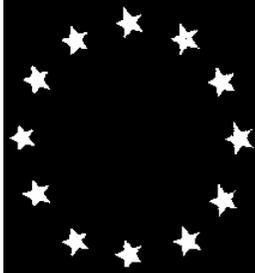


EURAMIS: Integrated Multilingual Services for a Large Multilingual Community.

EURAMIS stands for European Advanced Multilingual Information System. It combines hitherto separated multilingual services, like translation memories, terminology databases and machine translation into integrated services allowing for profile driven document preprocessing. Integration is implemented using a genuine linguistic resources data base together with a document repository. The data base is defined via SGML, providing for the necessary flexibility. The basic entry is a multilingual aligned text (MAT) allowing entries ranging from words to whole paragraphs. Considerable synergies arise from this approach and a general mechanism for feedback exploitation is provided. The system is based on e-mail and uses a HTML-3 derived pivot format with Unicode for the internal document representation. It is under construction and a first version will be available to all Commission staff in 1996.

Jean-Marie LEICK, curriculum:

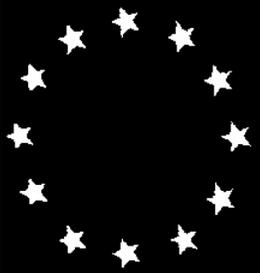
Born in 1945 in Luxembourg, graduate engineer in electronics (1968), assistant professor at the faculty of informatics, operating systems, at the Technical University of Karlsruhe, Germany (1970 - 1975), administrator at the European Commission since 1977, Directorate General XIII, Telecommunications, information market and exploitation of research, Information industry and market and language processing, Language processing and applications, (DG XIII), Statistical Office, since 1992 head of the Commission Action Plan for the Enhancement of Transfer of Information between languages, (DG XIII).



