not least, it will cover common internationalization requirements for the Middle East, including alternative calendars, local digits, and geopolitical sensitivities. The intended audience of this presentation are developers, software architects, and managers planning to bidirectionalize their software or add support for other requirements of the bidirectional language markets

---

Presenter:

Ilya Shtein

IT Architect

Metavante

Track 3 - Banking in the Cloud: Challenges of Internationalizing Banking Software (Case Study)

Based on the experience of building the Metavante Global Banking platform, we will discuss the challenges of internationalization in a distributed, service-oriented, heterogeneous banking environment.

Internationalization in the banking industry presents a number of challenges, such as the large number of legacy applications that do not share the same terminology and the need for further terminology customization on multiple hierarchy levels, as well as transactions spanning multiple locales and time zones.

We will talk about the applicability of Unicode and Unicode standards in different architecture layers, using W3C-i18n recommendations, and discuss the effect the listed challenges have on internationalization decisions.

四、 演講者簡介

Presenters' Biographies

Keynote Presenter:

Mr. Nicholas Ostler

Chairman, Foundation for Endangered Languages

Nicholas Ostler holds an MA in classics, philosophy and economics from Oxford, and a PhD in linguistics from MIT. His first job was teaching in Japan, later consulting on machine translation for Fujitsu. Returning to England, he worked in IT research during the 1980s and '90s, especially with the UK government, and the European Union. He has been Chairman of the Foundation for Endangered Languages (www.ogmios.org) since its inception in 1996. He also edited its newsletter Ogmios until 2006. Within descriptive linguistics, his main research field has been the grammar of the (extinct) Chibcha language of Colombia. He has served on the board of the British National Corpus, the LSA's Committee for Endangered Languages, and on the editorial board of the International Journal of American Linguistics. As a writer, his book "Empires of the Word: a language history of the world" (HarperCollins, 2005) traced the histories of the large literate languages, from Sumerian to English, considering the factors that make for large-scale expansion. Later, "Ad Infinitum: a biography of Latin" (Walker & Co., 2007) considered the attitudes that have accompanied the Latin language throughout its 2,500 year recorded history. He is now at work on a book about the prospects of English as a global lingua franca, in the light of the competition, past and present. This is due for publication in 2010.

Presenters:

Mr. Adil Allawi

Technical Director, Diwan Software Limited

Adil Allawi has been working in the field of multilingual computing for the past 20 years. He started with writing bilingual software for one of the first implementations of Arabic on a personal computer and has continued into the fields of word-processing, desktop publishing, and text rendering on small devices. Adil largest project was rewriting a high-end DTP application first to handle both Arabic and English and later to work with all international languages. Adil is the Technical Director and lead engineer of Diwan Software Limited.

Dr. Deborah Anderson

Project Leader, Script Encoding Initiative, Department of Linguistics, UC Berkeley

Deborah Anderson is a researcher in the Department of Linguistics at UC Berkeley. She runs the Script Encoding Initiative at UC Berkeley, and is also a Unicode Technical Director.

Mr. Adam Asnes

President, Lingoport, Inc.

Adam Asnes founded Lingoport in 2001 after seeing firsthand that the niche for software globalization engineering products and services was underserved in the localization industry. As Lingoport's President and CEO, he focuses on sales and marketing alliances while maintaining oversight of the company's internationalization services engineering and Globalyzer product development. Adam is a frequent speaker and columnist on globalization technology as it affects businesses expanding their worldwide reach. For creative inspiration and fun, Adam enjoys cycling and Colorado's Rocky Mountains.

Manish Bhargava

Google Inc.

Manish Bhargava is Product Manager for Internationalization(i18n) efforts at Google. He helps drive Google's 40-language initiative. He graduated from Stanford University with 2 Masters – Computer Science, Aeronautics & Astronautics. He has Bachelors from IIT Bombay, India.

Marco Aurelio Carvalho

Senior Software Engineer, Yahoo! Inc.

