

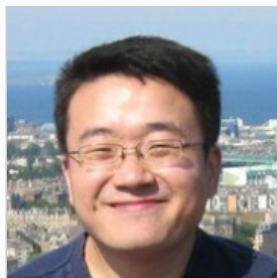
## 刘知远THU的博客



清华大学计算机系助理研究员

<http://blog.sina.com.cn/liuliudong> [订阅] [手机订阅][首页](#) | [博文目录](#) | [图片](#) | [关于我](#)

## 个人资料



刘知远THU



加好友

发纸条

写留言

加关注

博客地图 world map

博客等级: 14

博客积分: 510

博客访问: 66,105

关注人气: 284

获赠金笔: 22

赠出金笔: 0

荣誉徽章:

## 相关博文

- 图像速记汉字  
老铁
- 《米奇妙妙屋大搜寻》MP4格式128  
茉莉早教动画片下载
- 美图 | 简欧温馨小窝  
用户389421361
- 厦门买房时如何辨别户型的好坏  
用户373238863
- 2018年3月4日播出《相伴大国医之  
家庭养生堂
- 区块链技术内涵和应用领域  
理财师郭乐
- 【实拍】三十年一遇四月雪惊艳  
汽车兵1976
- 资金下周爆拉一股  
趋势老班长
- 雍正为给爱将治病剪了自己的胡

## 正文

字体大小: 大 中 小

## 初学者如何查阅自然语言处理 (NLP) 领域学术资料 (2013-07-31 12:57:59)

标签: it



昨天实验室一位刚进组的同学发邮件来问我如何查找学术论文, 这让我想起自己刚读研究生时茫然四顾的情形: 看着学长们高谈阔论领域动态, 却不知如何入门。经过研究生几年的耳濡目染, 现在终于能自信地知道去哪儿了解最新科研动态了。我想这可能是初学者们共通的困惑, 与其只告诉一个人知道, 不如将这些Folk Knowledge写下来, 来减少更多人的麻烦吧。当然, 这个总结不过是一家之言, 只盼有人能从中获得一点点益处, 受个人认知所限, 难免挂一漏万, 还望大家海涵指正。

## 1. 国际学术组织、学术会议与学术论文

自然语言处理 (natural language processing, NLP) 在很大程度上与计算语言学 (computational linguistics, CL) 重合。与其他计算机学科类似, NLP/CL有一个属于自己的最权威的国际专业学会, 叫做The Association for Computational Linguistics (ACL, URL: <http://aclweb.org/>), 这个协会主办了NLP/CL领域最权威的国际会议, 即ACL年会, ACL学会还会在北美和欧洲召开分年会, 分别称为NAACL和EACL。除此之外, ACL学会下设多个特殊兴趣小组 (special interest groups, SIGs), 聚集了NLP/CL不同子领域的学者, 性质类似一个大学校园的兴趣社团。其中比较有名的诸如SIGDAT (Linguistic data and corpus-based approaches to NLP)、SIGNLL (Natural Language Learning) 等。这些SIGs也会召开一些国际学术会议, 其中比较有名的就是SIGDAT组织的EMNLP (Conference on Empirical Methods on Natural Language Processing) 和SIGNLL组织的CoNLL (Conference on Natural Language Learning)。此外还有一个International Committee on Computational Linguistics的老牌NLP/CL学术组织, 它每两年组织一个称为International Conference on Computational Linguistics (COLING)的国际会议, 也是NLP/CL的重要学术会议。NLP/CL的主要学术论文就分布在这些会议上。

作为NLP/CL领域的学者最大的幸福在于, ACL学会网站建立了称作ACL Anthology的页面 (URL: <http://aclweb.org/anthology-new/>), 支持该领域绝大部分国际学术会议论文的免费下载, 甚至包含了其他组织主办的学术会议, 例如COLING、IJCNLP等, 并支持基于Google的全文检索功能, 可谓一站在手, NLP论文我有。由于这个论文集合非常庞大, 并且可以开放获取, 很多学者也基于它开展研究, 提供了更丰富的检索支持, 具体入口可以参考ACL Anthology页面上方搜索框右侧的不同检索按钮。

