



CCiD赛迪

中国电子信息产业发展研究院

机器翻译发展汇报

## 单位情况简介

- 我单位隶属于国家一级科研单位-中国电子信息产业发展研究院（赛迪集团）；其前身是原信息产业部下属中文信息处理开放实验室。
- 我单位从“八五”开始得到国防科工委（总装备部的前身）的军事预研项目—军用共性软件技术项目—的支持，从事智能文本信息处理：包括中英文信息处理和机器翻译研究以及倾向性分析研究直到现在。

## 我单位机器翻译研发与应用现状

赛迪智能翻译系统通过采用国际领先的**MT+TM**技术和基于混合引擎设计理念，利用强大的自动翻译功能和辅助功能来满足不同客户的多种需求，为中英互译需求者提供全面解决方案。



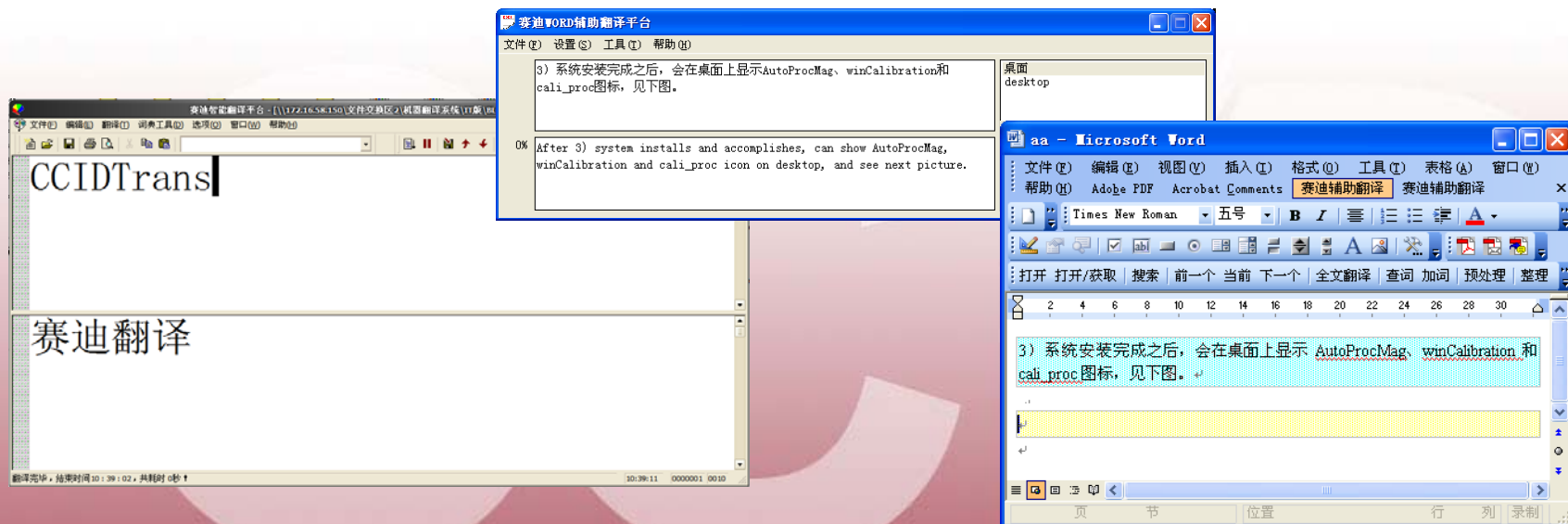
图为我公司为部队用户进行培训讲解赛迪智能翻译系统

利用赛迪智能翻译系统的自动翻译功能，可快速翻译各种英文情报、档案资料、科技文献，为有效处理海量英文信息提供解决方案；

将赛迪智能翻译系统集成到生产性翻译流中，成本和效率改进**30%**以上，而且保证各种术语、短语的译文准确、统一和规范；

## 我单位机器翻译研发与应用现状

- 利用赛迪智能翻译系统可全面整合各单位中英双语资料，打造中英双语智能知识库、资料库，有利于中英双语的积累。
- 多种文档格式处理能力，可翻译处理文本、Word、Excel、PPT、Html等文件格式，机器翻译系统的功能键已经嵌入到办公软件界面之中，非常便于客户操作使用。



