

# 建国际视野，做中国脊梁

■ 王小川

今天的演讲，我想讲一下搜狗公司的核心价值观——构建国际视野，做中国的脊梁。我将和大家分享搜狗在这方面的三个故事和一些经历，以及在智能 AI 方面的构想和行动。

今天，我们正在从信息文明走向智能文明。我分享的第一个故事是关于中文信息化，主要围绕搜狗输入法（如何解决中文信息化难题）来展开。

我中学时跟科学院院士做几何定理的机器证明，获得了由杨振宁先生颁发的一个奖项。李岚清副总理到学校视察，我当时作为小同学，对国家领导人将会提出的问题有很高的期待。后来李岚清问了我一个问题：“小同学，你一分钟打多少字？”当时我非常沮丧。我想，国家领导人提出的问题怎么才这个高度？我长大之后才开始理解，假如他问一个农民，会问“你一亩地产多少粮食”，这和问我“一个分钟打多少字”是一样的，一个代表物质文明的基础生产能力，一个代表信息化的基础生产能力。在八、九十年代的时候，中国当时的环境下学习电脑就是学习打字，会打字就代表了中国的信息化走出了第一步，所以当时我们就说应该做一个好的输入法。今天，我们实现了这样的诺言。

输入法在中国的今天是不可缺的产品。搜狗初代输入法的原理是把搜索引擎放在网上，把当时生产的 100 亿个网页都抓取下来，然后做大数据分析，理解中国人的单词和语法规则，由此塑造了一个超越前代的输入法，这个输入法能让我们打字更加省力。我们也是第一个把大数据、云计算，包括众包的概念用到输入法上的产品。现在有 99% 的中国人在 PC（个人电脑）上都在用搜狗输入法。我们也有很强的语音输入的能力，从打字开始走向声音的输入，有中国最大的在线的语音识别软件。

目前我们在研发下一步的输入法。输入法的下一步是不是做语音？我认为语音还不够。去年搜狗上市的时候，我在一天内收到了 3000 条祝福的短信，用语音去回复也是不够的。如果集体回复，对学长、老师、同事应该给予不同的回复。这样的情况下，我们期待未来的输入是智能的。让机器开始理解语言，成为人的助手。这是我想讲的第一个故事，搜狗输入法未来的道路。

第二个故事是关于机器翻译，（关于搜狗做 AI 翻译、用中文连接全世界的新思路）我曾经引用《圣经》里通天塔的故事来阐述翻译（其重要

性)。上帝跟人类签了一个和约，要人类在地球上好好生活，但是人类不满足，所以就造一个通天塔，去看看天上是什么样子。上帝很不开心，认为人类藐视他的权威，因此让不同的人说不同的语言，大家就无法进行沟通，通天塔就修不成了。这个故事反过来就强调了跨语言沟通交流的重要性。尤其在中国，这是一个很重要的事情。语言的隔阂依然阻碍着人类的交流与融合。

当提到翻译的时候，大家会认为谷歌或者其他美国公司的概念、技术是最好的。在基础研究领域，美国的深度学习比我们做得好很多，但是到了应用领域，就像输入法一样，中国人比美国人做得更好，因为我们更有痛点，有实实在在的语言的障碍，我们做翻译会更努力。一方面，我们想获得全球信息的时候需要用英文，但其实大多数中国人并不能学好英文，跨语言交流仍属于少数精英人群。另一方面，在国际场合和技术领域，比如联合国大会上要用英文发言。语言意味着话语权，虽然中国有十几亿人，但是这个世界的主流是英语。用别人的语言去沟通，其实是走进另外一种文化、另外一种文明。我们能不能把自己的文明带出去？所以我觉得在今天的中国，做翻译有重大的意义。

历史上曾出现过几次翻译高峰。例如在公元8世纪发生了阿拉伯百年翻译运动。当时阿拉伯帝国统治了欧亚非的大陆很多地方，所以他们希望通过翻译了解当地的文化，实现更好的统治。他们组织了很多专家用了100年时间做翻译，前



王小川演讲

后带来了500年的黄金盛世。中国历史上也有四次翻译高峰运动，每一次都带来了很大的影响。第一次是东汉末年至唐宋时期，当时最大的翻译是唐玄奘西天取经，带回的六百多部佛经中，19年只翻译了10%多一点（七十多部），即便这样，也极大地影响了中华文明。第二次是明清两代的科技翻译，《几何原本》等重要科技著作的翻译极大地帮助我们理解西方的科学文化。第三次是鸦片战争至五四运动时期的西学翻译，其中最重要的著作是《共产主义宣言》，引起了中国社会的重大变化。第四次，改革开放之后又迎来了一个翻译高潮。

今天，怎么能够让更多的大众、老百姓能享受国外的信息是我们的一个机会。我们做的两款AI翻译产品研究了最顶尖的翻译技术。一个是全球搜索，今天95%的全球信息是英文所写的，所以我们的搜索引擎启动了一个全球搜索版本，用中文去搜索全球信息，用中文做阅读，使大部分中国人有机会接触到更多的世

