

Social Impact of MT Panel

Jennifer DeCamp
jdecamp@mitre.org

AMTA
August 9, 2006

MITRE

Another dimension of the Digital Divide?

- **Numbers of language pairs**

- >10,000 languages and dialects (ISO 639-3) = 100M pairs
- >6000 languages (conservative figure) = 36M pairs

- **Similar to situation with Unicode and word processing**

- High end
 - Unicode; technical papers on implementing encoding
 - Word processing; web publishing; dictionaries; hyphenation; spellchecking; grammar checking; etc.
 - Localized interface
- Stripped down
 - Unicode; technical papers on implementing encoding
 - Word processing; web publishing
- Barely possible
 - Unicode
 - Make-your-own keyboard; add your own fonts; use existing rendering
- None
 - Not in Unicode

What might this look like with MT?

- **Tiered capabilities**

- High quality MT
- Less high quality MT
 - May be via one of the above programs as a hub
- Quick and dirty MT
 - MT development services or MT kits; possibly for very specific domains and applications
- No MT
 - Need for human translation and/or use of Language of Wider Communication (LWC)

- **Language pairs added according to**

- Commercial interests
- Economic and/or political concerns (e.g., World Bank)
- Public interest (e.g., “saving” a language)

- **Resulting in groups with fewer capabilities for broader communication**