与大部分计算机学科类似, 由于技术发展迅速, NLP/CL领域更重视发表学术会议论文, 原因是发表周期短, 并可以通过会议进行交流。当然NLP/CL也有自己的旗舰学术期刊, 发表过很多经典学术论文, 那就是Computational Linguistics (URL: <http://www.mitpressjournals.org/loi/coli>)。该期刊每期只有几篇文章, 平均质量高于会议论文, 时间允许的话值得及时追踪。此外, ACL学会为了提高学术影响力, 也刚刚创办了Transactions of ACL (TACL, URL: <http://www.transacl.org/>), 值得关注。值得一提的是这两份期刊也都是开放获取的。此外也有一些与NLP/CL有关的期刊, 如ACM Transactions on Speech and Language Processing, ACM Transactions on Asian Language Information Processing, Journal of Quantitative Linguistics等等。

根据Google Scholar Metrics 2013年对NLP/CL学术期刊和会议的评价, ACL、EMNLP、NAACL、COLING、LREC、Computational Linguistics位于前5位, 基本反映了本领域学者的关注程度。

傅华轩

- 不可辜负的樱花季—晶莹剔透的丹青妙厨

[更多>>](#)

#### 推荐博文

- iPhoneX再次砍单600万
- 网盘衰落之后，为何这款私人云盘
- 短期调整或将仍有反复
- 人工智能与隐私保护
- 台湾科技挣扎，人祸大于天灾？
- 收入份额=市场份额，虎嗅想干什
- 传奇的谢幕，谈岩田聪和他的任天堂
- 家常主食轻松做之——培根香葱花
- 盘点2015最惊艳流行的婚礼蛋
- 美团打车vs滴滴外卖，火力全开

[查看更多>>](#)

#### 谁看过这篇博文

 夜雨	0分钟前
 小雅wil	4月7日
 猴在火星	4月6日
 修罗	4月6日
 jack0077555	4月6日
 荒唐的世界	4月6日
 雨浮东海	4月6日
 足球迷_...	4月6日
 orangefancy	4月5日
 安防本科生	4月5日
 GrantYu	4月4日
 真常太清...	4月4日

NLP/CL作为交叉学科，其相关领域也值得关注。主要包括以下几个方面：（1）信息检索和数据挖掘领域。相关学术会议主要由美国计算机学会（ACM）主办，包括SIGIR、WWW、WSDM等；（2）人工智能领域。相关学术会议主要包括AAAI和IJCAI等，相关学术期刊主要包括Artificial Intelligence和Journal of AI Research；（3）机器学习领域，相关学术会议主要包括ICML，NIPS，AISTATS，UAI等，相关学术期刊主要包括Journal of Machine Learning Research（JMLR）和Machine Learning（ML）等。例如最近兴起的knowledge graph研究论文，就有相当一部分发表在人工智能和信息检索领域的会议和期刊上。实际上国内计算机学会（CCF）制定了“中国计算机学会推荐国际学术会议和期刊目录”（<http://www.ccf.org.cn/sites/ccf/aboutpm.jsp?contentId=2567814757463>），通过这个列表，可以迅速了解每个领域的主要期刊与学术会议。

最后，值得一提的是，美国Hal Daumé III维护了一个natural language processing的博客（<http://nlpers.blogspot.com/>），经常评论最新学术动态，值得关注。我经常看他关于ACL、NAACL等学术会议的参会感想和对论文的点评，很有启发。另外，ACL学会维护了一个Wiki页面（<http://aclweb.org/aclwiki/>），包含了大量NLP/CL的相关信息，如著名研究机构、历届会议录用率，等等，都是居家必备之良品，值得深挖。