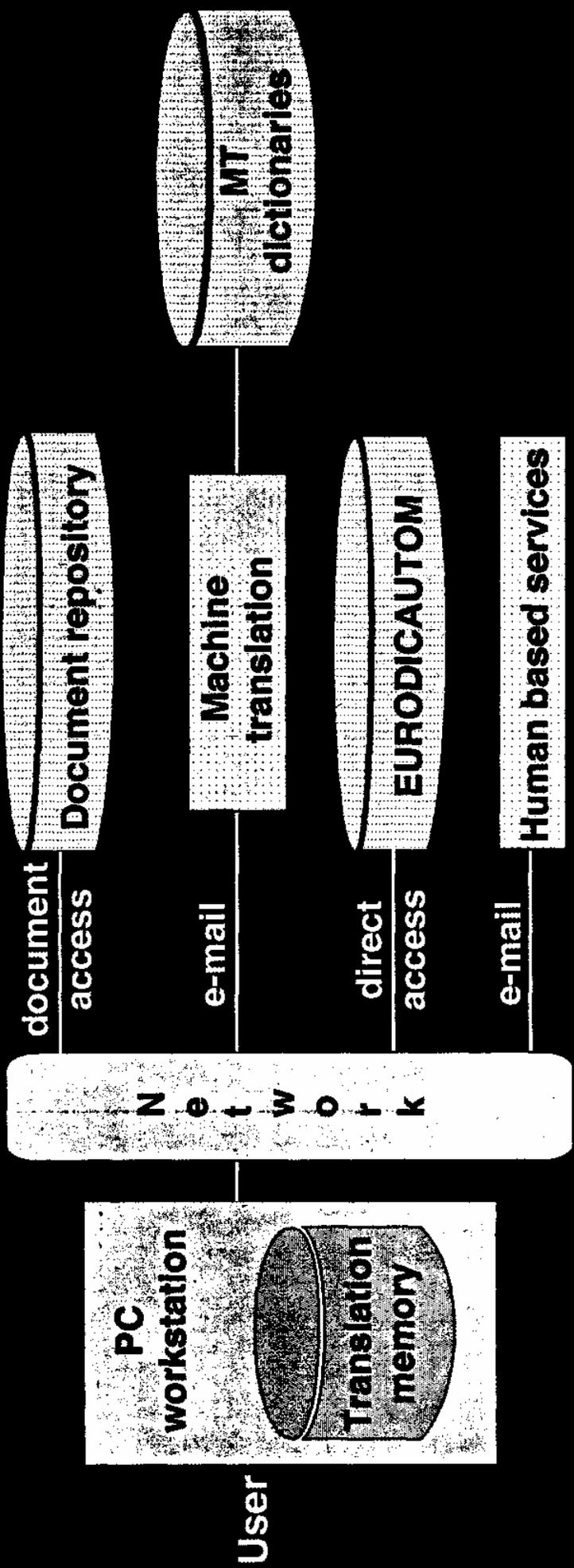
Multilingual Action Plan

EURAMIS

Advanced Information

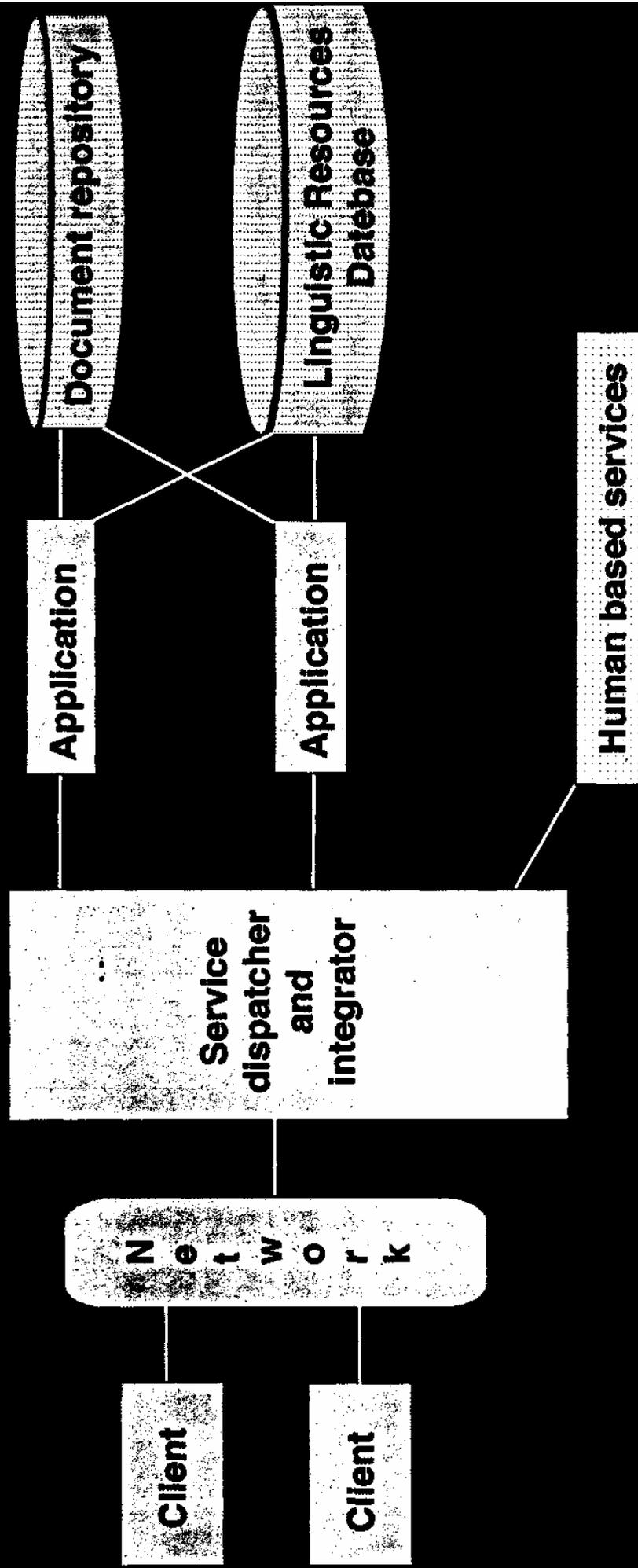


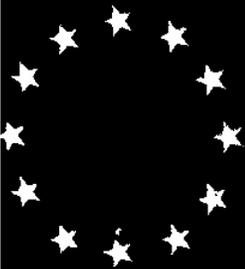
Multilingual Action Plan



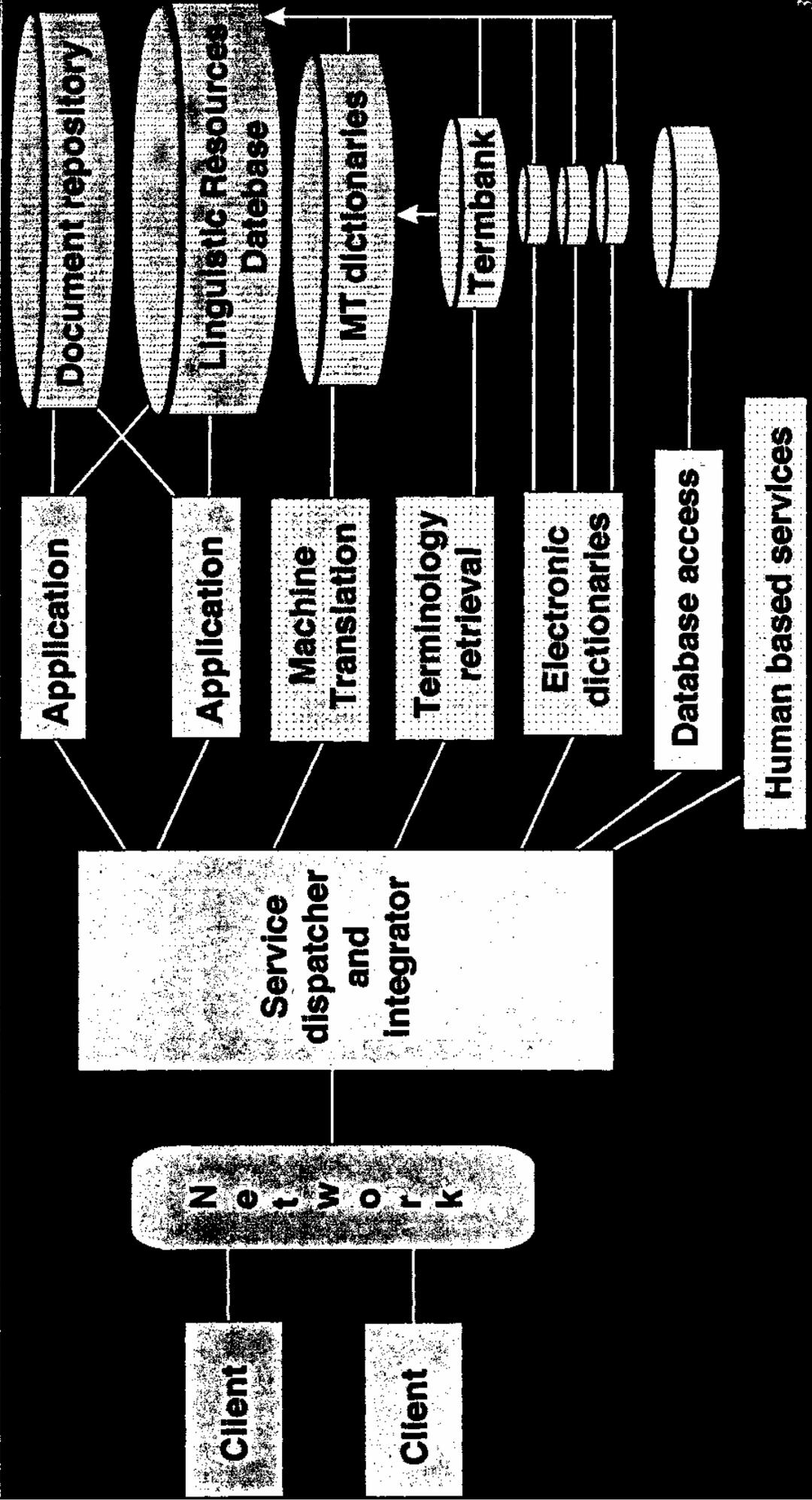


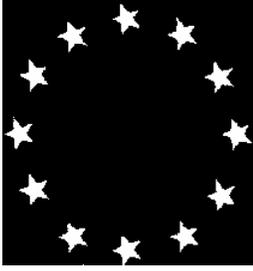
Multilingual Action Plan





Multilingual Action Plan

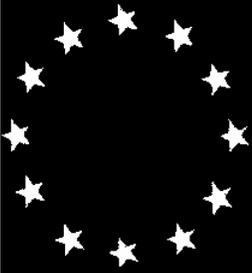




Multilingual Action Plan

Functional objectives:

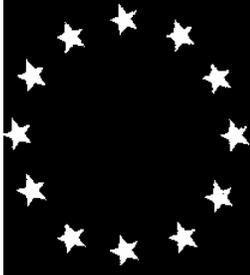
- ◆ Maximise reusability of translation resources (translation, terminology etc.)
by harvesting procedures for validation at different levels
- ◆ Optimise retrieval efficiency
- ◆ Allow synergy effects of currently separated applications
