



清华大学人工智能研究院
知识智能研究中心
Knowledge Intelligence Research Center

HowNet的发展历程

时间地点主讲人等信息



CONTENTS 目录

- 1/ 问题的提出
- 2/ 发展历程
- 3/ OpenHowNet





■ 问题的提出

1988年前后，董振东先生曾在他的几篇文章中提出以下的观点：

- 自然语言处理需要强大的知识库的支持。
- 什么是知识
- 如何构建
- 谁来构建



■ 问题的提出

相同与不同：

- ◆ 相同：自然语言处理需要知识已成为共识
- ◆ 不同：组织形式不同，知识图谱、树库、HowNet等
- ◆ 能否有机的结合



■ HowNet 30年的历程

1988年设计探索

1997年工程实施

2000年公布推出

2002年设计新的HowNet知识库描述语言 (KDML)

2006年HowNet专著出版《HowNet and the Computation of meaning》,并同年由世界科学出版社出版发行。

2012年12月荣获"钱伟长中文信息处理科学技术奖"一等奖。

2014年我们推出了基于HowNet的英汉机器翻译系统。

2016年我们推出了基于HowNet的中文文本分析系统。

2018年HowNet中文文本分析器正式应用

2019年与清华大学共同推出OpenHowNet



■ HowNet 30年的历程

探索与实施

- ◆ 1988年设计探索
- ◆ 1997年工程实施
- ◆ 2000年公布推出



■ HowNet 30年的历程

设计探索的难点：如何认定义原？

HowNet观察探寻义原大体方法如下：

- ◆ 选定4000个汉字，将它们的义项一一列出来；
- ◆ 对列出的义项同类归并；这样例如汉字中的“医”、“治”、“看”、“瞧”、“疗”所具有的“医治”暂定为一个义原。
- ◆ 对归并后余下的义原，观察它们可能存在的关系，研究它们是什么关系。例如，我们发现了它们主要体现了“横向”关系，如：“医”、“药”、“患”、“病”等



■ HowNet 30年的历程

调整：

- ◆ 2002年设计新的HowNet知识库描述语言 (KDML)
- ◆ 2006年HowNet专著出版《HowNet and the Computation of meaning》,并同年由世界科学出版社出版发行。
- ◆ 2012年12月荣获"钱伟长中文信息处理科学技术奖"一等奖。

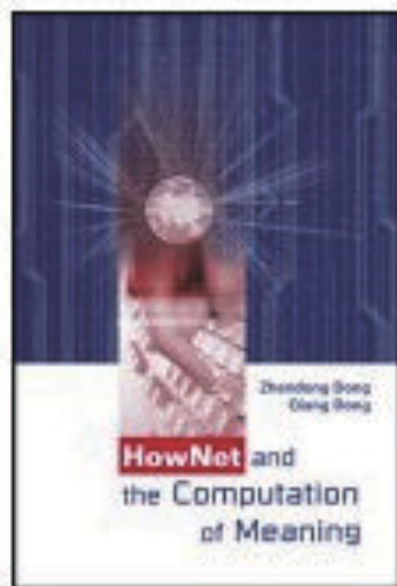


HowNet 30年的历程

HowNet专著

《HowNet and the Computation of meaning》

Hownet And the Computation of Meaning



Buy A Book!
amazon.com.

Source

Medium: Hardcover
Year of Publication: 2006
ISBN: 9812564918

Authors

[Zhendong Dong](#)
[Qiang Dong](#)

Publisher

World Scientific Publishing Co., Inc. River Edge, NJ, USA



■ HowNet 30年的历程

HowNet专著

《HowNet and the Computation of meaning》

I hope that this book will be read widely — and that HowNet will come to be used — not only by researchers in the field of natural language processing, but by many people throughout the world who are involved in the study of languages.

長尾真

Makoto Nagao

President, National Institute of Information and
Communications Technology (NICT)

Former President, Kyoto University



■ HowNet 30年的历程

面向应用的研究：

- ◆ 2012年开发出词语相关性计算器
- ◆ 2012年开发出语义群计算器
- ◆ 2014年我们推出了基于HowNet的英汉机器翻译系统
- ◆ 2016年我们推出了基于HowNet的中文文本分析系统
- ◆ 2016年我们推出了基于HowNet的英文文本分析系统



■ HowNet 30年的历程

HowNet的变化

- ◆ 知识词典信息变化
- ◆ 中英文词语同步增加和更新，与时俱进



■ HowNet 30年的历程

核心不变：

- ◆ 哲学、义原、关系、体系架构、知识描述语言不变
- ◆ 从而体现了知网的设计的正确性，系统的稳定性，与时俱进的适应能力。



■ HowNet 30年的历程

落地与推广：

- ◆ 2018年HowNet中文文本分析器正式应用
- ◆ 2019年与清华大学共同推出OpenHowNet



■ HowNet 30年的历程

近20年来得到海内外同行业界的可贵支持

近日倪光南院士在给董振东的信中写道：

“董老师：

知网已经发挥了很好的作用，现在语言信息处理发展很快，它的应用也越来越普遍，我相信你们的平台会受到大家的关注。”

“知网做得很好，我试了一下，大家都能享用自然语言处理的成果了！由于将高深的知识赋予给计算机，使自然语言计算成为可能，我相信这将会推动该领域的发展和扩大成果的推广应用。”



■ OpenHowNet

与深度学习相结合

- ◆ 用概念或者义原代替关键字
- ◆ 降低维度
- ◆ 减少歧义
- ◆ 不可控成为可控



■ OpenHowNet

项目地址：<https://openhownet.thunlp.org>

特点：

- 首次提供知网核心数据的开放下载
- 在线搜索知网词条、展示义原结构
- 提供Python 调用接口，方便查询使用
- 提供基于义原的语义表示学习结果下载（包含词、义项、义原向量）



清华大学人工智能研究院
知识智能研究中心
Knowledge Intelligence Research Center

谢谢大家

