All About Arabic

▌ Overview

Arabic is one of the, if not _the_, most difficult languages to deal with in software.

This workshop teaches the key aspects of Arabic Language Information Processing. It covers the Arabic language and culture and the special relationship between them. The Arabic writing system, with its contextual rules and complex ligatures, is presented. We will look at Arabic characters, how they are entered, stored and displayed. All important Arabic coding systems will be covered. The features of bi-directional writing will be explained: directional characters, neutral characters, mirror characters etc. The Unicode algorithm will be discussed, and simple rules to follow in Word and HTML will be presented.

The workshop then covers the various tools available today to help process Arabic data and concludes by presenting insights and experience on doing business in the Arabic world (partnerships, sponsorships, transactions, etc.)

▌ Target Audience

This course is intended for software developers, web developers, testers and team leaders, as well as for executives and product managers (who need to understand why they must plan more time for Arabic support than for Chinese, for example). Localization engineers and translators that deal with the Arabic language will also benefit from this workshop.

▌ Benefits

After taking this workshop you will be ready to target the Arabic market with your product: you will know how to do business, what requirements to consider, what changes your software or Web site or embedded system requires, and how to implement those changes.

▌ Duration

The agenda described below is for a one-day session.

▌ Pre-requisites

None.

▌ Agenda

1. Arabic Culture

This module will explore the religious importance of writing in the Islamic world. We will review the history of Arabic and the current demographics of the Arabic language and where it is used in the world.

2. Arabic Writing System

This module covers Arabic writing including the alphabet, context analysis, sliding, diacritics and ligatures. We will review different Arabic scripts.

3. Arabic Data Representation

This module describes how Arabic data is represented on computer systems. We will look at logical versus visual order of Arabic text. The coding and rendering of Arabic data is discussed. A review of Arabic fonts and basic Arabic data entry, storage and display with examples from MS-Word, HTML and Oracle.

4. Arabic Language Processing

Of paramount importance to security applications is the ability to scan, search and identify Arabic text. This module reviews searching approaches, scanning tools, OCR technology. We will also discuss advances in Speech recognition technology and Arabic Machine Translation.

5. Arabic Information Systems

This module explores how Arabic text is used in many information systems. How is Arabic included in e-mail, HTML web pages, PDF files or in text? We will look at how Arabic data is represented on the web including image (gif) and text methods. A live demonstration of Arabic data entry and display is included.

6. Doing Business in the Arabic World

This module provides some insights into how to do business in the Arabic world. Several aspects of partnering and example business transactions are included. Insights and real business experiences are provided.

7. Tools and Technologies

This module will look into capabilities of tools and technologies available to support Arabic Language handling, display and processing. Companies in the Arabization business will be discussed and different Arabization and localization solutions will be presented.

▍ Handouts

Each attendee will receive a 200+ page booklet, with ample room for notes, complete with table of contents and glossary. The booklet is designed to serve as a practical easy-to-use reference "book" for regular use during an internationalization project.

About our Instructor – Matta Saikali

Matta has more than 25 years' experience in internationalization and localization testing. His testing experience covers more than 30 languages including European, Asian, Arabic, Hindi, etc.

Formerly Director of Software Quality Assurance at Gemplus, Matta built up and managed a team of 50+ SQA professionals responsible for testing globalized Windows applications and mobile devices in European and Asian languages.

As Director of SQA at Purkinje, Matta managed the testing team for a multilingual multi-user client-server application for clinical data entry.

Matta was also SQA team leader at ALIS where he was involved in testing all ALIS products, notably their Arabic/Farsi product line.

Additionally, Matta holds a B. Sc. and M. Sc. in Electrical Engineering.