**「采访目的」**如何做好科学研究  
**「采访时间」**2021年4月9日  
**「采访地点」**北京邮电大学PRIS模式识别实验室  
**「采访对象」**陈光老师 (@爱可可-爱生活[1])  
  
**「小七：」**我们知道学术界存在着很强的马太效应，绝大部分论文的引用量都很小，我们平时阅读的论文大多都是有一定影响力的论文，而一个平凡的科研人员辛辛苦苦写出来的论文可能并没有多少人关注，而且大部分论文解决的问题都很小，似乎创造不了什么价值，甚至没什么意义，陈老师在做科研的时候会有这样的失落感和渺小感吗？🤔  
**「陈光：」** 首先我觉得在无数种做研究的出发点中，**「兴趣应该是最重要的」**。回到最开始的那个问题上，就是你为什么要发论文，对吧？你是单纯为了让更多人看到自己，为了让大家引用自己的论文，为了提升影响力，还是说，就是为了自己？让自己觉得确实是有需要总结一下的工作，确实觉得自己在某些方面某些领域有所贡献，并且需要把这些贡献流传下去，我觉得这些才是更好的出发点。  
引用这个事呢，当然论文的影响力肯定越大越好，越多人关注越好，但是**「这并不意味着没有足够的影响就不是好的研究」**，我觉得这个问题是非常重要的。好的研究首先是值得关注的，值得关注和最后真正有没有被关注，我觉得是不一样的，你看现在每天都有大量的论文被发表，成果太多，并不是说没有足够多引用的，或者没有引起足够关注的那些论文就不够好，其实是有一些好论文的。这些好论文没有受到足够关注的原因有很多，一方面是作者所在的团队，或者所在的学校，或者本人知名度可能没有那么大，另一方面的话，其实就是运气，如果你像我一样每天选读这些新发表的论文，你会发现其实运气也是很重要的。就是你某一天把论文发在arXiv上面是不是能引起足够多人的关注，可能也取决于跟你同一天发表的论文怎么样，如果这一天你尽碰上Google、Facebook的论文，还有Berkeley、MIT这类名校的论文，你的论文可能就不会受到足够的关注了，对吧？这就是运气的成分。🙂  
但是不管怎么样，不管是说运气也好，还是说出身也好，这些都是外在的，我觉得，最关键的问题是你觉得你写的东西是不是一个值得写出来的东西，郭军老师(陈光老师的导师)说过一句话很重要，就是「**把东西写下来是很重要的，但是比这个更重要的是你写的东西值不值得写」**。你看现在很多论文发表出来，之所以质量不行，是因为他们并不是为了纪念一个值得写的东西，或者写出来不是为了让大家去借鉴，可能更多的是要完成一个任务，或者说是为了毕业，或者是为了职称。一方面这给整个圈子添加了噪声，给大家加大了筛选的难度，同时作者自己其实可能也并不希望有太多的引用。😅  
所以我觉得这个问题，从外面来看，我们还是要呼吁大家不要仅仅去关注那些之前有过好工作的这些人和机构，也要**「更多的去关注这些相对默默无闻的人」**，他们其实也会有一些很好的工作，这里面其实就有很多被埋没的。我记得有一篇论文是专门研究论文引用这块的，其实有很多影响的因素，比方说像国外的推特，你发论文以后是不是发twitter推广了，或者是不是引用了某一个名人，并且被他关注到了，还有像现在很多会议的审稿人一样，如果你引用了他的论文的话，他就愿意帮你宣传一下你的工作，因为你的工作好，其实也间接的说明他的工作好。从里往外看，我觉得首先还是要端正思想，简单来说就是写值得写的东西，做值得做的工作，做有价值的事，不要仅仅为了发论文而发论文。  
  
