

CASIA OPENIR 数据存缴报告（2016 年 8 月）

科研支持部信息中心 张桂英

CASIA OPENIR 各年份数据量统计

表 1 CASIAIR 各年份数据表

年份	数据量	年份	数据量	年份	数据量	年份	数据量
2016	442	2007	305	1998	63	1989	45
2015	735	2006	329	1997	43	1988	26
2014	826	2005	246	1996	40	1987	19
2013	726	2004	202	1995	33	1986	9
2012	751	2003	171	1994	38	1985	3
2011	854	2002	73	1993	25	1984	19
2010	794	2001	71	1992	38	1983	1
2009	668	2000	66	1991	35	1982	3
2008	454	1999	58	1990	25	1900-1981	39
合计				8274			

截至 2016 年 8 月底，研究所机构知识库数据量达 8274 条，近三个月新增 583 条，增长率 7.6%。新增数据主要集中在 2015-2016 年，合计新增 432 条，占有增长量的 74.1%；其中，2016 年的数据比 5 月份增长 298 条，为 5 月份的三倍。另外，2012-2014 年新增数据 134 条，近年的历史数据同时增长中。

CASIA OPENIR 各内容类型数据量统计

在图 1 的九种类型文献中，七种类型有不同程度的增长，主要集中在会议论文、期刊论文和学位论文，增长量分别为 229、155 和 139 条，合计 523 条，占有增长量的 89.7%。前三类数据占有所有数据的 88.3%，其中，期刊论文数量突破 3000 条，占有所有数据的 37.7%。

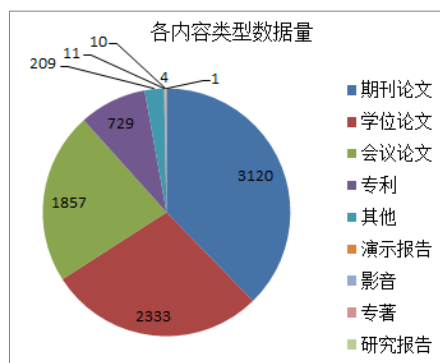


图 1 各内容类型数据量

CASIA OPENIR 各团队数据量统计

截至 2016 年 8 月底，41 个团队数据量排序如图 2。其中 31 个团队的数据量有增长，先进控制与自动化团队、中国科学院分子影像重点实验室和先进机器人控制团队依然排在前三位，先进控制与自动化团队以 497 条排在第一。

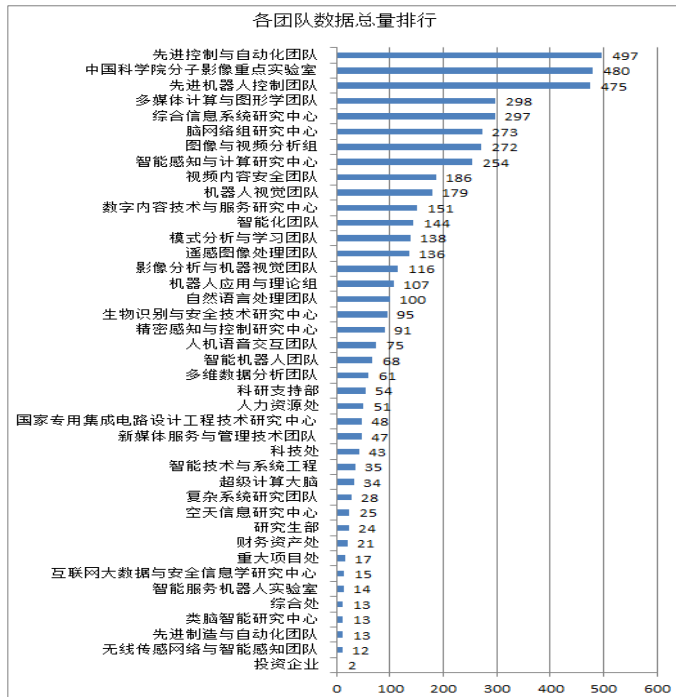


图 2 各团队数据量排行

CASIA OPENIR 各年份全文存储量统计

研究所机构知识库的全文主要来自近几年，近 10 年全文数量占有所有全文的 95%。截至 8 月底，2015 年全文存储最多为 510 篇，2016 年全文存储率最高至 83.3%，2014-2016 三年的平均全文存储率 68%。