Marco Aurelio Carvalho is a senior software engineer at Yahoo Inc. He works on the globalization group where he participates on the design and development of a company wide localization tool. He worked on several different projects and libraries, focusing on internationalization and globalization solutions. Marco worked previously on localizing products at Yahoo! Brazil for Latin America markets.

Mr. Ryan Cavalcante

Software Development Engineer, Microsoft

Ryan Cavalcante has been a Software Development Engineer in Test for the Globalization Services workgroup at Microsoft for the past 6 years, involved in such globalization technologies as collation, encoding, normalization, IDN, and most recently the Extended Linguistic Services platform.

Richard Cook

Post-Doctoral Researcher, Dept. of Linguistics, UC Berkeley

Dr. Richard Cook is a researcher in linguistics at UC Berkeley and at Berkeley's International

Computer Science Institute. He is an editor of Unicode 5.0, contributor to the Script Encoding Initiative, and author of the recent book Classical Chinese Combinatorics. He is co-author of the *Character Description Language (CDL) Specification*.

Ms. Cindy Conlin

Senior Engineer, The Church of Jesus Christ of Latter-Day Saints

Cindy Conlin is a software engineer at The Church of Jesus Christ of Latter-Day Saints and is the Church's liaison to the Unicode Consortium. Cindy is a former Oracle employee. She graduated from Brigham Young University.

Mr. Craig Cummings

Internationalization Architect, Yahoo! Inc.

Craig Cummings is an Internationalization Architect at Yahoo! Inc., where he helps set internal standards, participates with the Unicode Consortium, and helps drive corporate technical strategy for internationalization. In past lives, Craig was a key member of Oracle's Applications Internationalization team. Since J2SE v1.3, Craig has worked closely with Sun's internationalization team to help shape some of the pluggable locale, resource bundle, font, and supplementary character support in Java.

Laura Cuozzo

Google Inc.

Laura Cuozzo has been a senior user experience researcher at Google for 2.5 years. Laura was the lead UR on Adwards.

Mr. Douglas Davidson

Software Engineer, Apple, Inc.

Douglas Davidson has worked on Mac OS X and its predecessors at NeXT and Apple since 1996, primarily working on the Mac OS X subsystems dealing with localization and text. He currently leads a group at Apple dealing with natural language data and processing.

Mark Davis

Sr. Internationalization Architect, Google Inc.

Dr. Mark Davis co-founded the Unicode project and has been the president of the Unicode Consortium since its incorporation in 1991. He is one of the key technical contributors to the Unicode specifications.Mark founded and was responsible for the overall architecture of ICU (the premier Unicode software internationalization library), and architected the core of the Java internationalization classes. He also founded and is the chair of the Unicode CLDR project, and is

a co-author of BCP 47 "Tags for Identifying Languages" (RFC 4646 and RFC 4646), used for identifying languages in all XML and HTML documents. Since the start of 2006, Mark has been working on software internationalization at Google, focusing on effective and secure use of Unicode (especially in the index and search pipeline), the software internationalization libraries (including ICU), and stable international identifiers.

Mr. Jim DeLaHunt

Principal, Jim DeLaHunt & Associates

Jim DeLaHunt is a Vancouver, Canada-based software engineer and consultant in multilingual websites. He helps business reach culturally diverse markets through globalize technology products. He is an active contributor to the Joomla and Drupal ecosystems, and is a regular Unicode conference participant. He also writes, teaches, and develops software. Earlier, he worked 16 years in Silicon Valley for Adobe Systems. Jim is a licensed pilot, and has sung tenor with Opera San José. You can contact Jim at http://jdlh.com/.

