

· 科学论坛 ·

# 国家自然科学基金信息系统全在线 申请模式运行状况及改进设想

郝艳妮\* 李 东

(国家自然科学基金委员会信息中心, 北京 100085)

**[摘要]** 2015年国家自然科学基金项目首次实行全在线申请, 本文对今年申请阶段系统的处理情况进行总结, 对全在线模式下系统功能使用及用户行为数据进行了分析, 并由此提出对全在线环境下信息系统改进与展望的思考, 以期更好促进项目管理。

**[关键词]** 在线申请; 用户行为; 信息系统

在2015年3月5日十二届全国人大三次会议上, 李克强总理在政府工作报告中首次提出“互联网+”行动计划。互联网+是将互联网技术的最新成果应用于社会的各个领域, 形成更广泛的以互联网为基础设施和实现工具的经济发展新形态。具体到国家自然科学基金(以下简称: 科学基金)信息管理, 互联网+科学基金将意味着科学基金信息管理的变革。

科学基金项目管理系统经历了几次较大的更新换代。伴随着新技术、新应用的产生和信息系统的不断升级, 科学基金项目申报管理也经历了几次较大的变革。在2003年以前受限于当时信息系统技术的发展水平, 项目申报填报主要是采用单机数据录入或通过软盘等介质导入的方式采集信息, 因此采集的信息非常有限仅包括申请项目简表中的几项基本信息, 申请书的全文也只能接收纸质件报送。从2003年开始随着计算机与网络技术的飞速发展, 国家自然科学基金委员会(以下简称基金委)开始推行项目全文信息采集, 项目申报管理中建立起由申请人下载特定格式的申请书模板离线填写申请书内容, 再由依托单位通过网络集中上报申请书的管理模式。由于在申请书中应用了WORD+VBA宏做为数据的采集技术使用数据采集更加准确规范, 通过网络报送申请书实现了全文电子化收集。最新一代科学基金网络管理系统的建设目标之一是“对基

金委外部, 系统应具备支持10000人同时在线操作的能力”, 加大在线的功能是新平台的重心。从2012年部分类型的试用在线申报到2015年开始全面支持所有类型在线申报项目。

在线填报对项目申请人来说只是一种填报方式的改变, 而对信息系统来讲, 全在线填报申请对信息系统的要求更高, 特别是采集到的数据量更大, 原先很多离线方式无法收集到的信息, 通过在线方式在一个月甚至两周之内全部上传到信息系统中。本文旨在通过分析在线申报模式的改变对信息系统的要求, 以及这些数据对信息系统功能带来的影响进行思考, 并提出初步的系统解决方案。

## 1 2015年全在线申请情况与总结

项目在线申请始于2012年, 当年选择了项目申请体量较小的项目类型如国际(地区)合作类、国家杰出青年科学基金类项目试行在线申报项目, 只有几千份申请书; 在2013年科学基金在面上项目、青年科学基金项目、地区科学基金项目以外的其他类型项目中推广了在线申报的方式, 有1.3万份申请书通过在线方式提交; 有了这两年试行的基础后, 在2014年除了面上项目与地区基金项目外其他类型项目全部实行在线申报, 在线申请的项目数量达到了8万份之多, 超过了离线方式申请的项目数量; 2015年开始所有类型的项目全部实现在线填报, 集

收稿日期: 2015-11-16, 修回日期: 2016-03-01

\* 通信作者, Email: haoyan@nsfc.gov.cn

中接收期间共有 16.56 万份。在线填报大大增加了申请书信息的数据规范与申请书全文内容的有效管理。

在考虑项目全在线申报的同时,首先要将项目申请人员的身份信息尽量规范与完整,这样才能更好地体现出互联网+时代科学基金信息系统以人为本的理念。所以 2015 年在申请项目之前,对人员信息做了一定的规范与整理,并在申报时将信息系统中增加了以证件号作为人员唯一身份认证的限定。这样为今后信息系统以人的关键信息为依据进行管理奠定了良好的基础。

