

# 小米NLP首席科学家王斌：NLP在落地、商业化方面进展慢是假象

From: 刘燕 AI前线 Today



策划 & 作者 | 刘燕

采访嘉宾 | 王斌，小米 NLP 首席科学家

编辑 | Linda

**AI 前线导读：**这两年，BERT 和 Transformer 的出现和运用推动 NLP 迎来了历史上的黄金时代。机器翻译、机器阅读、人机对话、机器写作等领域进步喜人。与此同时，NLP 在工业界的落地与应用也呈现百花齐放之态，在搜索、推荐、信息流、互联网金融、社交网络等领域都有其足迹。

近日，InfoQ 有幸采访了小米人工智能实验室主任、NLP 首席科学家王斌博士，他详细介绍了 NLP 在小米的智能手机、IOT、搜索、推荐、广告等多个不同场景中的应用实践。此外，王斌还对 NLP 的发展现状和趋势发表了自己的看法，他表示，NLP 存在的两个重要挑战，其一是如何获得大规模的高质量标注数据，其二是，如何得到轻量级的优质模型。王斌预测，NLP 技术今天已经似乎真正触摸到“语义”的门槛，未来的 NLP 应该会真正跨过这道门槛，实现自然语言理解的突破。

以下是此次访谈的全部内容，如果你对该话题有进一步的兴趣，王斌还将在即将召开的 AIcon 全球软件开发大会（北京站）2019 上，带来题为《演讲：NLP 在小米的探索与实

践》的主题演讲》。

更多优质内容请关注微信公众号“AI 前线” (ID: ai-front)

InfoQ：很高兴采访您，首先请您做一下自我介绍，目前主要负责哪些工作？

**王斌：**我之前在中科院从事了 20 多年信息检索、NLP 相关的研究和开发工作，曾是中科院研究员、博导和中国科学院大学教授。去年 8 月全职加入小米，任 NLP 首席科学家。今年 3 月开始任人工智能实验室主任，领导视觉、声学、语音、知识图谱、NLP 和机器学习等 6 个团队，研发核心技术，全面赋能公司业务。

InfoQ：目前，小米在 NLP 领域的总体布局是怎样的？

**王斌：**目前小米 NLP 布局主要包括基础 NLP 能力建设和 NLP 业务应用两部分。前者打造云端、本地、设备端三端一体的 MiNLP 平台，后者是重点支持多个核心业务应用，包括智能问答、人机对话、机器翻译、内容过滤、搜索推荐等等。

InfoQ：自您去年 8 月份加入小米这一年多来，小爱同学的对话能力有了哪些提升？

**王斌：**小爱的对话理解能力正在从单轮的理解往连续对话能力上延伸，可以根据上下文理解用户的意图，不需要用户每次交互都是完整的表达。尤其是实现全双工交互后，用户跟小爱交互的习惯也在发生改变，表达越来越自然，更口语化。

InfoQ：上个月亮相的小爱同学 3.0 最大的优势功能是做到了自然连续语音对话（全双工），为实现该功能取得了哪些关键性的技术突破及背后运用的核心算法是什么？

**王斌：**上述能力涉及到整个语音交互链条中各个环节的技术，包括前端降噪和回音消除、连续语音识别、无效音识别、对话管理、上下文理解、回复上的节奏控制等等。

InfoQ：从行业来看，全双工到底是不是一个新技术？近期很多公司都在该技术上有突破，这对于语音交互和 NLP 领域来说意味着什么？

**王斌：**全双工这个词不是人工智能或者说语音 AI 时代才出现的，并不特指某一种技术。它是通信传输的一个术语，指可以同时进行信号的双向传输。那在我们这个产品里，其实就是指一种更像人与人之间在沟通交流时，一种实时的、双向的交互方式。

两年的时间，小爱同学主要的定位是“一句话的事儿”，目的是快速满足用户的需求。但随着用户对智能助理使用越来越熟练，以及小爱同学能力的延伸，会发现用户对小爱同学有了更多的预期，会希望小爱可以处理更多复杂任务，这种就很难用一句话来说清楚。

自然连续语音对话会使得对话更自然、更人性化，会进一步提高对话的体验。

**InfoQ：**在机器翻译方面，小米近期有哪些创新成果？

**王斌：**今年公司发布了小爱老师这样一款硬件，提供了语言学习和口语翻译等功能，其中一个核心技术是多语言离线机器翻译。受端上设备上存储空间和计算能力的限制，我们进行了相关的技术实践和探索，从而可以在端上提供高质量低延迟低功耗的离线神经机器翻译服务。

另外，在手机上，MIUI 在全球具有庞大的用户规模，其中大量用户有跨语言获取信息的强烈需求，包括扫一扫拍照翻译，小爱同学语音翻译，米聊对话翻译和浏览器网页翻译等场景，这些场景涉及到多领域、多语种以及不同模态下的翻译需求，针对这些场景，我们也进行了各种技术优化和尝试，开发了在线的机器翻译系统。

