

以下是采访摘要：

问：虚拟主播的推出，会不会让很多人失业？

王小川：首先 AI 技术分为感知技术和认知技术。感知是有声音的、有图像的，感知技术方向上，机器基本有机会可以做到跟人一样好。但是在认知技术方向上，机器背后的推理、知识、思考，以语言为核心的逻辑性的东西，机器的处理能力是有限的。这种情况下，涉及到人的高级活动的时候，机器现在是做不到的。

问：以后有可能做到吗？

王小川：现在没有这个技术在。如果是在垂直领域里，特别细分的特定领域，像法律或者医疗，面越窄，机器更有机会做到更接近人类。不能把 AI 当成大的词儿，AI 里面有很多内涵。在这种情况下，把 AI 上升到跟人们认知相关的事情，机器在里面只能做辅助，是没有办法取代人类的。但如果只是听个声音，看个图像，现在像 Face++ 或者商汤做的事情，机器就已经能取代一部分感知上的事情了。对于人类的高级活动，现在机器取代人类是没戏的。高级活动叫认知，低级活动叫感知，目前的机器可以做到感知，这是先给大的框。

问：您怎么看待 AI 合成主播，人类跟机器是怎样的关系？取代还是配合？

王小川：“AI 合成主播”这项技术里面涉及三个合成：声音合成、表情合成、唇型合成，特别是唇型合成。因为本身虚拟这个词，画个卡通画也叫虚拟，如果叫虚拟主播的话别人容易较真，说这个东西怎么叫虚拟，所以我们叫 AI 合成主

播。AI 合成主播和人类的关系是取代还是配合这个问题，如果只是感知问题，不涉及到后面的语言组织、撰写深度稿件的话，它是可以去跟真人接近的。

问：搜狗的 AI 合成主播跟微软小冰做的主播有什么差异？

王小川：搜狗的 AI 合成主播是真人的形象，小冰是虚拟形象，声音跟真人不一样，也没有表情和唇型的变化。AI 合成主播真的是合成之后，能够起到真人的替身作用。

问：这个技术除了在主播领域，还有哪些其他场景的应用？

王小川：比如说现在我们跟凯叔讲故事谈，以前是凯叔去讲故事，以后可能就会变成你的爸爸妈妈给你讲故事。在我们大的理念下，主播是其中一块，之后要把它个性化，变成其他的人。

问：在语言这块，门槛并不高？

王小川：语言这块门槛挺高，高德地图合成林志玲的声音的时候，林志玲是读了非常多句话的，不是固定的“向左转向右转”的句子。搜狗现在只要求十几分钟的数据量就够了，可以用很小的数据合成一个人的声音。

问：为什么？是有什么算法的突破吗？

王小川：所谓小数据其实也是离不开大数据的，机器要看很多声音，然后找到这个人的声音特征跟别人有什么不一样，所以小也是大。我们以前说婴儿学习很快，看一张图片就会了，其实他是看了很多图片之后再看一张图片才会。所以对特定领域来说，数据越小越好，但是要掌握足够多的通用领域的的数据。因此这里涉及到的技术，既要有海量声音的训练，同时能做到用更少的声音数据训练特定人的声音，这是有技术壁垒的。

问：合成主播商业化的节奏会不会比之前做的其他 AI 落地的项目要快一点？

王小川：最快的话，第一是翻译，翻译是刚需。不只是翻译宝，搜狗搜索支持用中文搜索全球信息，用中文阅读全球信息，这是应用了翻译技术的。翻译是搜狗的使命之一，输入法输入中文转成外文，搜索能搜外文转成中文。搜狗本身是做信息桥梁的公司，所以翻译是很重要的，这也是最快落地的技术。首先是语音、图像，然后是翻译，翻译之后是分身，训练一个人的数据之后，帮助他去做表达。最后是问答，就是个人助理，帮你回答问题。对于消费者而言，这是技术演进的路线。

问：刚才提到大数据，搜狗做通用训练的时候，语音素材是怎么拿到的，是平常语音输入的时候的语音吗，还是怎么样的？

王小川：我们有很多合作，很多标注。比如现在我们跟喜马拉雅有合作，他们读的东西我们可以拿过来用，尽可能多的采集不同的语音。今天变成 AI 技术的研讨会了。

问：财报中广告点击业务占比比重开始下滑，是因为流量获取成本太高了吗？

王小川：两个层面，用户端要做权威的内容，使得信息真实，甚至让医生参与到中间的内容创造，我们先给内容，再给广告，现在搜狗的医疗用户查询已经占超过 9.5%。我们用户端做得更努力，商业化是更少。

问：你说其实不想医疗广告做大，为什么？

王小川：我是希望产生差异化。医疗上，用户更需要一个内容或者引导到一个权威的医院去，现在这样的模式，它的天花板是很低的。

问：天花板你是指？

王小川：广告收入的天花板，往下不是特别可持续的发展模式，这个东西有限，我们就压低这块收入，做更好的内容，未来带到三甲医院或者社区医院。

问：您觉得这种路线会给我们在营收上或者商业化上带来更好的回报吗？

王小川：首先不是让医疗广告更多，反而是争取到医疗广告更少，医疗广告收入更低，但是我们最终的目的是用户更爱用搜狗搜索，使整个流量变大，用户变大，提升收入。

问：除了医疗以外，搜狗也在做法律相关的内容搜索，不知道我们对这种领域的选择标准以及下一步扩展垂直搜索的规划是什么？

王小川：首先最大的是医疗，法律是中间有明显的知识结构化能力在，这个知识边界是相对有权威和规范的。在这种情况下，我们可能去做一个选择。但是医疗还是远远优于法律的。

问：还会扩展其他领域吗？

王小川：其他领域，百科我们还是想努力能把它做得更加权威和真实。总之有些领域，互联网的信息不够好，我们希望能够在里面能用新的方法做好，用 AI 技术或者用不同的方式

问：有没有考虑相关多元化或者更多 AI 的软件或者科研或者硬件？

王小川：不考虑，我觉得我们已经张得很开了，我们在信息文明时代核心就是对知识或者对语言的理解。做 AI 我认为要满足几条，第一个有场景，有数据，所以你做 AI，如果你没场景、没有数据只做技术，这件事情是很困难的。我们数据和场景是在用户表达和信息获取里，在输入和搜索里，所以我们围绕场景做这件事情。