



# 人工智能与信息社会

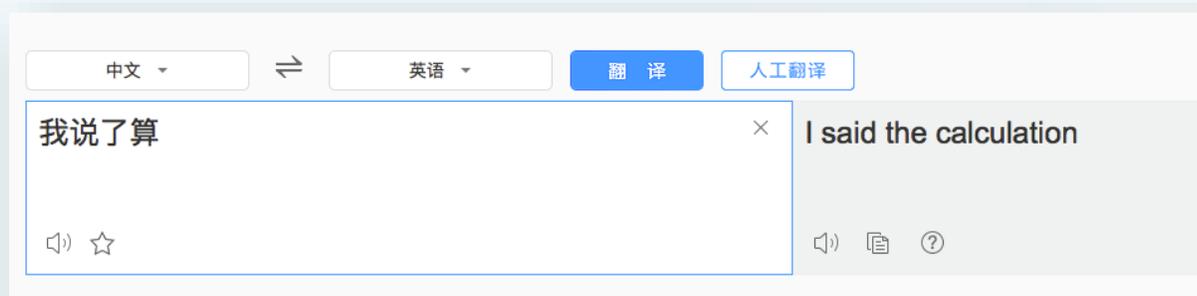
陈斌 北京大学地球与空间科学学院 [gischen@pku.edu.cn](mailto:gischen@pku.edu.cn)