**InfoQ：**随着神经网络翻译技术的发展，机器翻译质量有了很大提升，业内频有论调称机器翻译将取代人工翻译，您觉得二者的关系是怎样的？

**王斌：**个人认为机器翻译和人工翻译相辅相成。随着神经机器翻译技术的快速发展，对于资源丰富的语种以及专利等垂直领域，机器翻译的质量也在快速提升，在对翻译质量要求不苛刻，以及人工翻译无法快速满足海量文本翻译需求等场景，机器翻译得到广泛使用，机器翻译 + 译后编辑等人机结合模式也成为语言翻译行业的重要实践。

长期来看，机器翻译的快速发展会对一些低水平人工译员带来影响，但同时，机器翻译也有大量不适用的场景，例如对“信，达，雅”有很高要求的文学翻译需求是高水平译员大有可为的地方，同时，机器翻译的质量提高无法离开人工译员产生的大规模高质量双语数据。

因此，可以预期，机器翻译替代不了人工翻译，两者应该是长期相辅相成，互相促进的。另一方面，在很多场景下，机器翻译也可以看作是人工翻译的自然延展。比如跨语言搜索、跨语言电商、跨语言信息流等领域，人工翻译资源不足以完全覆盖这些领域，机器翻译可以看作是人工翻译的补充，与人工翻译一起为用户提供更完整的跨语言服务。

**InfoQ：**当下，机器翻译技术的发展现状和未来的趋势如何？

**王斌：**当前，在数据较为丰富的翻译场景下，神经机器翻译质量已经显著得到提升，神经机器翻译模型基本替代了统计机器翻译模型，已经成为工业界主流的技术方案，得到广泛应用。

但同时，对于一些深层次的理论和应用层面问题，目前神经机器翻译依然缺乏有效的解决方案，包括理论层面的可解释性以及应用层面的诸多问题 (如篇章建模和翻译、语音图片翻译的鲁棒性、低功耗翻译模型、语言领域自适应、小语种翻译等等)。未来对这些问题的解决，可以得到更精确、更鲁棒、领域自适应的优质机器翻译系统。

**InfoQ：**NLP 技术在小米落地“土壤”肥沃，主要的运用场景在硬件上，智能手机、IOT、搜索、推荐、广告等场景都有运用，可否详细介绍一下具体的运用？

**王斌：**小米的很多应用场景中都包含内容处理，应用了很多 NLP 技术。比如小爱同学中的查询意图理解中使用了文本分类、分词、命名实体识别、句法分析等技术，通过文本分类技术将查询意图分到意图体系，通过分词和命名实体识别技术从文本中提取关键字段，通过句法分析技术对查询进行结构化分析等等。在智能问答和对话中，使用实体链接、相似度计算、文本生成等技术进行问题理解、搜索和对话生成。在搜索和推荐场景，同样使用了大量的 NLP 技术。

**InfoQ：**可否用一组数据说明小米 NLP 在解决复杂运用场景时的效果和能力，有没有一些场景是难以完美解决的？

**王斌：**很多 NLP 的场景都非常复杂，超过普通人的想象，需要综合应用一系列 NLP 技术。就以刚刚提到的查询意图理解这一个模块为例，就用到了数十种 NLP 技术，很多技术做得非常非常细，之间还要组合和衔接才能形成有机的整体。即使是这样，该模块从技术上来看也没有得到完美解决。实际上，NLP 的很多场景都几乎不可能从技术上完美解决，只能在实际中通过技术能力和产品设计的平衡来提高用户的综合体验。

**InfoQ：**从行业应用的角度看，目前自然语言处理技术发展到什么程度了？

**王斌：**这些年 NLP 技术的发展很快，有些任务已经取得了大幅度进步。比如机器翻译、机器阅读、人机对话、机器写作等领域都取得了长足的发展。

随着 Transformer 和 Bert 的提出，NLP 技术这几年得到了较大发展。这些技术除了在学术界得到广泛推崇之外，在工业界也得到了实际验证。小米公司也在广泛地使用这些技术，来进一步提高 NLP 的实际应用效果。个人觉得 NLP 技术的进步还是很大的，当然由于 NLP 本身的一些特点，其实际水平一直不低，在实际中也早已落地，因此看上去并没有获得和其他某些领域一样的大幅度增长。

