# ONLINE MT AT COMPUSERVE

Mary Flanagan

CompuServe Applied Research
5000 Arlington Centre Blvd
Columbus, OH 43220
tel: 614-457-8600
email: mflanagan@csi.compuserve.com

## 1. Overview

MT has traditionally been positioned as a tool for speeding the production of high-quality translations. In this model, the ideal input to the MT system is a well-written, domain-specific text, and the machine translated output is postedited by a language or subject matter professional. Yet this conception of MT has struggled to gain acceptance among translators, companies and individual users for a number of reasons, including concerns about measurement of productivity and the infrastructure changes MT requires. But perhaps the most important concern is that MT output is often low in quality compared to human translation, and it is not always clear how much postediting will be needed. This is a perceived obstacle for potential MT consumers, who usually must learn as much about what an MT system CANNOT do for their organization, as what it can do. CompuServe has taken a pragmatic approach to MT technology, focusing on finding a market niche for what it can do—generate rapid, very rough draft, information-scanning quality translations in an environment where quick scanning for content is more important than high quality.

CompuServe uses a customized version of the Transcend machine translation software for all of its online MT services. Customizations included a fully automatic pre-editing filter developed by CompuServe which identifies and corrects punctuation anomalies, format problems and blocks some proper names from translation, and dictionaries of online terminology containing approximately 5000 words per language pair. Bidirectional translation between English-French,

English-Spanish and English-German is offered. All of CompuServe's translation services run on a single Windows NT Pentium 100, and translation speeds exceed 3000 words per minute.

## 2. World Community Forum (production service)

CompuServe's World Community Forum is a bulletin board which integrates fully automated machine translation into a message board. World Community focuses on chat, and members are interested in topics such as travel, penpals, world politics and different cultures and languages. MT is used to support communication between people who don't speak the same language. CompuServe has developed MT dictionaries for many of the sections on the message board. Four separate "views" of the forum exist—one each for French, German, Spanish and English. CompuServe has developed an automated process which continuously logs on to each of the four views of the forum, collects new messages, machine translates them, and posts the target (with source appended at the end) to the appropriate target forum. Translations appear in approximately 2 minutes. Bidirectional translation between English-French, English-German and English-Spanish is available and the service translates approximately 30,000 words per day.

## 3. CompuServe Document Translation Service (production service)

CompuServe members can upload documents in either Text or RTF formats for machine translation. Two quality levels are offered: unedited MT at 1 cent per word, or postedited MT at 10 cents per word. All completed jobs are mailed to the user's CompuServe mailbox. Unedited MT jobs are usually completed within minutes. Postedited MT is first machine translated, then routed to a translation bureau where it is edited at a rate of 2,500 words per day. The service translates approximately 250,000 words per month.

## 4. Mail MT (planned service)

CompuServe's Mail MT service will enable members to send translated e-mail to other CompuServe members, or to themselves at very low cost. CompuServe members can attach a translation prefix to their outgoing mail messages to route mail through a machine translation server running the Intergraph Transcend software. The available language pairs are bidirectional English/French, English/Spanish and English/German. The translation process is entirely automated and produces unedited translations at very high speed. The translations are usually

suitable in quality for information scanning. CompuServe members are charged by the size of their mail message. The service will place machine translation capabilities in a high profile role; e-mail is a cornerstone product of any online service.

## 5. Chat Translation (prototype)

Chat translation continuously collects chat text from live chat rooms in a CompuServe forum, translates the text and posts the output to a companion window. CompuServe members who speak different languages will be able to participate in cross-lingual discussions in real time, posting and reading in their own language. Each participant need only enter the room for his/her language. All text entries made by the participant and others actually in that chat room will appear in the chat room window, as well as translated text from synchronized rooms. The chat translation software uses "room update daemons" to monitor each open chat room for new text, collect the text, submit it to the translation server for translation, then collect the results and post to the output window. Prior to public release, a software filter will be developed to identify and correct the format and punctuation errors typical of chat. Some vocabulary development will also be necessary.

## 6. Contacts

For further information about CompuServe's online translation services, please contact Mary Flanagan, CompuServe Applied Research, 5000 Arlington Centre Blvd, Columbus, OH 43220, telephone 614-457-8600 or email mflanagan@csi.compuserve.com.