



**INFORMATION SOCIETY
TECHNOLOGIES PROGRAMME**



**INTERNATIONAL STANDARDS
FOR LANGUAGE ENGINEERING**

Deliverable D 13.2

Machine Translation Evaluation:

**Draft chapter for
an evaluation handbook**

August 2002

Deliverable Identification Sheet

Project ref. no.	IST-1999-10647
Project acronym	ISLE
Project full title	International Standards for Language Engineering

Security (distribution level)	Public document
Contractual date of delivery	M22 = November 2001
Actual date of delivery	August 2002
Deliverable number	D-13.2
Deliverable name	Machine translation evaluation: draft chapter for a handbook
Type	Report
Status & version	Final 1.0
Number of pages	40
WP contributing to the deliverable	WP 13
WP / Task responsible	ISSCO
Other contributors	ISI
Author(s)	Margaret King, ISSCO; Andrei Popescu-Belis, ISSCO; Eduard Hovy, ISI
EC Project Officer	Brian Macklin
Keywords	Machine Translation Evaluation, Context-based Evaluation
Abstract (for dissemination)	<p><i>This document represents the draft of a chapter on MT Evaluation. It is currently under review for the Journal of Machine Translation.</i></p> <p>The Framework for Machine Translation Evaluation (FEMTI) defined here relates the quality model used to evaluate a machine translation system to the purpose and context of the system. We first summarize previous attempts to structure this domain, then present an overview of the ISO/IEC guidelines for software evaluation (9126, 14598). FEMTI, an application of these guidelines to MT software, is a framework that is made up of two inter-related classifications or taxonomies. The first classification enables evaluators to define a context of use, and provide a quality model (quality characteristics and metrics) for the respective context. The second classification provides definitions of various metrics used by the community. Finally, we explain how consensus is sought for, and how feedback from the community is taken into account in the FEMTI life-cycle.</p>