针对 2015 年全部项目在线申请的要求,信息中心认真总结前两年部分项目类型在线申请的经验。从 2014 年底开始,信息中心就组织对系统的整体架构以及申报功能进行部署和多轮的压力测试,并针对政策的变化对系统功能上进行了相应的调整。整体上信息系统顺利支撑了 2015 年集中接收申请项目的业务要求,在 2015 年集中接收期间共接收项目申请 165 598 项,比 2014 年集中接收期增长了 9.35%。

在线申请对信息系统的处理能力提出了更高的要求,整个集中接收期处理了 18.3 万份申请,其中 16.5 万份申请正式上报基金委,有近 1.9 万份申请(占总处理量的 10%)是无效申请,但是这些项目同样占用了整个信息系统的各项资源。与离线申报方式相比在线申报系统的处理的冗余量是一个数量级的提升,2011 年是离线申报方式下集中接收申请书最多的一年,当年系统共处理 17.2 万份申请书,系统单日处理申请书转 PDF 量的峰值是 3.19 万份申请书,占当年总申请量的 21.7%;而 2015 年集中申报期系统的单日申请书转 PDF 的峰值是 28.47 万份,是 2015 年系统处理申请总量的 1.56 倍。在线申请的高峰还有多种功能点重叠的特点,在系统转 PDF 出现峰值的时刻系统同时在线人数、申请书内容保存、申请书相关文件处理等关键功能点都出现相应峰值,所以信息系统需要充分考虑到冗余处理量以及同时高并发的处理能力。

由于 2015 年是全在线申请的第一年,在集中申报过程中也出现了一些问题,如高峰期系统登录困难、申请书 PDF 转换不及时、网络拥堵等现象,所以在线填报对信息系统除了提供功能性的支持外还需要考虑到系统并发处理能力以及与网络环境、硬件容量等多个方面的合力配合的问题。在集中接收工作结束后,信息中心就今年申请阶段的事项进行了

认真的总结,并启动了下一年度集中申报的准备工作。

## 2 全在线申报模式下的用户行为数据分析

在线管理是科学基金网络管理平台建设之初所定的重心之一,在大数据技术的背景下,利用网络环境开展相应的管理工作是科学基金精细化科学化管理的趋势。全在线申请也是科学基金项目在线管理的一项重点功能之一。在全在线申请以后,我们可以很好地利用网络环境收集到与科学基金申请管理相关的数据,从而为今后的项目管理提供相应的支撑。我们收集数据的角度可以从信息管理与系统处理两个方面进行分析。

信息中心对 2015 年集中接收期信息系统中的主要功能点数据进行了统计分析,如表 1 所示。

表 1 2015 年集中接收功能点使用统计

分析点	总量	平均每天	平均每份	单日最高	日期
新增用户	91 933	1 108		5 488	1/16
发送邮件	540 282	6 283		151 077	2/12
系统登录	96 684 591	1 124 240		1 600 133	3/10
同时在线	617 100	7 260		55 628	3/10
申请保存	5 902 138	69 437	30 次	526 448	3/11
文件数	1 988 853	25 499	12 个	235 069	3/10
申请成果	3 430 387	39 889	20 个	347 223	3/11
PDF 请求	2 691 145	32 038	15 次	284 773	3/10
申请提交	411 384	4 840	2.4 次	46 884	3/11
单位退回	245 370	2 887	1.4 次	31 250	3/11