**InfoQ：**有观点认为，NLP 和语音技术这几年在技术和运用上都没有实质性的进步，您认为自然语言处理在发展过程中还存在哪些挑战，如何突破现有的瓶颈？

**王斌：**NLP 的发展还存在很多挑战，很多专家都有过总结。这里提我个人认为最重要的两点。第一是标注数据问题。当前主流方法的效果取决于标注数据的规模和质量。获得大规模的高质量标注数据永远是个难题。要解决这个问题，一种可能的方法是通过自动标注或者半自动标注或者自然标注来扩大标注的数据量。另一种可能的方法是通过弱监督或半监督方法来充分利用大规模的未标注数据。

第二是轻量级优质模型问题。当前的主流模型需要消耗大量资源进行训练，这种趋势目前看来有增无减。如何得到轻量级的优质模型是一个挑战性问题。可能的方法包括对现有模型的裁剪甚至另辟新路提出新的模型。

**InfoQ：**相较于计算机视觉，自然语言处理技术在落地、商业化方面进展较慢的原因是什么？

**王斌：**个人认为，深度学习的发展对计算机视觉的影响远比对 NLP 大。原来很多计算机视觉任务仅仅停留在学术研究层面，难以落地应用。深度学习使得很多计算机视觉任务突然具备了落地的可能，所以这些年来计算机视觉领域在商业化和落地方面取得了一些明显的进展。

相对而言，NLP 技术一直在工业界平稳落地，不论是搜索、推荐、信息流、互联网金融还是社交网络，NLP 技术都在其中起着十分重要的作用，正是因为应用已经非常广，加上 NLP 技术本身处于底层支撑地位，给人的感官不明显，所以给人造成落地、商业化进展很慢的假象。

**InfoQ：**小米在 NLP 领域商业化方面的进展如何，稳定的赚钱模式是否已经确立？

**王斌：**公司的 NLP 需求非常旺盛，目前小米 NLP 的技术主要给公司的业务和产品服务，未来也会考虑对外提供服务。

**InfoQ：**此前 InfoQ 接触到的一位专家曾预判，NLP 目前处在历史上最好的发展时期，迎来了发展的黄金时代，对此您怎么看？

**王斌：**个人同样认为，NLP 迎来了历史上的黄金时代。

可以从 2 个角度来看：第一，大家都知道 AI 的成功大致取决于数据 + 算力 + 算法。作为 AI 领域最活跃的方向之一，相对于其他一些领域，NLP 数据 (包含已标注和未标注数据) 的积累更加丰富。大量自然语言数据的存在，使得 NLP 的数据基础非常坚实。



此外，近年来 BERT 等预训练模型的出现大大促进了 NLP 领域的发展，再加上算力和算法的进步，NLP 确实具备了成功的前提。第二，NLP 的需求一直非常旺盛，学界和工业界形成良性循环。如果说很多研究是兴趣探索驱动的话，那么当前 NLP 的很多研究都是需求驱动的。工业界有大量的 NLP 需求，学术界对这些需求对应的自然语言处理任务进行研究，研究的成果很多可以应用到工业界，通过这种不断循环迭代，NLP 的研究得到了快速的发展。

**InfoQ：**您认为，未来 NLP 将会是怎样的发展趋势？

**王斌：**经过这些年的发展，我有两个个人的判断。

第一，虽然“语义”理解在学术界说了很多年，但是个人隐约感到 NLP 技术今天已经似乎真正触摸到“语义”的门槛。多层次、多粒度、语言无关的语义表示正在向我们走来。未来的 NLP 应该会真正跨过这道门槛，实现自然语言理解的突破。

第二，自动生成技术可能会有大的突破。通常把 NLP 分成自然语言理解和自然语言生成两大部分。前者的工作占据当前 NLP 的主要部分，但是随着 NLP 技术的发展，近年来自然语言生成技术得到了较大发展，已经在实际中逐渐体现出其重要价值。我相信，下一步 NLP 的发展中，自然语言生成也会成为 NLP 的重要组成部分。

## 嘉宾介绍

王斌博士，小米人工智能实验室主任，NLP 首席科学家。中国中文信息学会理事，计算语言学、信息检索、社交媒体处理、语言与知识计算等专委会委员及《中文信息学报》编委，中国计算机学会中文信息处理专业委员会委员。加入小米之前，在中科院计算所、信工所从事自然语言处理和信息检索相关的科研工作，任中科院博导、研究员，中国科学院大学 A 类岗位教授。发表学术论文近 150 篇，出版《信息检索导论》、《大数据：互联网大规模数据挖掘与分布式处理》、《机器学习实战》等译著 6 部，科研成果获得国家科技进步二等奖 2 项，省部级科技奖 3 项。是 SIGIR、CIKM、ACL、IJCAI、AAAI、WWW 等会议的程序委员会委员。

## 福利时刻

王斌老师将在 11 月 21-22 的 AICon 大会上带来更多关于 NLP 的精彩分享，想要跟王老师深入交流的，不要错过面基的机会。另外，大会还有关于机器学习、NLP、搜索推荐、智能