**「小七：」**那把问题再放大一点，陈老师觉得现在科学的发展是靠一群天才推进的，还是广大科研人员通过努力共同推进的？很多研究生都会觉得自己不够聪明，也没有那么多好的想法，搞不出什么很有影响力的研究，觉得科研是牛人才能做的事，从而对科研本身这件事就失去了兴趣。您觉得我们应该怎样摆脱这样的心态，发现并认可自己的价值？🤔  
**「陈光：」**首先做研究还是要沉得住气，说到牛人和普通人区别，我觉得**「首先很多牛人并不是说他比别人更聪明，他们厉害的点更多在于他们有更宽广的视野」**。其实整个研究领域的进步，都是大家一点一点的添砖加瓦逐步推进的，即使一些聪明人和牛人，他可能想到一个好的点子，比方说Kaming He，他提出的很多点子并不是说靠头脑风暴一下子就能想出来的，其实他也是在前人大量的工作的基础上想出来的，像Bengio、LeCun这些人，这些人一方面对于现有的进展有足够的视野，他们知道目前做到了什么程度，能做到什么程度，还有哪些坑，哪些方法还存在哪些问题，或者说哪一类的方法普遍还存在哪些问题，概括来说就是他们有足够多的经验，这种经验，对于他们提的一些方向也好，还是提的一些新的方法也好，还是一些改进也好，是起到了非常大的作用的。  
所以从这个角度上来讲，**「我觉得我们所有人应该追求的是对自己的一种超越」**，而不是说你觉得自己不如某些人聪明，或者不如他们视野宽广，实际上我觉得这些差距你是可以去追平的，或者没有差那么多，并不是说人家是少年班就比你要强多少，视野就要好多少，他可能是比你快一点，但是你有足够的时间，还是可以去拉平差距的。  
但是更重要的是，我觉得还是回到刚才的出发点上，做研究首先不能急功近利，我觉得怎么去看待别人的成果也是很重要的，别人的成果不是因为这个人在学术成就上超越了你，更重要的是他在这个领域里面给我们带来了一些新的思考，带来了一些新的角度，找到了一些新的方向，或者更明确了一些东西，这些我觉得才算是对于整个学术圈的贡献。而这就是咱们经常讲的，就是进学术圈做研究注定是要孤独的，**「大部分人可能都只能是昙花一现，但是别人绽放的那一刹那，你更多的应该是为别人感到高兴」**，这个不是为了他出名了而为他感到高兴，而是说为你所在的领域又有了一个进步，你又有了更多的你可以去提升的素材，你又有了一个更好的肩膀而高兴，其实就是为自己高兴，所以我觉得从心态上来讲，这才是一个正常的心态。🙂  
你首先要明确自己一定会走上一条孤独寂寞的路，但是同时有这么多人陪着你，而且不断的有一些新的能够帮助你去提升的这样一些成果和一些人出现，所以你不是一个人在战斗，再一个只要你能够持续的去推进，持续的认真的在一个足够小的点上去发力，我相信是一定会有一个让你自己觉得满意的一个成果，**「咱们说让自己满意的成果并不是说非得要发一个什么样的级别的论文，或者说非得让多少人去引用，我觉得是说你有没有说服自己做的所有的这些工作是有价值的，只要你觉得它是有价值的就可以了。」**  
人类几百年以来，科学研究最大的驱动力其实就在于每个人对于自己探索真理，追求科学进步的那一点点，而不是在于有多少人认可你的工作。先不用说你现在的工作有没有进展，即使有进展，后人也会很快把你取代或推翻，认为你的想法是一个愚蠢的，或者甚至于是错误的想法。但重点是，你的最大的收获是在于你在你所坚信的这条方向上前进了，所以动力更多的是来自于内在，而不是来自于外在，**「如果你靠外在的驱动，你可能走不了多远，你只能是从内部找到你的方向，找到你的动力，你才能走得足够的远」。  
「小七：」**陈老师的这番回答其实也把我下一个想问的问题回答了，也就是如何判断自己是否适合做科研，总结下来就是要有足够的好奇心，不能急功近利，还要沉得住气。😃  