(translation memories, machine translation, termbanks, general purpose dictionaries, encyclopedias)



Basic Client functions:

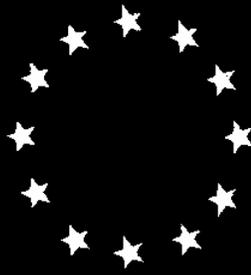
- ♦ Machine translation (using translation memory, termbanks and classical MT)
- ♦ Terminology lookup from text
- ♦ Terminology lookup from list
- ♦ Database query from text (e.g. CELEX)

All functions will be profile driven (end-users, translators, terminologists)



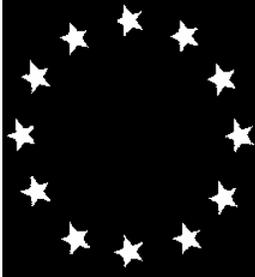
Advanced client functions:

- ♦ Text alignment
- ♦ Content identification
- ♦ Document version support
- ♦ Termbank workbench
- ♦ Text corpus management
- ♦ Private linguistic resources database management
- ♦ Access rights management



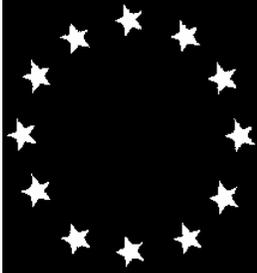
Multilingual Action Plan

- ♦ **Machine translation - analysis for generation of text specific glossaries**
- ♦ **Usage of multilingually aligned elements in LRDB for alignment of texts and their translations (by linguistic scoring)**
- ♦ **Content analysis and text characterisation via lemmatised terminology recognition**
- ♦ **Enrichment of terminological entries with grammatical informations (from MT)**
- ♦ **Crossreference and KWIC - procedures for consistency checks in MT and termbank development**
- ♦ **Advanced translation memory retrieval (lemmatised match)**
- ♦ **Document version control via document repository references**

