

“基于知识图谱的自动化新闻标签技术” 研究合作协议

甲 方： 北京奇虎科技有限公司

乙 方： 东南大学

技术研究合作协议

本协议由以下双方于 年 月 日在北京签署：

甲方： 北京奇虎科技有限公司

住 所 地： 北京市西城区新街口外大街 28 号 D 座 112 室（德胜园区）

项目联系人：李晓明

电 话： 18912380598 ； 传 真：

电子信箱： lixiaoming-dc@360.cn

乙方： 东南大学

住所地： 江苏省南京市四牌楼 2 号

项目联系人：漆桂林

通讯地址： 江苏省南京市四牌楼 2 号

电 话： 02552090910 ； 传 真： 02552090910

电子信箱： gqi@seu.edu.cn； 邮编： 211189

本合同基于甲方委托乙方研究开发“基于知识图谱的自动化新闻标签技术”项目，并支付研究开发经费，乙方接受委托并进行此项研究开发工作。甲乙双方经过平等协商，根据《中华人民共和国合同法》的规定，就合作事宜达成如下协议：

一、 项目名称

基于知识图谱的自动化新闻标签技术

二、 合作内容

本合作课题中乙方需要完成的服务内容包括：

(1) 乙方应提供一套可运行系统，用于将甲方给定标签集合，按照相关度标注在甲方给定的新闻集合上。具体要求如下：

将甲方给定的广告短语标注在新闻文本上，标注结果准确率不低于 60%，覆盖率不低于 50%；其中，准确率为所标注标签中正确的数量除以所标注的标签总量；覆盖率为所标注标签中正确的数量除以所有标签总量。

(2) 乙方应提供一套通过百科构建的知识图谱，该图谱用于保证上述系统的正常运行。

甲方应在本合同生效后向乙方提供研究开发计划。乙方应当按照研究开发计划以及下列进度要求完成研究开发工作：

(1) 交付的形式及内容

1、可运行系统源码及系统文档说明。

2、用于该系统运行的知识图谱数据。在合同有效期及终止之后，乙方增加知识图谱更新 5 次。

(2) 交付的时间：2019 年 1 月 10 日之前。

3. 验收

双方确定，依据甲方提出的合同标准及方法对乙方完成的研究开发成果进行验收。甲方验收后审核不符合合同标准的，乙方有义务进行修改，直到符合合同标准为止。

三、 研究合作经费及其支付、结算方式

研究合作经费总额为人民币 20 万元整，分两次支付。

1. 甲方于本合同签订后向乙方支付人民币 10万元；
2. 甲方在对乙方提交的研究开发成果及结题报告验收合格后向乙方支付
剩余款项人民币 10万元；

乙方应在甲方每次付款之前，向甲方开具正式等额的增值税专用发票。
甲方应在本协议签订后，并且收到乙方提供合法等额的增值税专用发票后
15 个工作日内按照合同约定的金额支付到乙方指定帐户内。

乙方帐户信息：

账户：东南大学

支行：建设银行南京市四牌楼分理处

账号：32001594138059123456

甲方开票信息：

名称：北京奇虎科技有限公司

税号： 911101026662879416

地址：北京市西城区新街口外大街 28 号 D 座 112 室(德胜园区)

开户行：招商银行北京建国路支行

账号： 862684692210001

四、 保密

- 1、研究期间，甲方提供的所有数据、文档及其他信息归属于甲方所有。

甲方向乙方提供与本项目有关的数据，说明文档等技术资料，乙方保证不

复制、不泄露上述技术资料以及甲方的其他保密信息给第三方，亦不用于本协议项目之外的目的。

2、乙方应保证对于本合同内容以及在本合同履行过程中所获知的甲方的技术、商业、管理等方面所有保密信息负有保密义务，未经甲方事先书面许可，不得以任何方式将前述保密信息向第三方泄露和公开，也不得用于履行本合同之外的其他目的。否则，因此给甲方造成的损失，应承担相应的赔偿责任。