**「陈光：」**我再补充一个非常重要的点，就是**「做科研并不是以苦为乐」**，不是说做科研就一定要吃得了苦，就像咱们写代码一样，比方说老师给你派了个活，是你很不愿意写的一个程序，或者说你觉得没什么挑战，就是一个工程性的代码，反正随便写写也能交差。但是反过来想，如果你能去充分的挖掘里面的乐趣，哪怕是说调试bug的乐趣，或者说重构代码让程序变得更好的乐趣，如果你能充分的挖掘里面的这些乐趣的话，其实这件事并不苦。经常有人会问，做研究是不是总要经历一个特别痛苦的过程，是不是要吃得了足够的苦，才能有最后的甜。我觉得做科研要吃很多苦是一种误导，因为这个过程它有苦也有甜。同样的一段过程，你可能切换一下看待它的角度，或者切换一下你所在的场景，或者说切换一下你的目标，这个过程其实它是可以变得丰富多彩，变得有意思起来的。  
咱们举个特别简单的例子，比如数据标注，大家也都做过标注，标注这件事，你说到底是苦还是甜？如果你仅仅把它看成是一个重复枯燥劳动的话，它就没什么乐趣。但如果你在这个过程里面你用点心，你带着一些问题，你带着一些对于数据的认知和了解深度的追求，去分析数据，筛选数据，去进一步的有效利用数据。带着这样一些思考，你会发现你看的越多，你的高度也会有所提升。这样的感受就会将这个过程轻松化甚至是愉悦化，**「如果你不能做你爱做的事，就爱上你做的事」**，这句话其实我觉得放在这儿是非常合适的。  
咱们很多同学不管是在研究方向，还是在具体工作具体方法的选择上，总是在想这可能并不是我最喜欢的方向，不是我最喜欢的工作，但是往往没有反过来去认真的去想一想，去挖掘一下你在做的这些事里面，它们有没有真正能吸引你的点。**「有时候你没有喜欢它，仅仅是因为你没有真正的去挖掘它」**，就像我们谈恋爱一样，你如果没有去认真的了解一个人，你可能爱上的仅仅是他的第一印象，你真正能爱上一个人的点其实是他内在的各种特性，如果你现在还不够喜欢这件事，不够喜欢工作，首先你要摆正你的态度，希望自己去喜欢它，这个很重要。  
就像刚才说的这些大牛，或者说真正有了有影响的成果的这些人，抛开别的不说，这个过程一定不是一个在无边无际的黑暗里面去找那一点光亮的过程。从外人看，可能是这样的，因为他一直都没有成果，然后突然一下有了成果，好像是在黑暗中找到了一点光亮，**「但实际上，在他心里始终是有那么一点光亮的，他知道那个方向在哪，而且找这个方向的过程，也让他觉得每一点进步都让他觉得高兴，每一点进步都让他觉得兴奋，所以他才能坚持下去，才能够靠近光亮」**。所以我觉得心态的调整，或者对于自己手头所做的事，态度上不是说别人要求你要怎么样对待你的工作，真的是你只有从内心里，你觉得这件事值得做，你觉得这件事是让你有所收获，能让你有所触动，或者能吸引着你不断的前进的这些事，才是真正能做好的。  
如果仅仅是说为了完成，或者是为了能写出一个有影响力的东西而做科研的话，八成是做不出来的。如果做出来可能也是撞大运。为什么说是潜心研究呢？外人看起来你是茶不思饭不想在做这件事，好像很辛苦，但是你置身其中的时候，实际上你是顾不上吃饭，不想睡觉的，我觉得这是一个好的状态。「如果你和外人的看法一样，你也是觉得自己苦哈哈的在这，真是一天都不想再坚持了，恨不得明天就发表了就完事了，我觉得这个肯定是不会有一个好的结果，任何事都是这样。」 🙂  
  
**「小七：」** 陈老师说的这些品质和精神，我觉得对一个如果是有志于做科研的同学来说是非常重要的，但是目前来看好像大部分同学都是没有做科研的想法的，比如我们实验室的大部分同学都是不会选择继续读博士的，这些同学在选择研究方向的时候，他们可能会选择一些比较好找工作的方向，而不会选一个他们真正感兴趣的方向去研究，因为他们真正感兴趣的方向可能工业界并没有需求，如果不选择做科研的话，我们在硕士阶段应该培养什么与本科阶段不同的能力呢？🤔 **「陈光：」**首先从最终的出路上，我觉得每个人都有自己的方向和选择，或者每个人目前对于人生和价值的看法不一样，我觉得这些都无可厚非，关键是你要知道怎么样才能让自己变得更好，所以我觉得目前你掌握了哪些能力是次要的，首先你得知道你想成为什么样的人。