

# 知识搜索引擎与翻译

徐意想, 倪蓉

(上海理工大学 外语学院, 上海 200093)

**摘要:** 信息化进程的纵深发展对翻译行业产生了深刻的影响, 译者如何从海量信息中高效获取信息关系到其翻译服务的质量。因此, 译者的信息素养和搜索能力显得愈发重要。通过探讨搜索引擎的补充和延伸——知识搜索引擎, 着重论述知识搜索引擎相对于传统搜索引擎, 在主题分类、一站式知识检索、互动式知识问答、个性化知识库等方面具备的突出优势, 并对比国内三大知识搜索引擎, 论述其在翻译实践中的重要启示。

**关键词:** 信息素养; 搜索引擎; 知识搜索引擎; 翻译

**中图分类号:** H 059      **文献标志码:** A      **文章编号:** 1009-895X(2020)04-0315-05

**DOI:** 10.13256/j.cnki.jusst.sse.2020.04.003

## Knowledge Search Engine and Translation

XU Yixiang, NI Rong

(College of Foreign Languages, University of Shanghai for Science and Technology, Shanghai 200093, China)

**Abstract:** The in-depth development of informatization has a profound impact on the translation industry. How a translator effectively obtains useful information from the massive information sources is critical to the quality of his translation service. Thus, information literacy and searching ability are of increasingly importance for translators. This paper focuses on knowledge search engine, a way of supplementing and extending search engine. The author compares the search engine with knowledge search engine, discussing its advantages over traditional search engines in thematic classification, one-stop knowledge searching, community question answering, personalized knowledge base and so on. The author also compares three major knowledge search engines in China, further stating enlightenment of knowledge search engine in translation practice.

**Keywords:** information literacy; search engine; knowledge search engine; translation

当今时代, 信息爆炸对翻译行业产生了深刻的影响, 尤其改变了传统译者的工作面貌。一方面, 海量的信息为译者翻译提供了超级信息平台; 另一方面, 信息资源呈现海量化、多元化、碎片化和非结构化, 为译者高效获取信息带来挑战。因此, 信息素养作为当今译者必备的能力则显得尤为重要。信息素养是指译者为胜任翻译工作, 积极利用各种

信息工具与资源, 主动开展信息实践活动而发生的批判性知识发现与建构体验<sup>[1]</sup>。

译者信息素养的重要性体现在哪方面呢? ——搜商。陈沛在《搜商: 人类的第三种能力》中阐述了通过搜索获取知识的能力已经成为与人类情商、智商并列的第三种能力<sup>[2]</sup>。王华树和张成智从译者翻译能力高度论证搜商是职业译员能力的重

收稿日期: 2019-11-24

作者简介: 徐意想(1994-), 女, 硕士研究生。研究方向: 英语科技笔译。E-mail: 1445529559@qq.com

要组成部分,明确指出搜索能力是译者翻译能力的一部分<sup>[3]</sup>。因此,译者如何从海量信息中挖掘有用信息为翻译实践服务,是决定其翻译质量和效能的关键因素。信息爆炸引发搜索革命,搜索引擎是译者翻译工作中有力的辅助工具,然而许多问题是传统搜索引擎难以有效解决的,因此作为搜索引擎的补充和延伸——知识搜索引擎应运而生。

一、什么是知识搜索引擎

随着互联网信息资源日益丰富,国内互联网公司如百度、腾讯、搜狗、新浪等以及国外的互联网公司为了满足用户更简单高效获取信息的需求,都相继开发了知识搜索引擎以弥补搜索引擎的不足,如国内有百度知道、新浪爱问、搜狗问问等,国外有美国的 WolframAlpha、韩国的知识 in、日本的知识 plus 等。

高雯雯和杨木锐对国内知识搜索引擎进行了探究,认为知识搜索涵盖五个核心要素即知识的创造、评价、利用、共享、积累<sup>[4]</sup>;彭陶和马张华指出,知识搜索引擎是 web2.0 智能化发展阶段新

一代互动式知识服务的产物,实现了知识共享社区和搜索引擎的结合<sup>[5]</sup>。因此,知识搜索引擎并非是一种单纯的信息检索工具,它可为网民提供个性化、条理化、知识活性化的信息交流平台,是传统搜索引擎的补充和延伸。知识搜索引擎引入主题分类和问答机制,为用户反馈更有针对性、更准确的知识,充分激活人类隐性知识,实现知识搜索与共享。