我们重点对申请项目填报过程中的添加用户、邮件通知、系统登录、保存申请、生成 PDF 文件、提交申请、单位退回申请等这几大关键功能点的数据进行了统计分析。通过表 1 的数据分析可以看出系统整体的压力点集中出现在 3 月 10 日和 11 日两天,这个数据与今年集中接收的实际工作也是相吻合的,从而为下一年度申请资源的合理配置提供参考依据。通过对项目申请过程中几大功能节点处理数据的分析可以发现新增用户与邮件发送这两个功能节点比系统峰值的时间提前对系统的冲击小,同时看出这两项功能的并发压力不是很强,这也是这两项功能支持的业务位于整个申请管理的前端性质所决定的。系统压力峰值最高时刻同时在线人数达到 5.56 万,系统用户登录与 PDF 转换请求这两项也达到了峰值,反应出这几项功能点的并发有连锁反应性,并发时间点也很接近,几乎没有延迟,这几

项功能点也是申请书填报过程中最为关键的几步,相关性非常紧密,因此对系统的并发处理能力有更高的要求。从系统功能点处理量分析看出平均每一份申请书保存 30 次、PDF 生成请求 15 次,平均一个项目包含 12 个文件,平均一个项目提交 2.4 次等,从这些数据上反应出用户的一些使用行为。

在线填写时申请书保存是与服务器交互最多的操作之一,平均每份申请书交互 30 次,也给信息系统带来了更多冗余的处理量;还有一项功能是申请书转 PDF 的操作也给系统带来更多的冗余处理量。这也是做为一名普通项目申请者申报项目时的正常操作,申请书的内容是要进行仔细调整与修改后才可以最后提交。因此信息系统需要把这个过程中的功能再进行细分,然后分析合理性,使得系统的资源更好地被使用。

在申请高峰期结束后,我们对科学基金项目管理系统中与申请管理相关的系统处理日志进行了梳理,对其中申请书 PDF 生成数、PDF 的大小、PDF 的页数、项目成果数的记录进行了分项目类型的整理,具体如表 2 所示。

从表 2 中的数据可以看出不同的项目类型用户的操作习惯也不尽相同。以 PDF 请求申请次数为例,面上项目、青年科学基金、地区基金的数据基本趋于平均值,国际(地区)合作与交流项目、应急管理项目的数据低于平均值;国家杰出青年科学基金项目、优秀青年科学基金项目、创新研究群体项目、重大仪器等都明显高于平均值。但是从 PDF 文件的

页数与大小的数据来看,人才类项目如国家杰出青年科学基金项目、优秀青年科学基金项目、青年科学基金项目类明显低于平均数,资助金额比较大的一些项目如重点、重大、重大仪器类的数据高出平均数很多。这些数据体现出根据各类型项目的管理特点,用户在具体操作上的侧重点也是不同的,人才类项目更注重内容的简练,研究类项目更侧重内容的丰富性。因此如果今后在申请书管理中需要加入对申请书大小控制的限制,就需要充分考虑到各类型项目的特点以及用户的操作习惯。

科学基金的面上项目、青年科学基金、地区基金这三类项目申请量最大,从申请书 PDF 文件的大小来看,其 PDF 文件大小的最大数与最小数的差分别是 57.32M、37.47M 和 34.53M,也是相差很大的,但是每类项目 PDF 文件的平均值却相差不大;而重大项目的申请书,PDF 文件的平均值最大;这一点也提醒我们,要限制申请书 PDF 文件大小时,需要按照项目类型分别设定,才能做到更好的个性化服务。