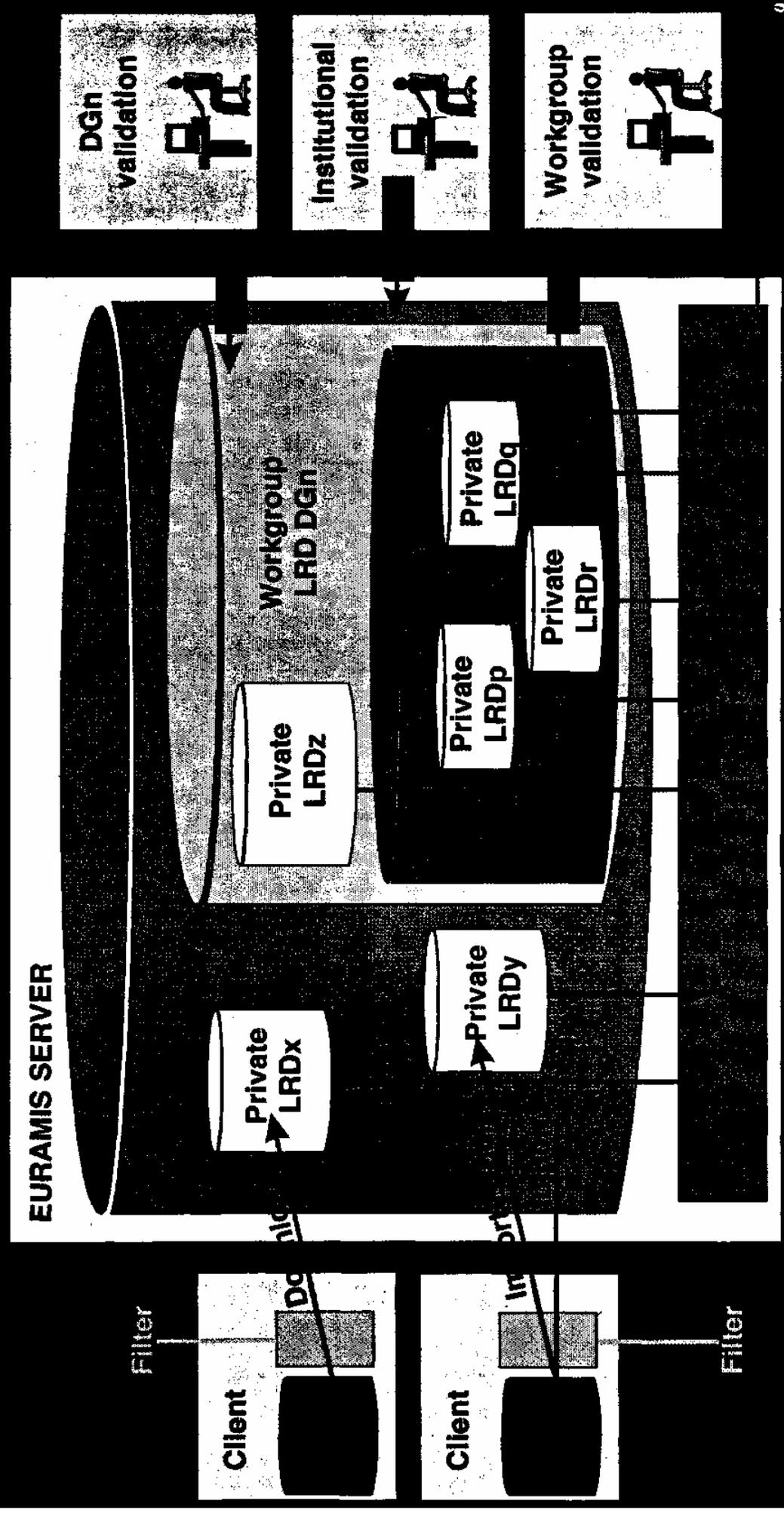


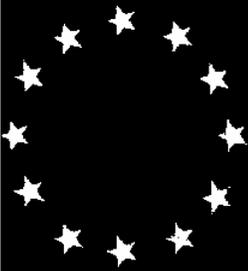
Design principles:

- ◆ **Basic elements: multilingually aligned entries**
from paragraph down to word form entries
with references to documents in the document repository
with SGML-oriented attribute-value pairs
capable of representing termbanks, MT, general purpose dictionaries
and encyclopedias
- ◆ **Structured access paths to areas of different validation levels**
private (may be local to a PC)
work groups of different levels
institutional
- ◆ **Distributed database support**