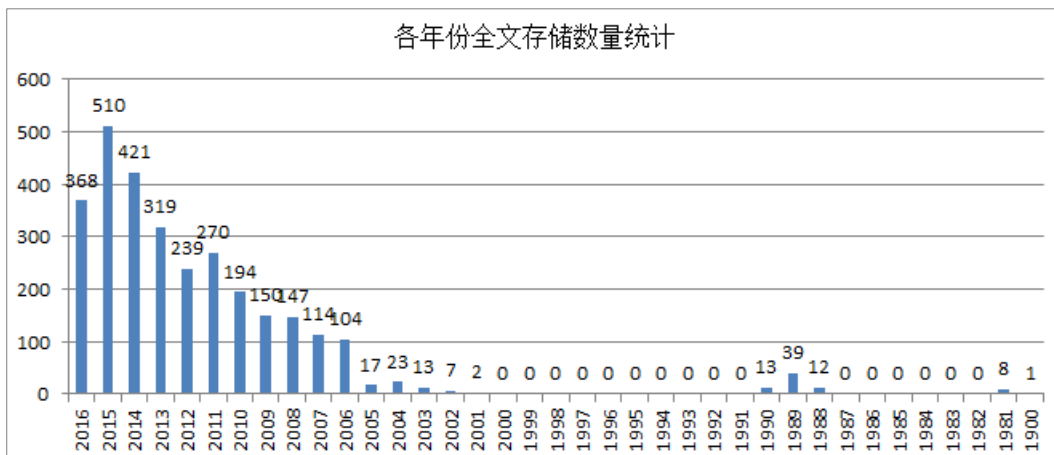


图 3 各年份全文存储量统计

CASIA OPENIR 各内容类型全文存储量统计

截至 8 月底，所有内容类型的全文合计 2972 篇，全文存储率为 35.9%，比 5 月份增高了 15.5%。在各内容类型的文献中，全文主要来自期刊论文、学位论文、会议论文，合计占有全文的 90.7%。其中，全文存储量最大的是学位论文，达 1514 篇，占有全文的 50.9%，学位论文全文存储率 64.9%，主要是近期科研支持部系统收集并集中上传了 1369 篇全文。