界文明。另外一个就是同声传译的系统，我们在很多比赛和会议等场合都使用了这样的系统。我们希望通过这样的技术辅助信息沟通，使得东西文化有更好的融合。

第三个故事关于医疗，搜狗明医搜索。今天很多人对互联网医疗有误解，我和很多医院的院长、副院长、主治医生沟通过，他们认为医疗就是医院的事情，诊断和医疗需要在医院进行，医院之外的互联网跟医院是没有关系的。而我认为病人在什么地方，我们的服务就在什么地方。病人不在医院里面，病人其实是遍布在城市里、农村里，医院并不覆盖，而这是互联网的机会。我问过投资人代表，他们也认为互联网医疗是一个伪命题，因为现在有很多投资在互联网上做分诊、问诊但都失败了。而我不这么认为。相对于国内“好大夫”“春雨医生”“微医”这些品牌，它们一天的用户数是 20 万，而搜狗搜索每一天有超过 3000 万次的健康相关的检索，百度超过 1 亿次。这个数字说明不管今天的搜索引擎做得如何糟糕，网民依然需要用搜索引擎去解决医疗的问题。我们总不能摆出这样一个结果，说“对不起，我不回答你的问题”。但网络中健康资讯存在错误、虚假，魏则西事件表明虚假医疗信息已经给用户带来了伤害。所以我认为，我们有义务在搜索引擎里提供更好的医疗信息内容。今天我们是能够通过引擎来给大家提供权威而优质的信息的，搜狗搜索中第一条一定是全面的信息而不是广告。

三个故事讲完后，我们来思考怎么真正从信息文明走向智能文明。在信息文明里，我们有两个



搜狗的 AI 理念

核心的产品，一个是输入法，一个是搜索引擎，一个帮你表达信息，一个帮你获取信息。人类表达、获取信息核心的关键词是语言，这也是人类区别于动物的关键点，人类通过语言去抽象这个世界。今天机器处理语言是 AI 领域最大的挑战。从人类发展的角度，掌握一种语言，就是理解一种灵魂，全世界有七千多语言要让机器去学习。语言在 AI 时代里称为人工智能皇冠上的明珠。我们天天都在跟语言打交道，一个是 output（输出），一个是 input（输入）。

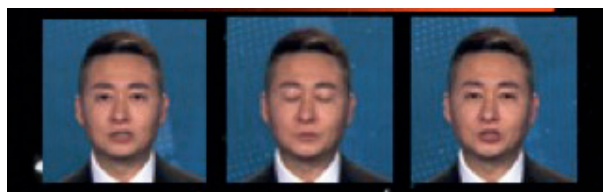
语言可以分为两个领域（如上图），左边是沟通，即让机器掌握、表达和理解人的语言。机器可以通过语音、图像识别人的语言，例如唇语，根据嘴形的运动判断出人在说什么。反过来，机器也能通过声音合成语言，甚至能合成表情，例如虚拟演员也是我们的技术之一。

语言不止是在于沟通，它还承载着思考和计算，因此右边的部分有对话、翻译、问答。一方面讲沟通，一方面讲智慧，这是让机器走进智能时代，走向智能语言的两个核心突破。

给大家看几个我们的小 demo（演示）。第一个是声音迁移。给机器 14 分钟的语音训练（比如王小川的声音），然后配上一首歌曲，就可以

生成王小川唱的这首歌曲。第二个是手写字体迁移，当你写出一些汉字后，机器学习你的字体，生成你的个性书法特征。第三个是人脸迁移，涉及到人脸识别、三维人脸拟合、表情控制，可以做出一个虚拟的形象，实现面部实时动态变化。能模仿你的声音、模仿你的表情，把一个人变成另外一个人。

接下来还有自然语言交互技术。机器不仅能够生成，还能够识别和个性化合成。这是搜狗虚拟主播的展示，视频中的主播是虚拟的，我们只要输一段文字，所有的语音和唇形都是合成的，和真人唇形是能对应上的。



搜狗虚拟主播

下面是问答领域的搜狗汪仔，它可以参加由组织出题、选手作答的比赛。这个比赛和我们的理念是一致的，因为搜索引擎最终要走向问答式，用户提出问题，它给出答案。今天的搜索引擎是输入一个关键词后给出十几条结果，不是自然的问答过程，因为机器听不懂人的问题。随着 AI 技术慢慢发展，现在我们可以慢慢做到把人类提出的问题理解。

今年还有一个工作是阅读理解，从众多网页里寻找用户需要的答案。IBM 在 2005 年就发布了类似的产品，但他们的产品是分领域的，提问的时候要人工指定这个答案是明星、地点还是某个门类；我们提供的是通用的产品，并且我们的



以语言为核心的 AI 战略布局

引擎可以用语音提问。IBM 系统是用公用数据库给出答案，而我们的搜集引擎是在整个互联网去搜索，从页面寻找答案。这个系统后来运用到今年年初的一个活动“直播打比赛”。比如网站出 12 道题目，奖金 100 万，如果 12 道题都答对了可以拿到 100 万，但是如果说 100 万人都答对了，每个人只能拿到 1 块钱。我们推出的“直播答题助手”能帮用户答题。后来发现人和机器相比取不了胜，出题网站就把题目改得很长、很复杂，我们的工程师也修改程序，能做到题目长到人读不了的情况下还能回答。不过对于事实类的问题机器能做得很好，但是对于含有推理、判断的题目，机器学习目前还做不了。

从信息文明走向智能文明的过程中，当机器通过语言掌握沟通的能力、自然交互的能力，让机器掌握智慧，掌握基于语言的计算能力，最终机器会成为人类的助手，它会和你沟通，能帮你回答问题，能帮你和他人互动，这是我们认为未来机器发展的方向，我称它为虚拟个人助理。这是我们在今天这个时代中最重要的一项功课。📖