知识搜索引擎的工作原理首先是用户搜索。其次,系统在基于问题分类上进入搜索模块,在本地知识库中搜索匹配答案。如果答案匹配失败,则用户可选择进入问答平台,由用户提出问题,通过社区回答或专家答疑反馈答案。得到的答案再进入评价监督环节进行审核,最终返回高质量的答案。最后用户的问答会累积到本地知识库,实现知识库的动态更新。简言之,知识搜索是实现知识管理、知识交互、知识共享和知识服务的搜索工具。

笔者基于检索对象、检索方式、在线问答、知识库管理和检索结果五个方面比较传统搜索引擎和知识搜索引擎的区别,便于读者更好地了解知识搜索引擎的特点。

表 1 传统搜索引擎与知识搜索引擎的比较  
Tab.1 Comparison between the traditional search engine and knowledge search engine

| 比较对象  | 知识搜索引擎                                                              | 传统搜索引擎                              |
|-------|---------------------------------------------------------------------|-------------------------------------|
| 检索对象  | 基于用户问答建立的庞大本地知识库,资源对象比较单纯                                           | 基于整个互联网资源,数据来源庞大、复杂                 |
| 检索方式  | 引入主题分类将知识库分门别类,或通过高级搜索、二次搜索等限定检索问题的性质进行搜索                           | 网络内容不设分类问题和主题分类体系,网络爬虫抓取和存储全网资源     |
| 在线问答  | 信息检索匹配失败,可进入在线问答,应用先进的自然语言提问技术,提供 24 小时问答内容审查,同时允许用户评价、交流、修改、维护问答信息 | 只设置信息检索,基于用户关键词提问返回检索结果             |
| 知识库管理 | 可保存已解决的答案、评价等,系统通过分类关联起来,形成动态知识库;实现优先自动检索和利用知识库里的知识,实现知识管理和共享       | 提供收藏夹供用户整理收藏的链接,收藏夹里的资源不能成为检索库里的新资源 |
| 检索结果  | 用户可对回答者的答案进行监督,系统自动纠错;检索结果返回基于评价、时效、相关度等综合权重系数,保证开放网络环境下的问答质量       | 注重内容为王,搜索结果往往返回数以万计的网页,内容鱼龙混杂       |

二、知识搜索引擎对翻译实践的启示

自 2005 年 6 月百度知道、新浪爱问相继上线后,国内知识搜索引擎得到了很大的发展,各大互联网公司也推出自己的知识搜索引擎。高雯雯、杨木锐等人集中对“百度知道”“新浪爱问”的发展进行了探究<sup>[5]</sup>,同时,“搜狗问问”作为搜狗旗下

最大的知识产品,成为中国最具影响力的社区问答平台之一。

国内三大搜索引擎在类目设置、页面设置以及搜索方式上既有相同点也有不同之处。三大搜索引擎在类目设置上均设有主题分类搜索,一级类目下根据主题不同设置不同数量的细分类目;在页面设置上均设有搜索、提问、问题分类、个人中心、热门推荐和意见反馈模块;在搜索方式上均设有基本

搜索和类目浏览搜索,此外三大搜索引擎在搜索结果中均设有相关推荐链接,可跳转到不同链接查看

相关搜索结果。另外针对三大搜索引擎的不同之处,笔者归纳如表2所示。

表2 国内三大知识搜索引擎比较

Tab. 2 Comparison of the three knowledge search engines in China

| 比较对象 | 百度知道                  | 新浪爱问                                      | 搜狗问问                                  |
|------|-----------------------|-------------------------------------------|---------------------------------------|
| 类目设置 | 一级类目:14<br>类目级别:4     | 一级类目:14<br>类目级别:3                         | 一级类目:14<br>类目级别:4                     |
| 页面设置 | 栏目、用户、权威合作、大家热议、影响力排行 | 发现、帮帮团、大家在问、爱问最强大脑、每周风云榜、最新问答、精选问答、常见热点问答 | 特色、问豆商城、问问医生、问问律师、问问专家、问问达人、问问专栏、问问动态 |
| 搜索方式 | 高级搜索                  | —                                         | 站内搜索、全网搜索                             |
| 其他   | 搜索结果提示相关超链接           | 设有爱问共享资料、爱问地区分类服务                         | 搜索结果提示相关超链接                           |