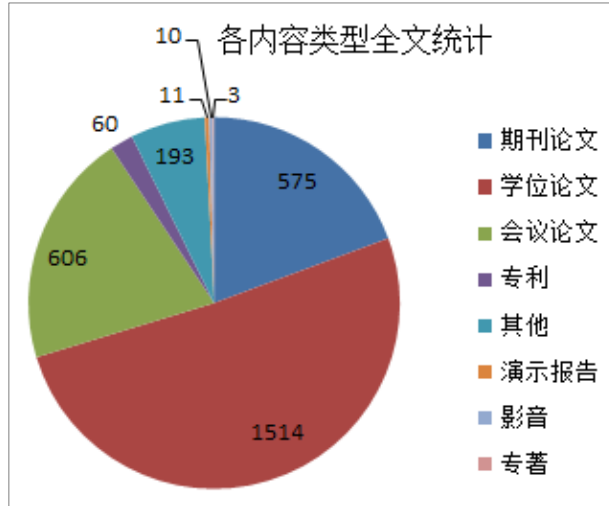


图 4 各内容类型全文存储量

CASIA OPENIR 各团队全文存储量统计

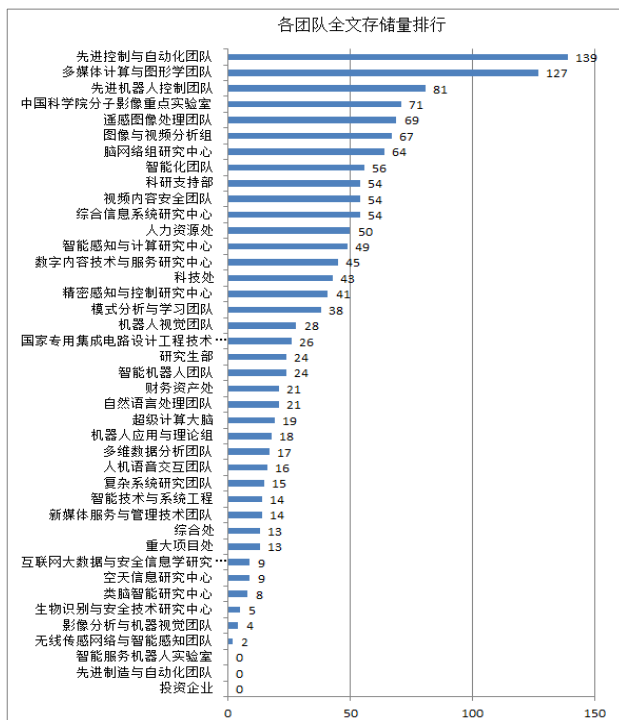


图 5 各团队全文存储量排行

研究所 41 个团队中，32 个团队全文数据有增长，且 12 个团队数据超过 50（含）篇，比 5 月份增加了 8 个。

其中，先进控制与自动化、多媒体计算与图形学团队分别以 139 和 127 篇全文排前两位。

CASIA OPENIR 各团队全文存储率统计

研究所机构知识库中，全文上载率大、排名靠前的团队还是来自管理和支撑部门，如图 6 所示。

另外，上载率超过 50% 的团队增加到 13 个，比 5 月份新增了 3 个，均为科研团队。

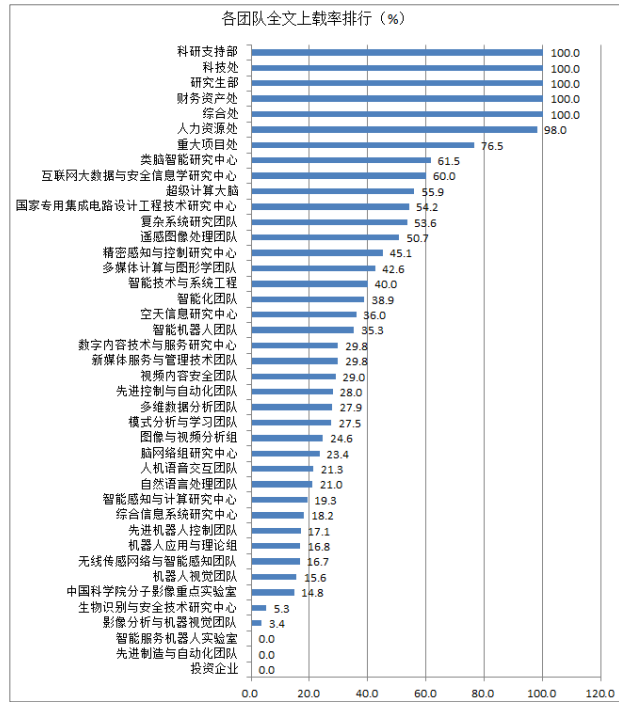


图 6 各团队全文存储率排行

CASIA OPENIR 5 月份和 8 月份访问量对比

截至 8 月底，研究所机构知识库访问量合计 85583，比 5 月份增长了 33110 次，增长量为 5 月份的 63%。6-8 月份，月均访问量 11037 次。

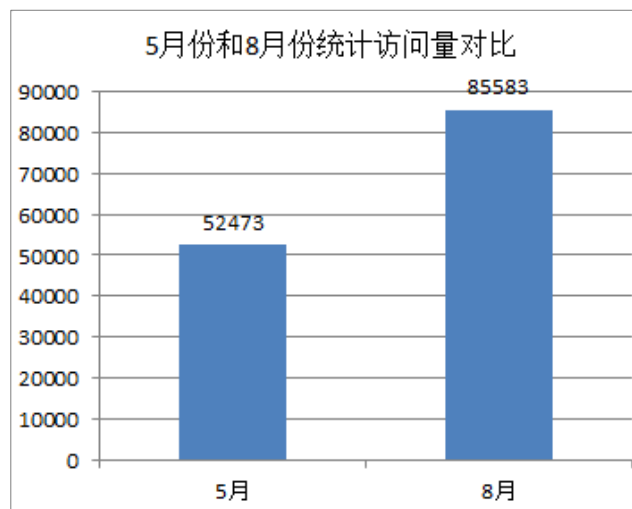


图 7 5 月份和 8 月份统计访问量对比

CASIA OPENIR 5 月份和 8 月份下载量对比

截至 8 月底，研究所机构知识库下载量合计 2302 篇次，比 5 月份统计增长了 1145 篇次，增长量为 5 月份的 99%。6-8 月份，月均下载量 382 篇次。