Dr. Martin Dürst

Professor, Aoyama Gakuin University

Dr. Martin J. Dürst is a Professor in the Department of Integrated Information Technology at Aoyama Gakuin University in Japan. Martin has been one of the main drivers of internationalization and the use of Unicode in a Web and Internet context. He published the first proposals for domain name Internationalization and composite character normalization, and is the main author of the W3C Character Model and the IRI (Internationalized Resource Identifier) specification. He has also been contributing to the Ruby implementation, mostly in the area of internationalization, since 2007. Martin teaches in Japanese and English, speaks fluent German, can get around in French, and studied It alian, Spanish, Russian, and Latin.

Mr. Behdad Esfahbod

Software Developer, Red Hat/GNOME

Behdad is an Iranian who grew up loving programming and typography. In high school, he was introduced to data structures and algorithms, and after a couple years of studying these concepts, he ended up pursuing a computer engineering BSc program at Sharif University of Technology, Tehran. It was around this time when he found about the true way of Unix, as well as Free Software, GNU, and GNOME projects. Nine years later, he's finished his MSc in computer science at the University of Toronto, and joined Red Hat in the Toronto office. He's become an expert in bidirectional scripts (like Arabic) and the Unicode standard, and would like to see Pango eventually used in a multilingual, internationalized, full-fledged print-quality desktop publishing system one day. He also dreams of a world where GNOME rocks on every desktop and laptop, and where he doesn't have to report bugs every other day.

Ms. Margie Foster

Localization Project Manager, Moblin Project, Intel Corporation

Margie Foster is currently the localization project manager for Moblin, an open source Linux-based operating system with an exciting new user interface designed for Intel's hottest new processor, the Atom. Margie has a love of localization going back many years, including working on videoconferencing software and help files, the Intel Year 2000 support website, and Intel's developer support website. She thoroughly enjoys working with people around the world, and has had opportunities to visit many of the countries of her L10N colleagues. Prior to her technology career at Intel, she was a certified operating room registered nurse. She whole-heartedly supports Loïc and Michael in their crusade for i18n and L10N awareness at Intel.

Benedicto Franco Jr.

Software Engineer, Yahoo! Inc.

Benedicto Franco Jr is an Internationalization Solution Architect and member of the Globalization team at Yahoo! Inc. He has more than two decades of software development experience.

Mr. Richard Ishida

Internationalization Lead, W3C

Richard Ishida is the W3C Internationalization Activity Lead. This activity has the mission of ensuring universal access to the Web, regardless of language, script or culture, by proposing & coordinating any techniques, conventions, guidelines and activities within the W3C that help to make and keep the Web international. For many years Richard's seminars and consulting have helped product groups around the world develop websites, documents, software, and on-screen information so that it can be easily localized for the international marketplace. His background includes translation and interpreting, computational linguistics, and translation tools. He has studied French, Spanish, Italian, German, Russian, Japanese and Arabic.

Mr. Tomas Galicia

Solutions Quality Analyst, IT Flex Service, Intel Corporation

Tomas Galicia is from Spain. He studied in Spain and France and came to the US after his University years. Soon after he started his career in translation and localization. He began working with Intel in 2003 focusing on localization of software and marketing material in Spanish. He currently works as a Solution Quality Analyst leveraging over 15 years of experience in the field.

Mr. Michael Kuperstein

Senior Localization Engineer, IT Flex Services, Intel Corporation

Michael Kuperstein has been working deep in the trenches of many localization projects, produced in partnership between Intel's in-house localization group and Intel business units. Michael was hired by Intel in 1996 as a software engineer, later transferring to Intel's localization team in 2001 as the dotcom bubble burst. He wears many hats as a localization engineer, software architect, application developer, tool wrangler, speaker, group historian, and all around go-to / fix-it guy for software internationalization. Armed with a vast array of creative concepts, software tools, internal social networking sites, defect reports and screenshots, financial data, and presentations, Michael is on a mission to evangelize proper internationalization and localization at Intel.

Mr. Norbert Lindenberg

Internationalization Architect, Yahoo! Inc.

Norbert Lindenberg is an internationalization architect at Yahoo!. He studied computer science at Universitat Karlsruhe and internationalization at Apple Computer, and then led internationalization projects at General Magic and Sun Microsystems.