# AI翻译

- › 谷歌翻译智能化
- › 微软机器翻译里程碑
- › 有道翻译官
- › 实时视频字幕  
语音识别+在线翻译

# 谷歌翻译

- › 生活和工作中少不了要接触英文，这时候就要依赖于机器翻译。
- › 但翻译有时候也会闹笑话，因为机器翻译总容易翻出一些奇怪的句子



# 谷歌翻译

## › 旧翻译方法：统计机器翻译SMT

将两种语言中同义的句子切分成词进行匹配

所有的结论都是机器根据统计数据得出的

## › 统计机器翻译类型

基于词

- 将句子分解成单词，最直接的想法

基于短语

- 分成 $n$ 个单词的组合短语，提高了准确性。

基于句法

- 对句子进行精确的句法分析，确定主谓宾等
- 效果没有预期的理想

# 谷歌翻译

## › 智能翻译方法：

2016年9月，谷歌宣布了神经机器翻译

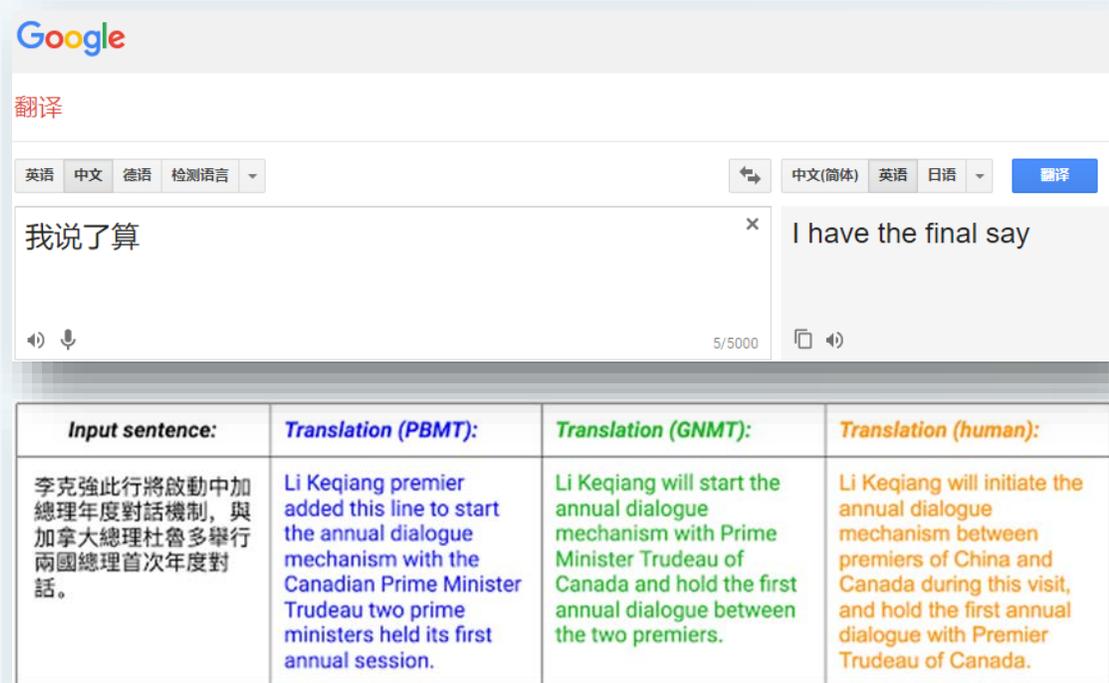
## › 神经机器翻译可以说是一种颠覆。

神经机器翻译的单词错误减少了50%，词汇错误减少17%，语法错误减少19%。

# 谷歌翻译

## › 结果逼近人工翻译：

这个全新的机器学习翻译系统，将全面启用在“中文翻英文”，其翻译结果更通顺流畅



The screenshot shows the Google Translate interface. The input text is "我说了算" (I have the final say). The output shows three translations: PBMT (blue), GNMT (green), and human (orange). The GNMT translation is significantly more accurate and fluent than the PBMT translation.

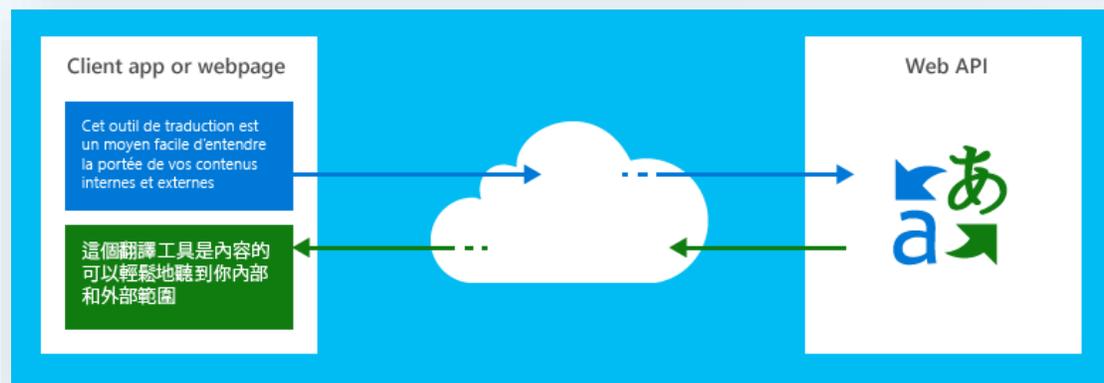
Input sentence:	Translation (PBMT):	Translation (GNMT):	Translation (human):
李克強此行將啟動中加總理年度對話機制，與加拿大總理杜魯多舉行兩國總理首次年度對話。	Li Keqiang premier added this line to start the annual dialogue mechanism with the Canadian Prime Minister Trudeau two prime ministers held its first annual session.	Li Keqiang will start the annual dialogue mechanism with Prime Minister Trudeau of Canada and hold the first annual dialogue between the two premiers.	Li Keqiang will initiate the annual dialogue mechanism between premiers of China and Canada during this visit, and hold the first annual dialogue with Premier Trudeau of Canada.