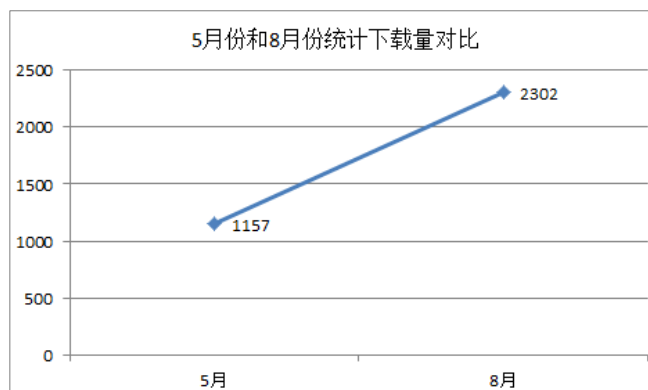


图 8 5 月份和 8 月份下载量对比

CASIA OPENIR 各内容类型按访问量统计

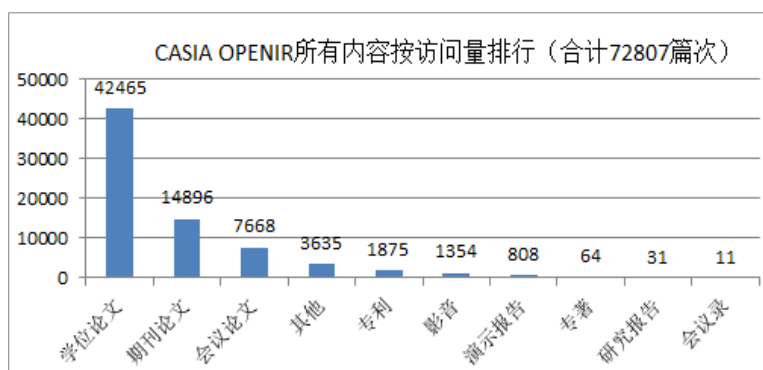


图 9 各内容类型访问量排行

截至 8 月底，研究所机构知识库的 18 种类型文献中，10 种类型文献访问量合计 72807 篇次，是 5 月份的 1.7 倍。

与 5 月份相同的是，主要访问量依然分布在学位论文和期刊论文，两类文献访问共计 57361 篇次，占有所有访问量的 78.8%。其中，学位论文的访问量高达 42465 篇次，占有所有访问量的 58.3%。6-8 月份间，学位论文访问量共计 21716 篇次，月均访问量 7239 篇次。