### (一) 主题分类对翻译实践的启示

主题分类原理是用户就其提出的问题自行贴标签,归类到知识搜索引擎某一类目中或者由系统为用户所提的问题进行自定义归类,最终基于用户的问题-回答集合累积形成知识库。因此,知识搜索引擎可将碎片化的知识进行主题分类,整合成一个系统的知识库,知识库涵盖的内容相比基于网络随机抓取的内容较为单纯,可有效实现从堆到筛,极大优化检索效能。

笔者在翻译一篇英中转基因工程的科普文时,文中涉及大量的术语和概念。因为对转基因生物技术领域所知甚少,如果笔者对一些专业术语只知其表不知其里,必然会导致背景知识理解不到位,影响语篇翻译的准确性。例如文章提到“*These techniques are much more precise than mutagenesis (mutation breeding) where an organism is exposed to radiation or chemicals to create a non-specific but stable change*”,如果只是简单地查词典知道*mutation breeding*的字面意思是诱变育种,但对其概念不甚了解,将会给笔者后续的翻译工作开展带来困难。笔者利用网络资源在百度搜索上输入“诱变育种”,其返回结果约4 120 000个(截至2020年1月19日),而笔者在百度知道上输入相同的标题,知识搜索引擎则优先检索系统已经累积收集的问题-回答知识库,直接根据知识库中已有问答返回237 027条回答(截至2020年1月19日),且知识库中的回答基本都围绕诱变育种展开,包括诱变育种的原理、概念、优缺点、过程及相关概念比对,可见知识搜索引擎返回的是经过整理的“知识”,而不是一堆“信息”,且能够确保

搜索知识的关联度,极大地节省了用户从海量网页中查找和筛选的时间和精力。

译者在拿到翻译文本时,往往都会遇到自己不熟悉的题材,而在翻译实践中,要真正透彻地理解原文,背景知识搜索的重要性不言而喻。知识搜索引擎汇聚网民专家的知识库,通过知识搜索引擎主题分类组成庞大的动态知识库,能更高效地筛选和过滤冗余的信息,返回结果更直接,更有针对性。

### (二) 一站式知识检索对翻译实践的启示

知识搜索引擎提供合理的检索结构和其他信息导航链接,实现一个知识点一站式处理,无须打开多个网站或多个搜索工具获取资源,极大地方便译者全方位搜索翻译所需的资源。“百度知道”为用户搜索提供了一系列相关的知识链接,如资讯、贴吧、视频、图片、文库等服务;“搜狗问问”的站内搜索提供了微信、知乎、名医、学术、英文、翻译、百科等相关链接。如此,信息搜索处理更全面,有效缩短了用户的决策路径。例如,笔者翻译的一篇中译英的转基因工程科普文——“将新的基因导入到植物中需要一个用于表达基因的特定区域的启动子”(译文:*Introducing new genes into plants requires a promoter specific to the area where the gene is to be expressed.*)。原文中提到“启动子”这一术语。笔者在搜狗问问上输入“启动子”,第一步:了解概念,搜索返回结果让笔者了解到转基因中“启动子”与“引物”两个容易混淆的概念的区别。

第二步,通过点击导航链接中的“视频”或“图片”,将文字表现的内容以视频图片呈现,结果更直观可视化。笔者从图片搜索中发现关于“启动子”术语及相关术语的中英文标注,如图1所示。