# 微软机器翻译

## 通用新闻报道：

微软机器翻译系统在通用新闻报道测试集上，达到了可与人工翻译媲美的水平

首个在新闻报道的翻译质量和准确率上可以比肩人工翻译的翻译系统



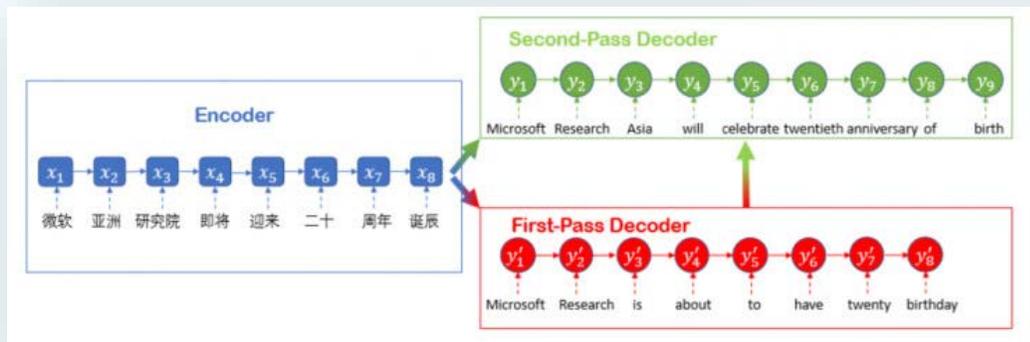
# 微软机器翻译

## 对偶学习:

中译英和英译中两个任务存在对偶结构  
两个任务可以互相提供反馈信息

## 推敲网络:

任务完成之后, 继续反复推敲  
基于前一阶段翻译结果进行润色



# 国内智能翻译产品

## > 有道翻译官：

语音对讲翻译

拍照翻译



## > 网易见外

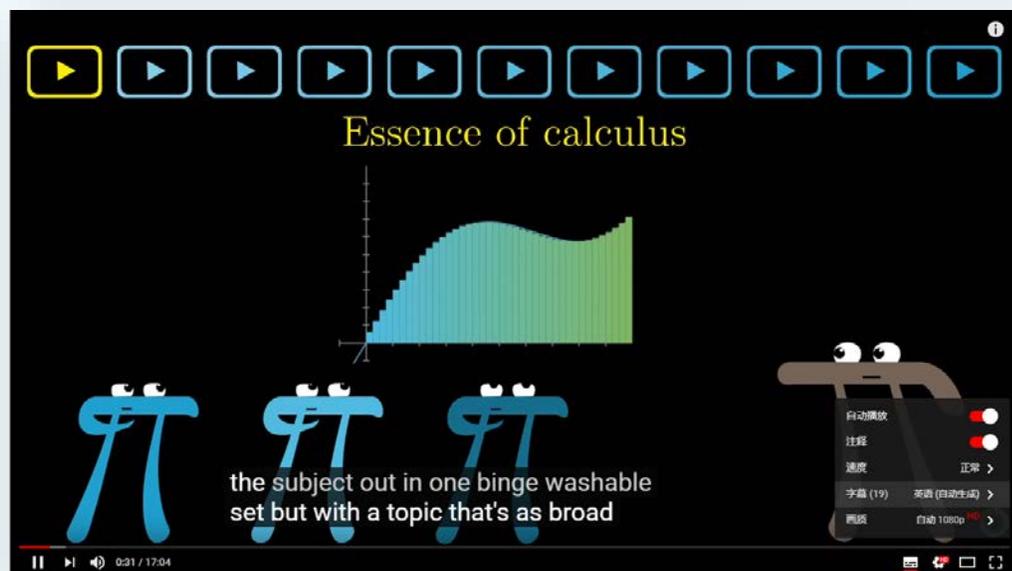
视频即时翻译



# 实时视频字幕

## › YouTube自动字幕

超过 10 亿个线上影片的字幕，是由系统自动产生



# 实时视频字幕

## › 2009年Youtube启动自动字幕

利用语音识别技术生成英文字幕

## › 人工智能语音识别技术发展

大幅度提高语音识别准确率

## › 机器翻译技术发展

将识别的英文通过翻译技术实时翻译成中文等其他非英语字幕