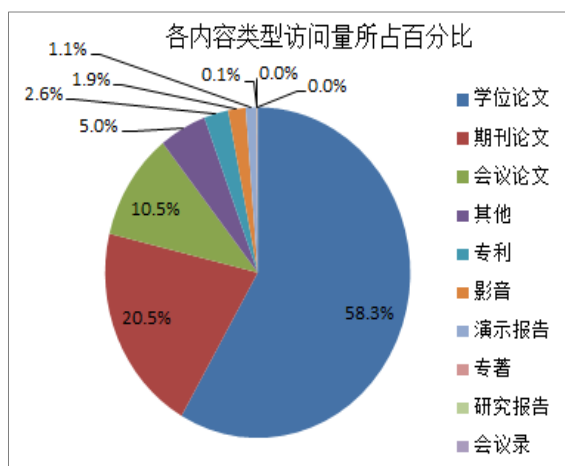


图 10 各内容类型访问量所占百分比

另外，尽管我所的期刊论文多数都可以在电子数据库中找到，但机构知识库中仍有较大的访问量。

CASIA OPENIR 各内容类型按下载量统计

截至 8 月底，各内容类型文献下载量合计 2302 篇次，是 5 月份的 2 倍。类型为“其他”的文献下载量最高占 40.4%，其他类型的文献下载量则相对分散。

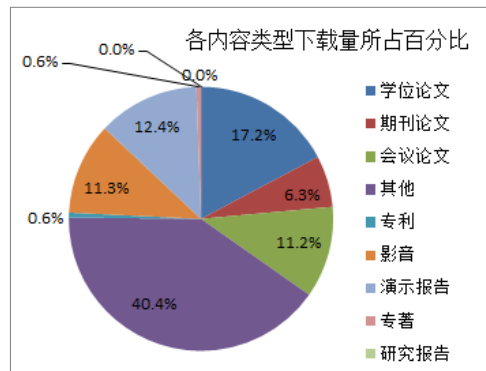


图 11 各内容类型下载量所占百分比

期刊论文下载量排名第 6、共 145 篇次，比起排名第 2、合计 1.49 万的访问量较少。

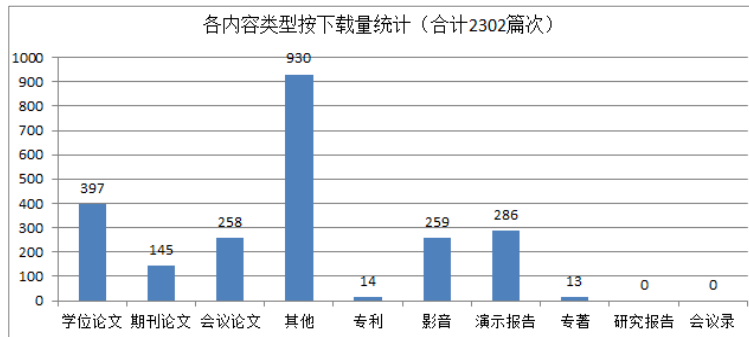


图 12 各内容类型下载量统计

CASIA OPENIR 各团队按访问量统计

41 个团队访问量总计 37687 次，比 5 月份访问量新增 9997 次，是 5 月份访问量的约 1.36 倍。

其中，科研支持部、中国科学院分子影像重点实验室和先进机器人控制团队 3 个团队访问量超过两千，超过一千的团队上增加到 17 个，比 5 月份多了 9 个团队。

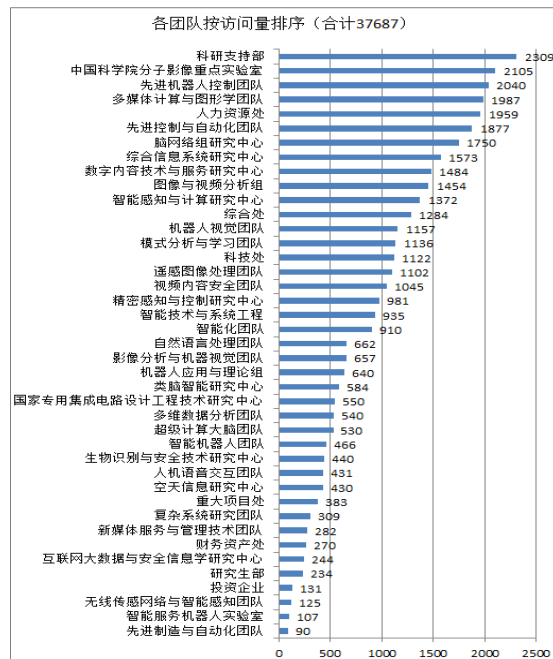
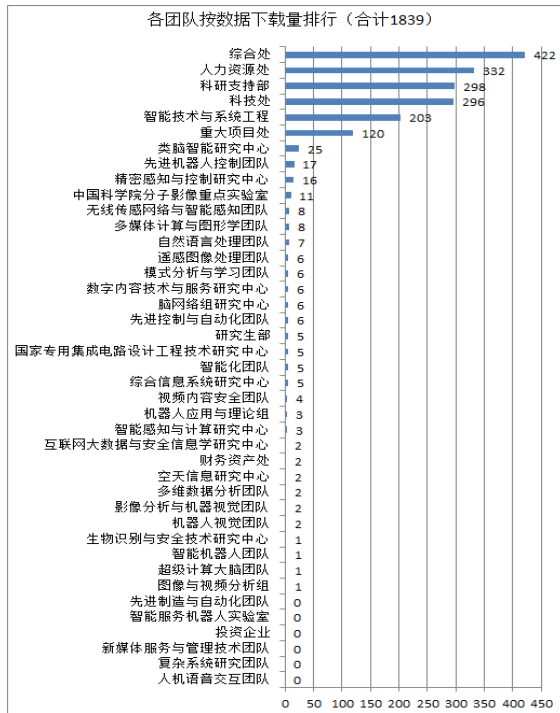


图 13 各团队数据访问量排行

CASIA OPENIR 各团队按下载量统计



41 个团队的所有下载量合计 1839 篇次，是 5 月份下载量的 1.6 倍，排位前 5 的团队没有变化。另外，重大项目处下载量从 5 月份的 7 篇次增长为 120 篇次，增长较多。7 个管理和支撑部门的下载量合计 1475 篇次，占 80%；34 个研究单元下载量合计 364 篇次，占 20%。研究单元数据下载量比例有提升。

图 14 各团队数据被访问次数排行

据另外统计，34 个科研团队（含投资企业）中，期刊全文开放的有 5 个，会议和专利全文开放的团队各 4 个，开放比例分别约为 14.7%、11.8%、11.8%，而管理和支撑团队共 46 个专题，38 个是开放的，约 82.6%。图 14 中下载量前 3 名的科研团队中，期刊全文全部为开放获取，2 个会议论文和专利为开放获取。以上数字均可说明下载量跟开放程度有直接的关系。