Mr. Su Liu

AIX Globalization Architect, IBM

Su Liu, advisory software engineer, is working on globalization enablement for IBM's AIX operating system. His major areas of expertise are in character set conversion, ideographic language IME, keyboard, and complex text layout technologies.

Mr. Steven Loomis

Software Engineer, IBM

Steven R. Loomis is a member of the Globalization Center of Competency at IBM San Jose, where he is the Technical Lead for the International Components for Unicode for C/C++ (ICU4C). His ICU contributions include the Locale Explorer demo and the CLDR Survey Tool. After discovering the world of internationalization during a temporary assignment to a bidirectional text project, he joined the ICU team in 1998. His hobbies include Maltese language advocacy.

Dr. Ken Lunde

Senior Computer Scientist, Adobe

Ken Lunde has been working for Adobe Systems Incorporated, headquartered in San Jose, California, for over eighteen years, and is currently a Senior Computer Scientist in CJKV Type Development. He spent a year-and-a-half revising "CJKV Information Processing," and the second edition was published at the end of 2008.

Mr. James Lyle - *Program Manager, Microsoft*

Mr. Michael Manca

Project Manager and Solution Quality Analyst, IT Flex Services, Intel Corporation

Michael Manca was born in Portland, raised in southern Italy around the city of Lecce, and moved back to Oregon in '97 to complete his higher education studies (International Business and Marketing). He spent most of his professional career at Intel, approximately 5 years, as a Solution Quality Analyst for the localization of server management software products. In recent times, he transitioned to localization Project Manager and Translation Lead roles, still focusing on some of the more technical aspects of the localization process.

Mr. Kamal Mansour

Manager of Non-Latin Products, Monotype Imaging

An early multilingual education served to stimulate Kamal’s interest in languages, alphabets, and later on, typography. His studies have spanned Computer Science, Linguistics, and Product Design. Having worked at Monotype for over 13 years, Kamal has been involved in the many aspects of multilingual typography and font development. During that time, he has also participated actively in various activities related to the Unicode Standard. In the last few years, his work has included OpenType implementations for various scripts including Arabic, Greek, Latin, Hebrew, Thai, Lao, Khmer, and Japanese. Since 2006, he has served on the Board of Adviser of the Script Encoding Initiative of UC Berkeley. In Spring 2009, he taught a linguistics course at Stanford University entitled "Writing Systems in a Digital Age".

Mr. Mike McKenna

Internationalization Architect, Yahoo! Inc.

Michael is a specialist in globalization of applications and distributed systems with over one and a half decades of internationalization experience. He is a licensed professional engineer with extensive experience consulting or leading globalization projects for a number Fortune 500 companies and has a background in global e-commerce, application design, database internals, distributed bibliographic systems, test engineering, and ethnographic research. He is currently leading the I18n Architecture team at Yahoo! Inc.

Mr. Thomas Milo

President, DecoType

Thomas Milo is the president of DecoType and has been working on Arabic script technology since 1982, in the course of which pioneered the concept of Dynamic Font technology. DecoType has a close partnership with WinSoft, France. Tom served as a captain in the Royal Netherlands Army and did two tours of duty as an Arabic speaking officer in the United Nations Interim Force

in South Lebanon as a member of a contingent of armored infantry on peacekeeping duty. At the behest of the Netherlands Ministry of Defence, he wrote the Handbook of Lebanese Spoken Arabic for the Royal Netherlands Army (1981). Tom acts as a consultant for Basis Technology in Cambridge, MA, contributing to their Arabic and Persian Technology projects. Tom holds a Unicode Bulldog Award – whenever he remembers where he put it.

Katsuhiko Momoi

Staff Test Engineer & I18n Consultant, Google Inc.