比方说你可能更喜欢做开发，或者说你就是喜欢做研究，或者你就喜欢赚钱，这些都没有问题，但是你要知道为了将来自己成为这样一个人，哪些东西是和你想成为的那个人是直接相关的，而且**「我觉得每个人一生可能会有几个阶段，每个阶段可能会有不同的追求，但是你要清楚离你最近的那个阶段，这个阶段的目标可能不是最远大的一个目标，但是在这个阶段里面，什么东西，或者什么品质能够帮助你成为你想成为那个人」**，我觉得这一点可能是目前大家是需要去思考的。  
但实际上汇成一句话就是，好钢要用在刀刃上，因为读研究生就这么短的时间，对吧？**「就像咱们玩游戏，你去均衡你的各项能力指标一样，你的技能点到底是加在哪个上面，你是加在体力上，还是防御上，还是在攻击力上？这和你最终你是想成为一个法师，还是想成为一个战士相关的」**。所以你首先要清楚自己未来的目标是什么样的，然后现在用最多的时间去打造你觉得你最值得锻炼的品质。  
**「我觉得现在如果武断的说研究生期间就应该重点培养某个能力，我觉得都是不负责任的，你一定是要根据自己的方向来的」**，比方说你将来想搞投资，投资需要的能力是什么？要有冷静的态度，还要有足够广泛的视野，还要有相关的专业知识，包括和你的专业结合在一起的那部分知识，比如用机器学习对股市做分析预测，比如怎么去做项目的评价，这些相关的项目你就多去关注一些，然后去看一看相关的这些方向。我觉得首先你有了足够准确的目标，你现在在做的事，你只要是冲着那个方向去强化你自己的都没有错。你不管是时间管理，还是说对于知识的学习的能力，这些实际上都是和你未来目标相关的。其实现在大家每个人心里面其实是有数的，自己在学什么，在增长哪方面的能力，在强化哪方面的能力，都是有数的。  
有时候你不知道哪个更重要的时候，更本质的原因可能是在于你还没有想清楚未来自己想成为一个什么样的人，所以我觉得如果是靠现在的能力的强化来决定你未来干什么，我觉得莫不如你反过来先想清楚未来你想干什么，再来针对性的来优化你现在得学习，因为按第一种方式的话，很有可能就迷茫，很有可能时间的利用就不够有效，**「就像打游戏一样，如果你不知道未来想成为什么样的人，你可能就随机的把这些技能点分配在各项能力上，最后就是一个各项能力平平的人」**，这肯定不如你为了将来的一个目标，你把技能点堆在一个能力上面，你在这个能力上你是超过一般人，甚至是能够独挡一面的，我觉得这才是最好的。  
如果从通用的能力的角度来讲，我觉得「**研究生最重要的一个能力其实就是自我学习，或者终身学习的能力」**，这个是最重要的。将来你哪怕转方向也好，换了不同的阶段也好，你有了新的人生目标，你想去培养新的能力，你想去学习新的知识，你需要快速的，而且有足够深入的去了解一个领域，我觉得这个能力是非常重要的，而且研究生阶段有可能是你在学校的最后一个阶段，读研的最重要的一个目标就是学习能力的提高，所以学习一定是最重要的。  
如果说第二重要的，应该就是耐力，或者说坚持，或者再从另外一个角度来说，就是眼光要足够的远，眼光足够的远，你才有一个足够长期坚定的目标，你才能去坚持。你要是今天想一样，明天想一样，这就不是有没有耐力的问题，而是你的想法都变了，你没法有耐力。**「所以一定是要在中长期有一个足够坚定的，让你至少这一段时间你都觉得是正确的一个方向，你才能去坚持」**，所以这些我觉得对于研究生的同学来讲可能是最重要的。  
  