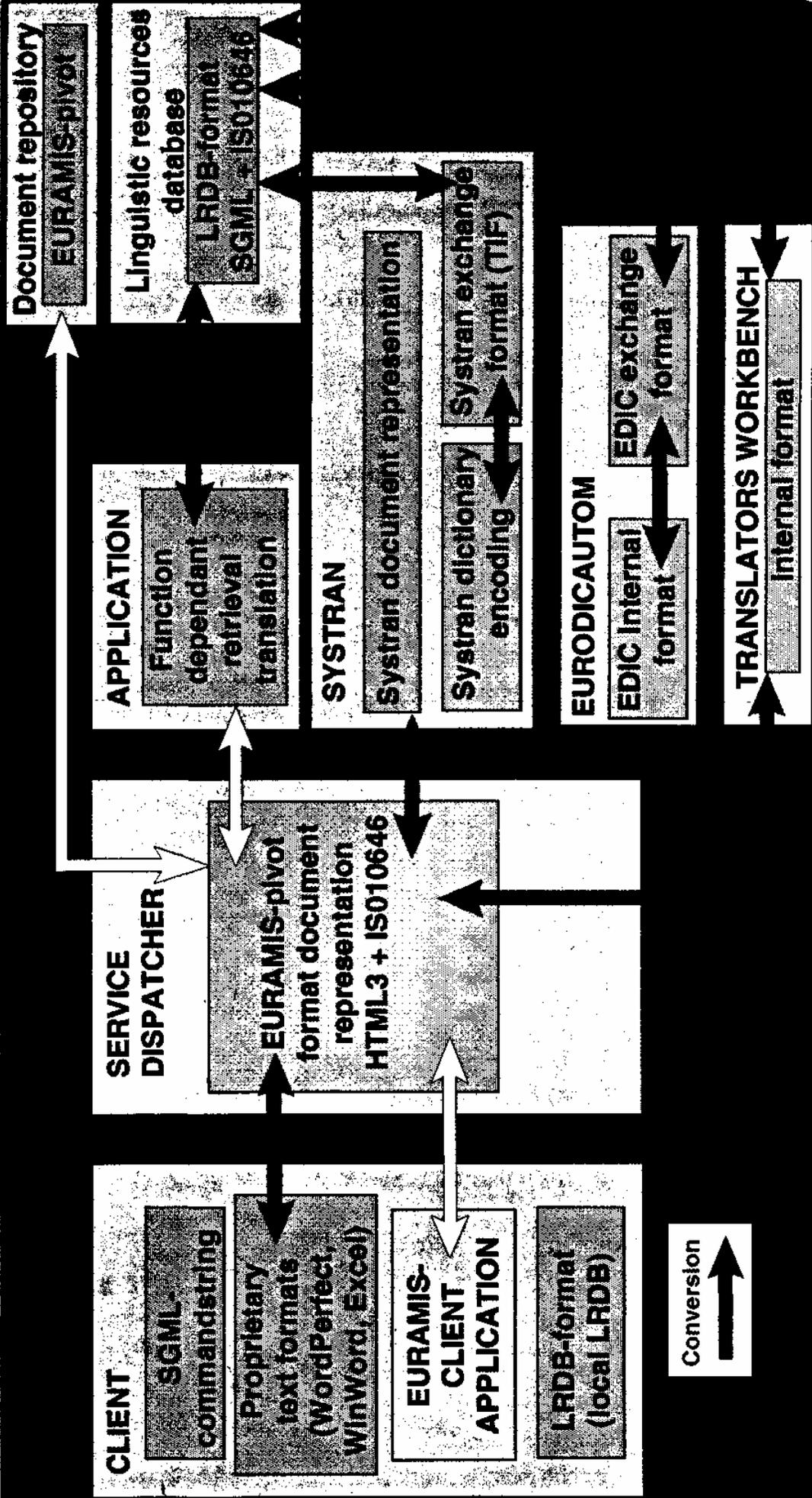


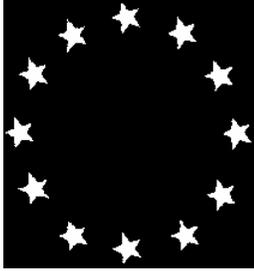
Multilingual Action Plan





Multilingual Action Plan





Multilingual Action Plan

Current approach: Electronic mail based

Advantages:

- Easily expandable (X.400 based EURAMIS third party server)
- No security problems (no logon)
- Standardised API interfaces available (MAPI)
- Response times sufficient for in-house mail (5-60 minutes)
- Batch service provision simple to organise

Disadvantages:

- No interactive functionalities
- Response times may be long in international communications

The overall design is not dependent on the network.