3、乙方保证乙方的技术人员知悉并遵守本保密条款，由于乙方技术人员泄露保密信息给甲方造成损失的，乙方应当向甲方承担赔偿责任。

4、保密期限自本协议生效之日起至协议终止后一年。

五、 风险责任的承担

由于不可抗拒情况或出现因现有技术水平和条件下难以克服的技术困难导致本协议不能履行，双方另行协商，但受不可抗力影响或技术条件限制的一方应当及时通知对方并采取合理措施防止损失的扩大。不可抗力持续 30 日以上的，任何一方均有权要求终止本合同，合同终止后，甲方有权要求乙方交付已经完成的技术成果以及返还尚未完成部分的价款。

六、 技术成果的归属和分享

1. 甲乙双方在签署本合同之前各自拥有的知识产权及相应权益均归各自所有，不因本协议的签署而改变；

2. 因执行本协议的需要，甲方向乙方提供的信息、资料，不构成向乙方授予任何关于专利、著作权、商标权等知识产权的许可行为，并不因本协议而发生转移；
3. 乙方与甲方项目合作中所取得的知识图谱数据，从公网获取部分归甲乙双方共同所有；
4. 乙方向甲方交付后的可运行系统源码及系统文档说明，其知识产权及所有其他权利均归甲方所有；
5. 乙方基于本合作项目的成果发表论文（包括会议论文、期刊论文、学位论文等）、发表演讲报告、参加学术会议、技术交流等时，均需注明成果受甲方资助完成。其中，发表论文由甲乙双方共同署名，甲乙双方可就项目中的技术各自申请专利。
6. 乙方承诺研究成果不会侵犯第三方知识产权以及其他合法权益，如第三方主张权利，由乙方单独承担责任，如对甲方造成损失，还应当承担甲方因此遭受的所有损失。
7. 乙方不得在向甲方交付研究开发成果之前，自行将研究开发成果转让、泄露、公开给第三人。
8. 双方确定，甲方有权利用乙方按照本合同约定提供的研究开发成果进行后续改进。由此产生的具有实质性或创造性技术进步特征新的技术成果及其权属，由甲方享有。

乙方有权在完成本合同约定的研究开发工作后，可继续对数据进行

更新，更新的数据归乙方所有。但乙方应根据本协议约定，将知识图谱 5 次更新数据提供给甲方，具体更新频率和内容须由甲乙双方协商确定。

七、 项目联系人

双方商定，在本协议有效期内，李晓明为甲方项目联系人，漆桂林为乙方项目联系人。

八、 协议变更或解除

本协议经双方协商可以变更，变更时应签署变更协议。本协议到期前，如有一方希望解除协议，应与另一方协商。对本合同任何条款的修改、补充或更改，双方必须签订书面协议并签字盖章后方可生效。

九、 违约赔偿

1. 甲、乙双方任何一方直接或间接违反本合同的任何条款或不及时、充分地承担本合同项下其应当承担的义务而构成违约行为的，守约方有权以书面通知要求违约方纠正其违约行为并采取充分、有效、及时的措施消除违约后果并赔偿守约方因违约方之违约行为而遭致的损失。若违约方在收到守约方关于其违约行为的上述通知后十五个工作日内未纠正其违约行为，守约方有权书面通知合同相对方终止本合同。乙方违约的，还应返还甲方已支付的合同款。
2. 在违约事实发生以后，经守约方的合理及客观的判断该等违约事实已造成守约方履行本合同项下其相应的义务已不可能或不公平，守约方有权以书面通知要求违约的一方违约方纠正其违约行为，并采取充分、有效、及时的措施消除违约后果，并赔偿守约方因此而遭致的损失；若违约方在收到守约方关于其违约行为的上述通知后十五个工作日内未纠正其违约行为，守约方有权书面通知合同相对方终止本合同。
3. 一方违约的，合同相对方应当在合理范围内采取适当措施防止损失的扩大；没有采取适当措施致使损失扩大的，不得就扩大的损失要求赔偿。一方因防止损失扩大而支出的合理费用应由违约方承担。如守约方不采取措施减少损失，对于本应减轻的损失数额，违约方有权要求从损害赔偿中扣除。