**「小七：」**感谢陈老师如此精彩的回答，下一个问题是一个经典问题了，我们做科研都是站在巨人的肩膀上，因此论文的阅读是必不可少的，那么如何选择合适的论文，读论文的时候我们需要注意什么，陈老师有什么建议吗？🤔  
**「陈光：」** 怎么读论文这都是老生常谈的话题了，读论文首先也是要有选择性的读，如果你不会选呢，你就先读那些相对比较有影响力的论文，至少是大家觉得好的论文本身应该也还不错，像这些顶会的论文都是人家审稿人审阅过的，还包括大家普遍关注的一些最新的论文，这些论文一定有它的可取之处，要么是他的工作做的确实扎实，要么他的想法和思路是有可取之处的，当然你有自己的研究方向，你去借鉴相关领域的一些进展，这是一个方面，但是更重要的其实还是说去看一看**「别人是怎么样去分析和思考问题的。」**论文表面上是一个人对于他已有的工作或者进展的一个记录，但是你把很多的论文放在一起，你再去看的时候，**「实际上作者在记录的是他对于目前这个领域的进展和状态的一个思考」**，有时候论文的价值真的不在于提的方法有多好，而是在于他对于这个领域目前存在的问题和现在的前景以及挑战的一种认识，一种观察，他会把他观察到的现有的一个一个的问题和一个一个的现象以及方法做一个抽象，抽象以后得到的对于整个领域的，对于类似方法的，对于类似问题的这种思考，我觉得这个是最重要的。  
所以看论文其实最重要的看的是作者背后的思考，所以我倒建议大家对于很多的论文，先不用急于去细读它的Method，这些东西其实相对来讲不是最重要的，反而是在它的Abstract也好或者是Conclusion也好，在这些里面他提到的一些思考，**「如果它里面单纯是就方法说方法，或者就只提他对于指标的这种改善的话，我觉得这种论文不读也罢」**，可能是对于你的具体的在做的任务上可能是有帮助的，但是对于你将来写论文，我觉得帮助并不是特别大，顶多是一些格式上的参考。  
我觉得最有帮助的还是它能够提供一些思考的角度，怎么样去发现问题，怎么样去看待问题，然后从哪个角度去吸取前人已有工作的这种经验，然后怎么样去把这个领域的工作向前推进一点点，或者从哪个角度去推进，怎么一个推进法？是在指标上去推进，还是在思考的角度上，还是回溯到任务和数据本身去找这里面存在的一些问题。我觉得这种应该算是一个方法，或者说一种思路，这些东西应该是最重要的。**「如果读论文仅仅是读到了论文的结构，或者仅仅是读到了这个论文的贡献的话，我觉得可能就很难真正的吸收到这个论文能够给你带来的那些启发。」  
「小七：」**现在我读了很多论文，就感觉整篇论文最有价值的一个板块就是Introduction，陈老师也有这样的感觉吗？😀 **「陈光：」**对，因为你看很多写的好的论文，其实他在Introduction里面是下了不少功夫的，一方面是，它对领域的进展也会做一个比较全面的小的综述。然后再一个就是，Introduction其实是在说服人，或者说它需要引起你的关注，Introduction里面往往也有很多的技巧，对吧，对其他人的工作，不是说一味的批判，但是他很明显的是为他后面的工作做了一些铺垫，这些就是一些技巧上的或者手法上的特点，然后再一个呢，有一些思考性的东西，因为放在后面不合适，所以往往它也是在Introduction里面，所以这些东西往往是能够给大家带来比较多的启发的。所以有时候我们也经常说怎么去找问题对吧？有的人是从很多论文里面他最后的Future里面找问题，但是可能更多的人是能够在Introduction里面找到问题的，虽然作者是在为他自己的工作做铺垫，但是他提到的很多方法，很多进展，已有的一些问题，可能他这个方法未必能解决得了，**「这时候这些方法或者是这些角度，这些问题，其实是给你也留了一个空间」**。还是刚才说的，他的这种分析和对这些问题的一些思考，对于你来讲应该是最有借鉴价值的。  
  