从申请书成果数的分析也反应出不同项目类型的特点,重点、杰青、群体等资助强度较大的项目申请书中的成果数量也大。

### 3 信息系统的改进与展望

通过分析,我们认为全在线申请的实现,可以从更多角度来获取用户行为数据,进而加强系统的数据分析与服务功能。例如用户提交一份申请书的平

表 2 2015 年集中接收系统处理分项统计

项目类型	生成 PDF 数		PDF 页数			PDF 大小(单位 M)			成果数	
	最大	平均	最大	最小	平均	最大	最小	平均	最大	平均
面上项目	278	15.32	135	12	39.61	57.83	0.51	2.04	299	20.12
重点项目	133	18.63	373	15	58.62	21.69	0.55	2.91	348	42.02
重大项目	46	16.86	119	33	65	6.48	1.69	2.99	41	20.67
重大研究计划	69	11.54	101	23	41.42	13.64	0.88	2.13	249	23.9
国家杰出青年科学基金	208	17.7	198	17	38.06	19.56	0.57	1.94	428	83.78
创新研究群体科学基金	77	18.6	172	24	64.55	21.4	1.11	2.34	467	79.4
国际(地区)合作与交流	378	11.49	185	5	47.21	17.98	0.36	2.37	10	4.94
联合基金	124	14.49	139	16	37.48	15.15	0.57	1.8	265	22.68
青年科学基金	266	15	98	11	33.82	37.97	0.5	1.78	164	11.19
地区科学基金	206	14.08	113	12	36.21	35.07	0.54	1.82	151	15.64
海外及港澳学者合作研究基金	80	15.47	114	21	35.19	19.24	0.73	1.69	253	28.85
国家重大科研仪器研制	119	18.51	189	20	69.52	20.77	1	3.49	275	43.55
优秀青年科学基金	276	15.94	135	16	33.28	20.58	0.55	1.87	308	45.31

均在线时间,通过这个数据再结合用户的年龄、地域等属性进行深入分析,就可以得出为用户做一些个性化的帮助服务。对生成 PDF 比较频繁的数据再进行深入分析,比如两次请求生成 PDF 的时间间隔,从而发现一些用户的操作习惯同时可以从信息系统的功能上加以调整,比如控制请求生成 PDF 的时间间隔,一方面给用户一种友好的提示另一方面也可以减轻服务器端的处理压力。

另外,我们还可以从申请书基本信息内容的调整上分析用户行为,例如在申请书填写过程中项目基本信息中的申请代码、项目名称、项目类型等要素的更换频次,看出申请者最看重的信息,这些数据信息经过一定时间的积累后可以在后期的辅助匹配专家过程中提供相应的参考。

从系统中提供的常见问题被访问的次数可以看到用户端存在的主要问题集中在哪些地方,进而分析系统功能需要做什么样的调整。

由于 2015 年是全在线申请项目的第一年,从用户习惯与信息系统的支撑上都还在试用期,在系统

功能上有很多需要改进的地方,特别是如何利用与分析用户行为数据更好的为科学家服务还需要几年的数据积累。

在科学基金信息化建设的“十三五”规划中明确指出,未来五年科学基金信息化管理将包括六个平台,其中的业务应用平台、服务共享平台、知识服务平台,都是需要以信息大数据作为支撑的,信息系统将从数据服务向知识服务、从服务单位向服务社会迈进,我们将进一步提供功能丰富、安全可靠、持续发展的信息系统服务环境,促进科学基金信息管理。

### 参 考 文 献

- [1] 贺赛龙,郭维森,王其冬. 浅析国家自然科学基金项目管理流程再造. 中国科学基金, 2006, 20(2): 105—109.
- [2] 李东,郝艳妮,李建军. 2014 年国家自然科学基金青年科学基金项目在线申请试行的总结与思考. 中国科学基金, 2014, 28(5): 362—365.
- [3] 王斌,张艳. 浅谈国家自然科学基金信息交流的精细化管理. 中国科学基金, 2013, 27(4): 231—233.
- [4] 李朗. 关于国家自然科学基金网络信息系统的思考与建议. 中国科学基金, 2008, 22(1): 52—54.

## The operation status of online applying mode and improvement ideas of NSFC information system

HaoYanni      Li Dong

Information Centre, National Natural Science Foundation of China, Beijing 10085

**Abstract** In 2015 National Natural Science Fund Project implemented the full online application for the first time. In this paper, I, we sum up the treatment on the application phase of the system, and then analyze the system functions and users' behaviors on the full online mode. At last we put forward to information system-wide improvements and prospect, so as to promote scientific fund project management better.

**Key words** online applying; user behavior; information system