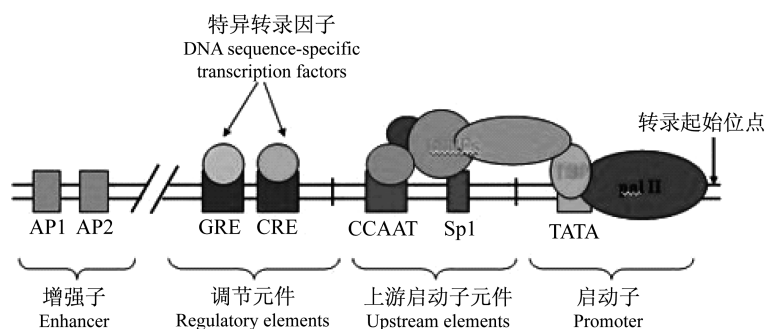


图1 “启动子”模型图

Fig.1 Model of promoter

第三步, 点击“英语”链接, 页面跳转到全球的综合百科网站, 如 <https://en.wikipedia.org/>、[www.dictionnaire.com](http://www.dictionnaire.com)、<https://www.addgene.org/>等, 这些英语网页可为笔者翻译提供参考。笔者在维基百科网站上搜索到有关“启动子”的英文术语的百科信息: In genetics, a promoter is a region of DNA that leads to initiation of transcription of a particular gene. Promoters are located near the transcription start sites of genes, upstream on the DNA (towards the 5' region of the sense strand). Promoters can be about 100-1 000 pairs long<sup>[6]</sup>. 如果译者想要了解所搜索知识点的更多详情, 可点击导航链接中的“微信”“知乎”等, 提高每个知识点的检全率和查准率, 有助于译者填补知识背景空白。

### (三) 线上互动问答对翻译实践的启示

知识搜索引擎最大的特点在于“人”的互动问答, 它不仅只是一个简单的搜索系统, 且可用自然语言对问题进行详细描述和标签设置, 充分激活人类的隐性知识, 能更快获得满意答案。2007年3月, 百度知道成为全球最大的中文互动问答社区, 截至2020年1月, 搜狗问问已为用户解决263 487 528个问题, 而新浪爱问提问后平均5分钟就有回答。

在翻译实践中, 如果译者在知识搜索引擎中搜索不到答案或对搜到的答案不满意, 则可通过提问的方式进一步获取有用信息。搜狗问问、新浪爱问、百度知道都设置了“我要提问”环节, 译者可通过提交问题寻求网络社区帮助以获取所需信息。这一举措弥补了普通搜索引擎的不足, 实现搜索引擎社区化, 拓宽了译者获取信息的来源和途径。笔者曾翻译过一篇中英文学作品, 文中写到:“曹雪芹

能写出《红楼梦》, 是因为他曾经有过这样的生活, 知道什么叫烈火烹油。”显然, 原文中“这样的生活”和“烈火烹油”存在呼应关系, 而“烈火烹油”则为理解的关键点。笔者首先搜了百度, 在百度百科中显示: 烈火烹油, 比喻声势气焰很盛。出处: 清·曹雪芹《红楼梦》第13回, “眼见不日又有一件非常的喜事, 真是烈火烹油、鲜花着锦之盛”。

Sperber & Wilson提出的关联理论认为人们的语言交际是一种认知活动, 交谈双方通过明示-推理过程理解话语的含义<sup>[7]</sup>。翻译也属于一种语言交际行为, 因此在关联理论的指导下, 译者在翻译实践中应根据上下文和语境寻求最佳关联, 以理解话语的含义, 从而准确传达原作者的意图。所以笔者认为, 若直接套用“气焰很盛”这一解析, 并不符合原文表达的意图, 于是在“百度知道”上使用自然语言提问: 烈火烹油、鲜花着锦在红楼梦中对应该什么? 得到热心网民回复(回复于2019-10-05): “该句出自《红楼梦》, 是秦可卿死前给凤姐托梦中说过的话, ‘鲜花着锦’即在鲜花上再加上锦绣, 鲜花更鲜艳; ‘烈火烹油’即在熊熊燃烧的烈火上再加上一瓢油, 火势会烧得更旺。两句连在一起, 意思同‘锦上添花’近似, 都是说好中更好。”

结合用户回复可知, “鲜花着锦, 烈火烹油”二者一起可释义为“锦上添花, 好中更好”的意思。因此笔者可将“烈火烹油”理解为曹雪芹家境富裕、养尊处优的意思, 遂将“曹雪芹能写出《红楼梦》, 是因为他曾经有过这样的生活, 知道什么叫烈火烹油”译为: Cao Xueqin, a well-known Chinese novelist living in Qing Dynasty (1616 - 1912), was able to write *A Dream of Red Mansions*, because he experienced such well-fixed life and knew what was glory and voluptuousness.