**「小七：」**下面问一个比较个性化的问题吧。刚才陈老师说每一个科研人员可能会有他自己喜欢的研究方向，或者说有他独特的研究品味，那么陈老师个人是为什么会选择当前的研究方向，也就是自然语言处理，或者说更具体一点的问答系统呢？  
**「陈光：」**这应该是一个历史传承问题😅，因为我们最早的时候是做手写汉字识别，做图像OCR的，然后后来整个实验室又开始做网络搜索，在那会儿就开始接触NLP了，我觉得一个是这其实是当时的那个阶段做出的决定，当时觉得文本相关的处理的开放性和可能性更大一些，因为当时对于图像来讲也没有深度网络，都是手提特征，手提特征呢，那会儿感觉也是有一个天花板，然后文本这一块感觉天花板好像还没有那么明确，因为待解决的问题非常的多，而且像那会儿像舆情里面做什么过滤啊分类啊，包括一些敏感的有害的信息的抽取啊，越做越觉得还是有挺多有待解决的挑战，然后不断的新的一些方法出来以后，你会明显的看到它的一些比较大的进步，然后你就会对它越来越感兴趣。  
如果说要是问我一个问题，为什么选文本不选图像？其实…就是这样🙂，并不是说我原来不是这个领域不是这个圈的，突然让我进入这个圈，让我在文本和图像里面去选，这可能是另外一个问题。但是就我个人来讲，其实就是这么传承下来的，而且就和刚才说的那几个问题一样，我觉得在做文本的这种处理的时候，**「它里面的这些挑战也好，或者说我能够解决的一些问题，给我带来的这种反馈让我很高兴，所以能一直的往前去推进，所以其实还是一个内在驱动在不断的在这个方向上越走越远。」**  
**「小七：」** 如果当年陈老师就选择图像的话，估计也会像现在一样做得很好。 **「陈光：」**好不好另说，至少还是应该挺喜欢的。😂 **「小七：」** 好像陈老师在任何领域，都能发现那个领域的一些很有意思的点。🤔 **「陈光：」**就是还是得让自己兴奋起来吧。😄  
  
**「小七：」** 最后问一个个性化的问题吧，陈老师也是在北邮呆了好几十年了，曾经也是在北邮读的博士，陈老师当年在做科研的时候，有没有遇到什么挫折或者是困难呢？或者说是有什么教训之类的？🤔  
**「陈光：」**那肯定有啊，教训的话，其实刚才也说了，最重要的还是要明确，我觉得不管是读博，读研，还是工作都是一样的，很多时候你迷茫，拖延，或者觉得无助，或者低落，往往就是因为你还没有一个具体的目标，更确切的讲，**「你有没有想把这个目标给它分解成你下一步该做的行动？」**这个很重要。我觉得如果我早一点认识到这一点的话，就会少一些拖延少一些迷茫，进度也会加快一些。我觉得这一点其实是整个读博期间，包括读研，包括毕业以后可能是最大的一个感触吧。  
然后再一个呢就是，觉得可能主动性上，这当然也是和刚才说的目标是有关系的，因为你目标不足够明确，所以你就没有那么多的思考，没有那么多的问题，所以你就没有那么强烈的需要和你的导师来沟通，所以我想说的下一点就是，**「真的是应该和导师多沟通」**，如果你不去主动的话，别人也不知道你在想什么，也不知道你目前的问题是什么样的，对吧？所以我觉得沟通的意义可能很重要的一点是在于相互启发。  
不管是博士阶段还是毕业以后，每次和郭老师一起讨论的时候，其实我都觉得能有一些新的启发。当然有一些胡思乱想的点子也能启发郭老师😂，这个过程会让你觉得非常的好，而且现在自己当导师以后，我觉得最大的一个感受就是说有时候真的是不用去想太多，你可能有时候觉得自己是不是想的方法太简单，或者一个太幼稚的，太天真的想法，这些我觉得都不妨和老师去讨论。所以当学生的时候，我觉得有很多想法都可以一起去讨论一下，**「不要因为觉得自己的想法可能太单纯或者太幼稚，而不去表达或者不去讨论，这可能最终会埋没不少好的点子。」** 所以要多去分享和讨论，不管是和老师还是和身边的同学，我觉得这也是非常重要的，就是沟通，从内来讲就是要明确你的目标和方向，从外就是要多注重沟通，就是一方面要有广泛的信息的来源，另一方面也要多方的去讨论，多方的去求证，多方的去启发。