Kat is currently a Staff Test Engineer at Google. He joined Netscape in 1996, where he initially worked as I18n Evangelist, then as Principal I18n Software QA Engineer and as Mozilla Technology Evangelist. Since joining Google in 2005, he has been working as an I18n Consultant and a Test Engineer for a variety of web applications. He has presented papers at W3C and Unicode Conferences as well as other Industry conferences in Japan and the US.

Mr. Derek Murnam

Senior Program Manager, Microsoft Corporation

Derek Murman is the Senior Lead Program Manager for MUI Technologies in the Windows International group. He has 12 years of experience at Microsoft. In addition his current position in .Windows International, Derek has worked on the Windows Update service and several enterprise server products.

Mr. Umesh Nair

Software Engineer, Google Inc.

Umesh P. Nair has been a software engineer for 18 years and is a part of the Internationalization team at Google since 2007. In addition to projects related to Unicode and Internationalization, he has contributed to several projects related to International calendars in Google as well as his previous jobs.

Mr. Mihai Nita

Globalization Architect, Adobe Systems, Inc.

Mihai Nita has been working in the localization/internationalization field for 12 years now, and still learning. Nita tries to cover internationalization not only for C/C++ and Windows, but also for Java, C#, Mac OS, Linux/UNIX, web technologies, client/server, the life, the universe and everything. As part of a small localization company in Silicon Valley, he has completed i18n/l10n projects, large and small, for well known or less known companies. He has written numerous guides for internal use or for customers, covering best practice for internationalization and localization, development, building, single sourcing, and testing. He held presentations and classes

on Java and XML internationalization, Web technologies internationalization, and single sourcing. Now he is doing more of the same at Adobe Systems, Inc. He spends most of his free time divided between various i18n newsgroups and learning more about his other passions: C++, system internals, security.

Anshuman Pandey

C.Phil. History, University of Michigan

Mr. Addison Phillips

Globalization Architect, Lab126 (Amazon)

Addison Phillips is the Globalization Architect for Lab126, creator of the Amazon Kindle e-book. He is the chair of the W3C Internationalization Working Group, a member of the Unicode Editorial Committee, and co-editor of IETF BCP 47.

Mr. Roozbeh Pournader

Internationalization Specialist, HighTech Passport

Roozbeh Pournader is an internationalization expert at HighTech Passport, a company specializing in localization and internationalization services. Roozbeh has been working on software bidirectionalization for his whole professional life, since 1996. As a native speaker of Persian, Roozbeh has had ample opportunity to experience the challenges of bidirectionalization firsthand. He is an author of multiple standards and technical reports for Persian support in software for Iran and Afghanistan, and a contributor to the Unicode Standard. Roozbeh contributes to various free software projects and is an advocate of open source software development. He represents the GNOME Foundation in the Unicode Consortium.

Elizabeth Pyatt

Instructional Designer, Penn State

Elizabeth Pyatt is an instructional designer at Penn State with a background in linguistics. She maintains the Penn State "Computing with Accents and Symbols Web Site" at tlt.its.psu.edu/suggestions/international/ and maintains a blog documenting her Unicode experiences at www.personal.psu.edu/ejp10/blogs/gotunicode/index.html.

Mr. Brent Ramerth

Software Engineer, Apple, Inc.

Brent Ramerth works on internationalization and natural language processing for the iPhone and Mac OS X platforms at Apple Inc.

Jens Riegelsberger

Google Inc.

Jens is a senior researcher in Google's User Experience team in London. Prior to Google he conducted research for LBi, Microsoft Research, and Amazon. Jens holds a PhD in Human Computer Interaction from University College London (UCL).

Charles Riley

Catalog Librarian for African Languages, Yale University

Charles Riley is a catalog librarian for African Languages at Sterling Memorial Library, Yale University. His time in the field includes research on scripts of six West African countries over the past five years, most recently in Guinea, Côte d'Ivoire, and a fourth return to Senegal, where he served as a Peace Corps Volunteer.