## 机器翻译研发与应用现状

- 现在，**MT**的用途已经划分的很细，根据**MT**的特性，主要有两个用途：一个是信息浏览，要求的是**MT**的可读性；另一个当做翻译平台工具，要求的是准确度。以上用户的需求对**MT**的要求已经截然不同。
- 信息浏览更注重语句的可读性和连续性，单个词汇的准确度放到了次要位置，语句的逻辑性错误和不连续造成可读性的下降会导致阅读者信息识别率的降低和阅读速度的下降。基于以上原因，对**MT**系统的调试更注重语法规则的调试。

## 机器翻译研发与应用现状

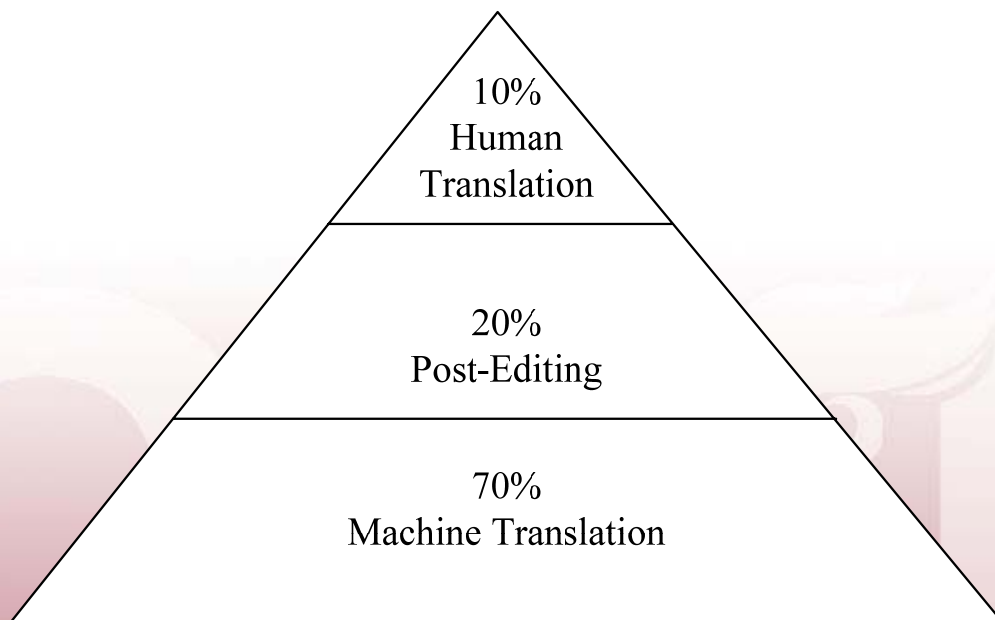
- **MT**作为翻译平台工具，对**MT**要求更多的是准确度。译者在平台上直接对翻译过的译文进行修改，对于大多数译文来说，无论译文的可读性如何，离出版精度的书面语都会有差距，译者都要重新组织语句架构，故可读性的重要性下降，单词尤其是短语的翻译准确性的重要性上升。译者更注重的是要把这些已经精确翻译过短语片段重新“组装”起来。当然，以上事件的发生会有一个例外：如果翻译说明书或模板化语言的时候，即使是**MT**作为翻译平台工具，可读性的重要性仍会大幅上升。

## 机器翻译研发与应用现状

- 另外，机器翻译在全球来讲是个未来趋势。随着全球化的到来，海量的本地化语言不可能都翻译成外语或外语也不可能都翻译成本地化语言。根据国外的成熟经验，一个金字塔形状的翻译市场正在形成：



# 机器翻译研发与应用现状





## 机器翻译研发与应用现状

- 从上述图形中我们可以看出，受限于成本，只有10%的资源需要人工翻译；70%的文档都会采用机器翻译。中间的20%的资源会采用post-editing 即后编辑的方式进行翻译。其精度介乎于人工翻译和机器翻译之间，成本也大约只有人工翻译的一半左右；同时在此处会形成一个很大的市场并创造出新的工作岗位-机器翻译后编辑：这个岗位不是翻译，机器翻译后编辑的岗位人员对MT的翻译非常熟悉，知道什么地方MT会出错或者翻译的不太通畅，他们会离开原文直接对译文进行polishing；处理译文的速度是人工翻译的2-3倍。

谢谢

CCiD