事实上，期刊论文和会议论文的开放获取，对于提高我所科研成果的引用率、提升研究所在同领域的影响力，应该是有益的。

CASIA OPENIR 条目访问量 TOP20

表 2 访问量 TOP20 的条目排行

序号	条目	作者&导师	年份	访问量
1	视觉 SLAM 中的多传感器标定及稠密地图创建的研究	李立杭 吴毅红	2015	1188
2	基于深度学习的特征表示和图像分类方法研究	刘炳源 卢汉清	2015	1111

3	图像内容表示与分类方法研究	张琳波 王春恒; 肖柏华	2011	681
4	图像识别中的特征表达方法研究	刘颖璐 刘成林	2015	600
5	中医证候的复杂系统建模及其与疾病的相关性研究	陈建新 西广成	2008	593
6	不同光照下的异构人脸图像的融合识别方法	许力 刘昌平	2008	489
7*	自动化所 2015 年度战略研讨会主报告	徐波	2015	454
8*	神经科学的回顾与前瞻	蒲慕明	2015	444
9	基于计算机视觉理论和方法的数字艺术	曾坤 杨青	2008	436
10	基于无线传感网络的滚动轴承在线监测系统	姚康平 李学恩	2013	430
11	基于无线传感网络的管道监测系统研究	吴泽华 李学恩	2013	424
12	面向非结构化文本的关系抽取关键技术研究	曾道建 赵军	2015	382
13	文本定位研究	欧文武 刘昌平	2004	359
14	磁共振神经影像数据分类方法的研究及其应用	戴岱 何晖光	2012	359
15	氯化釜温度的先进控制算法研究与实现	刘绍辉 庄诚	2014	345
16	基于弱监督学习的网络图像语义解析	刘洋 卢汉清	2014	337
17	面向网络的中文实体关系抽取的研究	王昊 徐波	2015	331
18	自然场景中文字检测与识别方法研究	史存召 王春恒	2014	330
19	人机口语对话系统的知识自动生成技术	黄韵竹 李成荣	2011	322
20	汽油自动调合系统关键技术研究	王伟 谭民; 李 泽飞	2007	314

访问量 TOP20 的文献，总访问量合计 9929 次，是 5 月份的近 1.6 倍。以上 20 条数据中，除了第 7、8 两条，其他 18 条均为学位论文，访问量 9031 次，约占 91%。大家对学位论文一直保持非常高的关注度。

CASIA OPENIR 条目下载量 TOP20

下载量 TOP20 的文献，下载量合计 1163 篇次。下载量最大的是蒲慕明院士的演示报告，下载条目最多的是“其他”，有 10 条、583 篇次，占 TOP20 下载总量的一半。

表 3 下载量 TOP20 的条目排行

序号	条目	内容类型	第一作者	下载次数
1	神经科学的回顾与前瞻	演示报告	蒲慕明	158
2	自动化所 2015 年度战略研讨会主报告	其他	徐波	149
3	自动化所程序文件汇编	其他	廉敬阳	102
4	《自动化天地》2015	其他	陈昭	100
5	A Integrated Vehicle Health Management Framework	会议论文	张桂刚	82
6	紫冬讲师礼仪系列 1——第一印象重要性	影音	高思宇	65
7	Prognostics Health Management of Condition-Based Maintenance for Aircraft Engine Systems	会议论文	吕鹭	59
8	研究所十三五规划部署情况汇报	演示报告	杨鑫	54
9	《自动化天地》2014	其他	陈昭	49
10	紫冬讲师-云计算及大数据技术 1——Google 的核心技术	影音	郑楠	46
11	自动化所年鉴（2006-2013）	其他	陈昭	43
12	【紫冬讲师】礼仪系列 2——仪态和气质	影音	姜扬	35
13	自动化所标准 logo	其他	陈昭	32
14	excel 系列-身份证号信息处理小技巧	影音	孙伟	30
15	ESI 和 JCR 培训 PPT	其他	何薇	28
16	201604 动态资讯	其他	鲁宁	27
17	201605 动态资讯	其他	鲁宁	27
18	幸福工程宣传片	影音	李琳琳	26
19	自动化所通用软件采购申请表	其他	胡蓉	26
20	礼仪系列 3——气场管理	影音	高思宇	25