Dr. Murry Sargent III

Partner Software Design Engineer, Microsoft

Murray is a Partner Software Design Engineer at Microsoft, working mostly on the RichEdit editing engine and math editing and display in MS Office. He completed BS, MS, and PhD degrees in theoretical physics at Yale University and worked for 22 years in the theory & application of lasers, first at Bell Labs and then as a Professor at the University of Arizona. He also worked on technical word processing, writing the first math display program SCROLL (1969) and later (1980s) the PS technical word processor. More info is given in his blog on Math in Office (blogs.msdn.com/murrays).

Mr. Sumit Sarkar

i18n Product Specialist, DataDirect Technologies

Sumit Sarkar has been working on internationalization and localization data access issues using DataDirect products for 7 years. His focus is on i18n and performance of the data access layer for which he has developed a patent pending technology for its analysis. DataDirect products connect applications to an unparalleled range of data sources using standard-based interfaces such as ODBC, JDBC, ADO.NET, XQuery and SOAP. 96 of the Fortune 100 turn to DataDirect for their data access needs. Sumit holds a B.S. in Computer Science from N.C. State University.

Mr. Markus Scherer

Unicode Software Engineer, Google Inc.

Markus Scherer is a member of the Google software internationalization team, focusing on the effective use of Unicode and on the development and deployment of cross-product

internationalization libraries. Previously, he was manager, tech lead and software engineer at IBM. He has been a major contributor to ICU for ten years and designed and developed significant portions of the character conversion, bidi, normalization, Unicode properties, and collation functionality. Markus is an alumnus of the University of Kaiserslautern, Germany.

Mr. Ilya Shtein

IT Architect, Metavante

Ilya Shtein is an IT Architect with the Enterprise Software Architecture team at Metavante Corporation. His focus is on creating an architectural foundation for software internationalization across Metavante's Global Banking platform, as well as evangelizing Internationalization within the company.

Dr. Toshiya Suzuki

Research Assistant, Hiroshima University

Suzuki Toshiya (ACM) earned his Ph.D. degree in Dept. of Physics from Tohoku University, Sendai, 1998. Since 2000, he has been working around printing-, fonts- and text-layout-related softwares in Open Source Softwares, especially around PostScript technology: GNU Yellow Vector Editor, GNUStep, Display Ghostscript, gs-cjk project, Ghostscript and FreeType. Now he is a member of Information Media Centre of Hiroshima University.

Andrew Swerdlow

Internationalization Tech Program Mng, Google Inc.

Andrew joined Google in December 2005 and is currently the technical program manager for Internationalization(i18n). He holds a MSc/BSc of Computer Science from the University of Victoria, as well as professional certifications from Stanford University.

Mr. Frank Yung-Fong Tang

Sr. Software Engineer, Google Inc

Frank works for Google Inc. as software engineer since fall 2005. Frank spent the past 16 yeas developing global software. Before joining Google, Frank worked for AOL, Netscape, Apple Computer and III (Institute for Information Industry in Taiwan, ROC). In the last 10 years, Frank architected and led the Mozilla internationalization development and managed the Netscape Client International and Text Engineering team. Frank received his M.S. degree in Computer Science from Northeastern University, Boston and College Graduate Diploma from Ming-Hsing Junior Engineering College, Taiwan, ROC.

Mr. Tex Texin

Xen Master, XenCraft

Tex Texin has been providing globalization services including architecture, strategy, training, and implementation to the software industry for many years. Tex has created numerous globalize products, managed internationalization development teams, developed internationalization and localization tools, and guided companies in taking business to new regional markets. Tex is also an advocate for internationalization standards in software and on the Web. He is a representative to the Unicode Consortium and the World Wide Web Consortium. Tex maintains two Web sites for internationalization, the popular www.I18nGuy.com and www.XenCraft.com.

Mr. Wenchao Tong

Software Engineer, Google Inc.

Wenchao Tong is a software engineer at Google Inc. He is a member of the Google software internationalization team. He works on input technology projects, including the virtual keyboard api.

Ms. Kirti Velankar

Senior Software Engineer, Yahoo! Inc.