4. 违约方因其违约行为而应赔偿的守约方的损失包括但不限于守约方因违约方的违约行为而遭致的直接的经济损失及任何可预期的间接损失及其他合理的费用，包括但不限于律师费用、诉讼及仲裁费用、财务费用及差旅费等。

十、附则

1. 为确保合作过程中共同遵守商业诚信行为规范、制止商业贿赂行为，乙方确认已完全阅读及并承诺遵守甲方的商业诚信约定，具体链接网址如下：<http://www.360.cn/about/shangyechengxinxieyi.html>

2. 本协议一式肆份，双方各持贰份，每份具有同等法律效力。

3. 本协议有效期：自合同生效之日起 2019 年 1 月 10 日

4. 本协议自甲乙双方盖章之日起生效。

5. 本协议未尽事宜，适用中国法律法规，同时甲、乙双方可协商签订补充条款解决，并与本协议具有同等法律效力。

6. 双方在履行协议过程中所产生的一切纠纷，由双方协商解决，协商不成的，应通过被告方有管辖权人民法院通过诉讼解决。

（本行以下无正文）

甲方：北京奇虎科技有限公司（盖章）

法定代表人/委托代理人：（签名）

年 月 日

乙方：东南大学（盖章）

法定代表人/委托代理人：（签名）

项目负责人：（签名）

年 月 日

附件 1: 课题研究计划

①首先,使用关键词抽取和命名实体识别对新闻文本进行重新表示;

为了获取新闻中的有用信息,需要获取新闻的语义表示,使用关键词抽取和命名实体识别技术对原新闻文本进行处理,获取关键词和实体,用于对该新闻重新表示。

②使用实体链接技术将抽取出来的实体(关键词)链接到知识库中;

拿到步骤①中获取到的实体集合,将每个实体链接到知识库中,Zhishi.me 实体丰富,大部分实体都可以映射上去,由于广告主购买的广告词动态改变,因此知识库也需要动态更新。将实体映射到知识库中后,可以得到该实体的上下位关系等大量属性以及文本描述。

③将广告主购买的关键词预处理,与步骤②类似,将处理结果链接到知识库;

④对广告主购买的关键词进行分词、去停用词,同样将这些关键词映射到知识库中,获取这些关键词的属性和文本描述。

⑤将步骤②、③的结果计算语义相似度并排序,获取其中 TopK 的关键词对应的广告短语,标注在新闻文本上。

获取新闻关键实体在知识库中的各种属性和文本描述,以及广告主购买的关键短语在知识库中的属性和文本描述,分别计算属性和文

本描述之间的相似度，取其中 TopK 对应的广告短语标注在新闻上，或者设置指定阈值，超过阈值的即可标注在新闻上。

附件 2：预算安排

说明项	数量	单位	单价 (人民币)	金额 (人民币)
人员支出费：项目参加人员的劳务支出费，3 个月：教授 1 人 10,000/月，副教授 1 人 10,000/月，博后 2 人 4,000/月，博士 2 人 3,000/月，硕士 4 人 1,500/人月	1.00	项	120,000.00	120,000.00
标注测试费：前期数据标注及后面测试的相关费用	1.00	项	10,000.00	10,000.00
差旅会议费：开会讨论，做实验往返学校与公司的交通费用	1.00	项	36,000.00	36,000.00
科研业务费：项目研发过程中所发生的文献检索费、通讯费、图书资料等费用	1.00	项	10,000.00	10,000.00
学校管理费：学校的协作管理费	1.00	项	24,000.00	24,000.00
合计				200,000.00