

Machine Translation for the Office Automation

Key-Sun Choi

Center for Artificial Intelligence Research
Korea Advanced Institute of Science and Technology
Taejon 305-701 Republic of Korea
kschoi@cair.kaist.ac.kr

The office automation (OA) is for efficient interpersonal communication. This communication has two aspects: control and information. The control has two processes: workflow control and cooperation control. On the other hand, the information automation leads two problems: information sharing and information extraction. Terms of “sharing” and “control” assume the standardization for communication.

The workflow means a sequence of document processing. Here, a document is a representation of information. Initialization of workflow is to search a right path. One of search methods can come from the information classification. Each process of workflow is human investigation of document or its improvement. A human involved in each step has different knowledge and different experience. Office workers have different roles on, for example, bookkeeping, information extraction, decision-making, etc. Each of them requires different knowledge level: for example, foreign language ability, terminology knowledge, and expert knowledge. What can MT contribute to in this situation? What can MT help them, each of who has different knowledge level and different viewpoints or purpose? How is MT adept in workflow environment?

MT should integrate with information retrieval. This approach contributes to the right application of workflow to the foreign document. MT should also integrate with knowledge base for supporting the terminology and expert knowledge. However, a reader of document cannot interpret completely. The writer of the document can interpret it completely.

We can see one solution, that is, *document interchange format (DIF)*. DIF contains the basic terminology and writers' analyzed knowledge. This means that each document contains the standardized interchange format for right translation. Incoming DIF includes the knowledge to interpret it. The initial phase of workflow interprets an incoming document according to the local knowledge base plus the attached DIF. Why should each document contain such complex information? The authors of document can best analyze the right meaning of the document. DIF has several levels according to the use level of recipient. This means that the different use requires different contents of DIF.

The different user role and their knowledge level require the different MT strategy. DIF contributes to such flexible MT strategy.