Kirti Velankar is a senior software engineer at Yahoo Inc. She is a part of the internationalization engineering group working on the internationalization libraries , backend platform work on worldwide Yahoo products and has contributed to the PHP 5 development with open source project with Zend Technologies. She has extensive experience in Object Oriented design and development in Java/J2EE. Prior to Yahoo, she was at Sun Microsystems, Inc.

Mr. Loic Dufresne de Virel

Localization Strategist, IT Flex Services, Intel Corporation

Loic Dufresne de Virel is currently a Localization Strategist with Intel in-house localization team. In this role, his main activities include promoting the use across Intel business units of a recently deployed Translation Management System, and constantly advocating for proper and improved Internationalization (I18N) and Localization (L10N) practices and processes, for Web, Software, and "print" collateral. Prior to moving to Oregon, and joining Intel, where he has been involved in Localization for the past 10 years, Loïc spent a few years in Costa Rica, working as a regional technical adviser for the UNCTAD.

Mr. Chris Weber

Casaba Security

Chris Weber is co-founder at Casaba Security where he’s leading product development for new tools to assist in the field of Unicode and Web-application security. He has spent years focusing on

software security testing for some of the world’s leading software development companies and online properties. He’s authored several security books, articles and presentations, and regularly speaks at industry conferences. He’s worked as a security researcher and consultant for over a decade identifying hundreds of security vulnerabilities in many widely used products.

五、 展示區廠商列表

Internationalization & Unicode Conference 33
October 14-16, 2009, San Jose, CA U.S.A

October 14-16, 2009 - San Jose, CA

Sponsor & Exhibitor Listing

The IUC 33 Exhibits will be open on Thursday from 10:00 to 20:00. Thursday refreshments breaks as well as the Thursday Evening Conference Reception will take place in the exhibit area in the Market Room on the Lobby Level.

Gold Sponsors

Adobe revolutionizes how the world engages with ideas and information. For 25 years, the company's award-winning software and technologies have redefined business, entertainment, and personal communications by setting new standards for producing and delivering content that engages people virtually anywhere at anytime. From rich images in print, video, and film to dynamic digital content for a variety of media, the impact of Adobe solutions is evident across industries and felt by anyone who creates, views, and interacts with information. With a reputation for excellence and a portfolio of many of the most respected and recognizable software brands, Adobe is one of the world's largest and most diversified software companies.

WinSoft

The successful internationalization and localization of software relies on an expert combination of linguistics, engineering and marketing knowledge. WinSoft has fine-tuned these skills during the nearly 25 years we have been localizing, creating and distributing international versions of branded software around the world. At WinSoft, we believe that users around the world should have access to quality, branded software in their language. Our key activity is creating and adapting software that fits the unique linguistic requirements not often met by standard product versions. WinSoft's product range is specifically adapted to Middle Eastern, South Asian and Central and Eastern European languages. Through our long-term partnership with major software companies like Apple, Adobe Systems & FileMaker, Inc, WinSoft has acquired strong references of efficiency and reliability.

For more information, and to see our complete product range, visit our website: www.winsoft-international.com.

Organizational & Media Sponsors

The Localization Industry Standards Association
Becoming a global enterprise is one of the most important challenges that your organization will ever face. There is no one right way to do it, but you should not have to reinvent the wheel. LISA is the leading international forum for organizations doing business globally. It has distilled the right ways and wrong ways of supporting international customers, products and services over the last fifteen years from more than 500 corporate members, public & private institutions, government ministries and trade organizations. You can now leverage this invaluable expertise - available nowhere else - in the form of clear guidelines, best practice and standards to enable your organization to deliver on its international objectives.

MultiLingual magazine covers the rapidly growing field of international language technology and global business communications. The content emphasizes information on international software development, localization and translation, with technical programming articles and market-related pieces. The magazine is mailed and posted online nine times a year (eight issues plus an annual resource directory/index) www.multilingual.com)