#### (四) 个性化知识库对翻译实践的启示

信息素养要求译者不仅能够高效搜索和利用信息,同时也应具备知识管理能力。国内知识搜索引擎如新浪爱问,不仅为用户提供提问、解答问题、搜索答案的功能,还设置了共享资料搜索、上传文档、文件夹管理、收藏夹的功能。因此用户可以充分利用这些功能,对答案、问题、下载的资料或自己上传的资料进行整理积累,形成用户的知识体系,建立个性化知识库。

如此一来,用户形成的知识库又可作为日后他人和自己信息搜索的结果,在信息搜索时可优先检索自己知识库里的答案,达到知识消费可循环,并及时实现知识库的维护和更新。

笔者认为在翻译实践过程中,相比传统搜索引擎,知识搜索引擎在准确性、针对性、信息降噪上更有优势,所呈现的答案是用户的知识体系,可帮助译者节省时间,提高效能。

笔者在下面总结了对国内三大知识搜索引擎的体验感受。

百度知道主题分类设置更详细。作为全球最大的中文问答平台,其知识库沉淀的知识容量更大;其次,百度知道对搜索到的答案还可进一步实现高级检索,如限定检索的时间范围;再者,百度知道在提问环节增加了“向TA提问”,可直接在用户回答的基础上,进一步向该用户“求助”。

搜狗问问在信息导航链接设置上更全面,查全率和查准率更高。搜狗问问设置的“英文”链接与全球各大综合百科知识网站合作,可直接实现搜索相关知识点的英语信息。因此,译者在翻译专业名词方面想要了解该词条的英文解析,使用搜狗问问更胜一筹。

新浪爱问在知识管理设置上更人性化,其设有“爱问资料共享”,方便译者按需收集、整理和利用知识。新浪爱问提供收藏夹,可将上传用户积累的知识分门别类,用户也可收藏在本地知识库中收集的问答及下载的资料,形成自己的个性化知识库。因此新浪爱问在帮助译者知识管理方面更胜一筹。另外,在提问环节中用户可同时将问题分享到

微博,提高知识网络中用户的参与度,更大化地汇聚网民智慧。

### 三、结束语

搜索能力在翻译中的应用已成为翻译行业越来越重视的领域和学者不断探讨的课题。如今人们对翻译行业不仅在翻译量上,而且在时间效能上的要求都越来越高。如何善用“搜商”获取准确高效的知識的能力已成为职业译者必备的能力。而知识搜索引擎相对于传统搜索引擎,增加了互动式问答和主题分类,在搜索上能更准确、更高效地满足译者个性化的信息搜索需求,大大提高了译者获取知识的时效。

知识搜索引擎尽管还处于探索发展阶段,在回答问题权威性、专业学术性以及检索来源上还存在很大发展空间,但一方面其对译者在翻译实践中掌握多元化搜索资源能力,从而对翻译实践服务和知识管理上具有很大的意义,另一方面,知识搜索引擎提供的知识服务也是将来必然的发展趋势,为翻译教学实践也提供了借鉴。各MTI高等院校可以基于学院需求,建立具备自己学院特色的知识搜索引擎,实现院系资源的共享服务和知识检索。

#### 参考文献:

- [1] 王少爽. 职业化时代译者信息素养研究:需求分析、概念阐释与模型构建[J]. 外语界,2017(1):55-63.
- [2] 陈沛. 搜商:人类的第三种能力[M]. 北京:清华大学出版社,2006.
- [3] 王华树,张成智. 大数据时代译者的搜索能力探究[J]. 中国科技翻译,2018,31(4):26-29.
- [4] 高雯雯,杨木锐. 知识搜索引擎探究[J]. 图书情报工作,2006,50(10):87-89.
- [5] 彭陶,马张华. Web2.0环境下的知识搜索引擎发展研究[J]. 图书馆学研究,2011(11):63-67,73.
- [6] 搜狗问问. 启动子[DB/OL]. (2019-05-20)[2019-10-20]. <https://baike.sogou.com/v1289402.htm?ch=ww.xqy.xgbk>.
- [7] SPERBER D, WILSON D. Relevance: Communication and Cognition[M]. Oxford:Blackwell,1986.

(编辑:朱渭波)