

**Submitted by: Basis Technology**

**Presenter: Joel Ross**

**Topic: The Conundrum of New Online Languages: Translating Arabic Chat**

Online communications are playing an unprecedented role in propelling the revolutionary changes that are sweeping throughout the Middle East. A significant portion of that communication is in Romanized Arabic chat (Arabizi), which uses a combination of numerals and Roman characters, as well as non-Arabic words, to write Arabic in place of conventional Arabic script. Language purists in the Arabic-speaking world are lamenting that the use of Arabizi is becoming so profound that it is “destroying the Arabic language.”

Despite its widespread use, and significant effect on emerging societies, Government agencies and others have been unable to extract any useful data from Arabizi because of its unconventional characteristics. Therefore, they have had to rely on human, computer-savvy translators, who often are a burden on dwindling resources, and are easily overwhelmed by the sheer volume of incoming data.

Our presentation will explore the challenges of triaging and analyzing the Romanized Arabic format and describe Basis Technology’s Arabic chat translation software. This system will convert, for instance, *mo2amrat*, *mo2amaraat*, or *mou’amret* to مؤامرات. The output of standard Arabic can then be exploited for relevant information with a full set of other tools that will index/search, carry out linguistic analyses, extract entities, translate/transliterate names, and machine translate from the Arabic into English or other languages. Because of the nature of Arabizi – writers are able to express themselves in their native Arabic dialects, something that is not so easily done with Modern Standard Arabic – there is a bonus feature in that now we are also able to identify the probable geographical origins of each writer, something that is of great intelligence value.

Looking at real-world scenarios, we will discuss how the chat translator can be built into solutions for users to overcome technological, linguistic, and cultural obstacles to achieve operational success and complete tasks.