## 2. 国内学术组织、学术会议与学术论文

与国际上相似，国内也有一个与NLP/CL相关的学会，叫做中国中文信息学会（URL：<http://www.cipsc.org.cn/>）。通过学会的理事名单（<http://www.cipsc.org.cn/lingdao.php>）基本可以了解国内从事NLP/CL的主要单位和学者。学会每年组织很多学术会议，例如全国计算语言学学术会议（CCL）、全国青年计算语言学研讨会（YCCL）、全国信息检索学术会议（CCIR）、全国机器翻译研讨会（CWMT），等等，是国内NLP/CL学者进行学术交流的重要平台。尤其值得一提的是，全国青年计算语言学研讨会是专门面向国内NLP/CL研究生的学术会议，从组织到审稿都由该领域研究生担任，非常有特色，也是NLP/CL同学们学术交流、快速成长的好去处。值得一提的是，2010年在北京召开的COLING以及2015年即将在北京召开的ACL，学会都是主要承办者，这也一定程度上反映了学会在国内NLP/CL领域的重要地位。此外，计算机学会中文信息技术专委会组织的自然语言处理与中文计算会议（NLP&CC）也是最近崛起的重要学术会议。中文信息学会主编了一份历史悠久的《中文信息学报》，是国内该领域的重要学术期刊，发表过很多篇重量级论文。此外，国内著名的《计算机学报》、《软件学报》等期刊上也经常有NLP/CL论文发表，值得关注。

过去几年，在水木社区BBS上开设的AI、NLP版面曾经是国内NLP/CL领域在线交流讨论的重要平台。这几年随着社会媒体的发展，越来越多学者转战新浪微博，有浓厚的交流氛围。如何找到这些学者呢，一个简单的方法就是在新浪微博搜索的“找人”功能中检索“自然语言处理”、“计算语言学”、“信息检索”、“机器学习”等字样，马上就能跟过去只在论文中看到名字的老师同学们近距离交流了。还有一种办法，清华大学梁斌开发的“微博寻人”系统（<http://xunren.thuir.org/>）可以检索每个领域的有影响力人士，因此也可以用来寻找NLP/CL领域的重要学者。值得一提的是，很多在国外任教的老师和求学的同学也活跃在新浪微博上，例如王威廉

（<http://weibo.com/u/1657470871>）、李沐（<http://weibo.com/mli65>）等，经常爆料业内新闻，值得关注。还有，国内NLP/CL的著名博客是52nlp（<http://www.52nlp.cn/>），影响力比较大。总之，学术研究既需要苦练内功，也需要与人交流。所谓言者无意、听者有心，也许其他人的一句话就能点醒你苦思良久的课题。无疑，博客微博等提供了很好的交流平台，当然也注意不要沉迷哦。

## 3. 如何快速了解某个领域研究进展

最后简单说一下快速了解某领域研究进展的经验。你会发现，搜索引擎是查阅文献的重要工具，尤其是谷歌提供的Google Scholar，由于其庞大的索引量，将是我们披荆斩棘的利器。

当需要了解某个领域，如果能找到一篇该领域的最新研究综述，就省劲多了。最方便的方法还是在Google Scholar中搜索“领域名称 + survey / review / tutorial / 综述”来查找。也有一些出版社专门出版各领域的综述文章，例如NOW Publisher出版的Foundations and Trends系列，Morgan & Claypool Publisher出版的Synthesis Lectures on Human Language Technologies系列等。它们发表了很多热门方向的综述，如文档摘要、情感分析和意见挖掘、学习排序、语言模型等。

如果方向太新还没有相关综述，一般还可以查找该方向发表的最新论文，阅读它们的“相关工作”章节，顺着列出的参考文献，就基本能够了解相关研究脉络了。当然，还有很多其他办法，例如去videlectures.net上看著名学者在各大学术会议或暑期学校上做的tutorial报告，去直接咨询这个领域的研究者，等等。