



(本文选自微金融 50 人论坛演讲汇编系列之一：微金融的基础设施，独家推出，转载请注明出处)

一、余额宝与大数据

近期来，大数据备受整个媒体和业界的关注。

余额宝与大数据：互联网金融的大数据基因



图1 大数据的4V特征

笔者更希望把大数据的“大”字先去掉，而谈谈天弘在以数据为驱动的情况下在整个余额宝业务的产品管理过程中做的一些事情。

首先需要阐述的是产品管理的概念。产品部原本在基金公司里面，尤其是余额宝业务之前，承担的工作有三条，第一是战略研究，给基金公司找到发展方向；第二是产品定义和产品管理，给证监会写申报材料，报证监会批准；第三是营销支持，给市场体系做文案，让渠道客户经理了解产品。而在互联网金融下，尤其是余额宝产品里，产品部的职责在实践中得到了拓展。第一和第三个职责还是给公司找战略方向和给营销找支持，只不过尤其营销端面对互联网有所不同，战略研究紧密结合电商的发展和技术的的发展，注重利用现有数据对业务进行指导，营销支持更通俗化地去解释。

产品管理在一般的基金公司产品部都是只管生不管养的，孩子生下来，拿到报汇的批文就可以了，而天弘基金互联网金融产品部承担的职责实际上是跟互联网企业的产品岗位非常类似，一个产品经理在这里承担是整个产品从生到养的全生命周期管理。互联网思维很大程度上就是数据思维，要用数据去说话。在整个管理中可能遇到一些问题，在实践中也带着一些探索。

做数据的分析首先要解决的就是客户是谁的问题。传统银行或者传统金融普遍存在二八定律。只有知道自己的衣食父母是谁，利润来源在哪，才能知道怎么做业务，以及如何对待客户。用图来说明天弘客户的情况。简单的男女比例，好像很简单，加起来百分之百，总的来说可以看到上线之初和后来中间有一个分界线，后来就走平了，分界线很明显就是去年双11。余额宝刚上线的时候男性的比例是非常高的，在双11之后女性比例追上来，一直保持平稳，这是资金保有。如果用购物订单的消费图来说，这两个曲线最后是交叉的，背后这意味着大约有超过10%的用户是女性用男性的账户买单，这是应当被了解的用户行为，也就有了现在亲密付业务的灵感。

余额宝与大数据：客户是谁

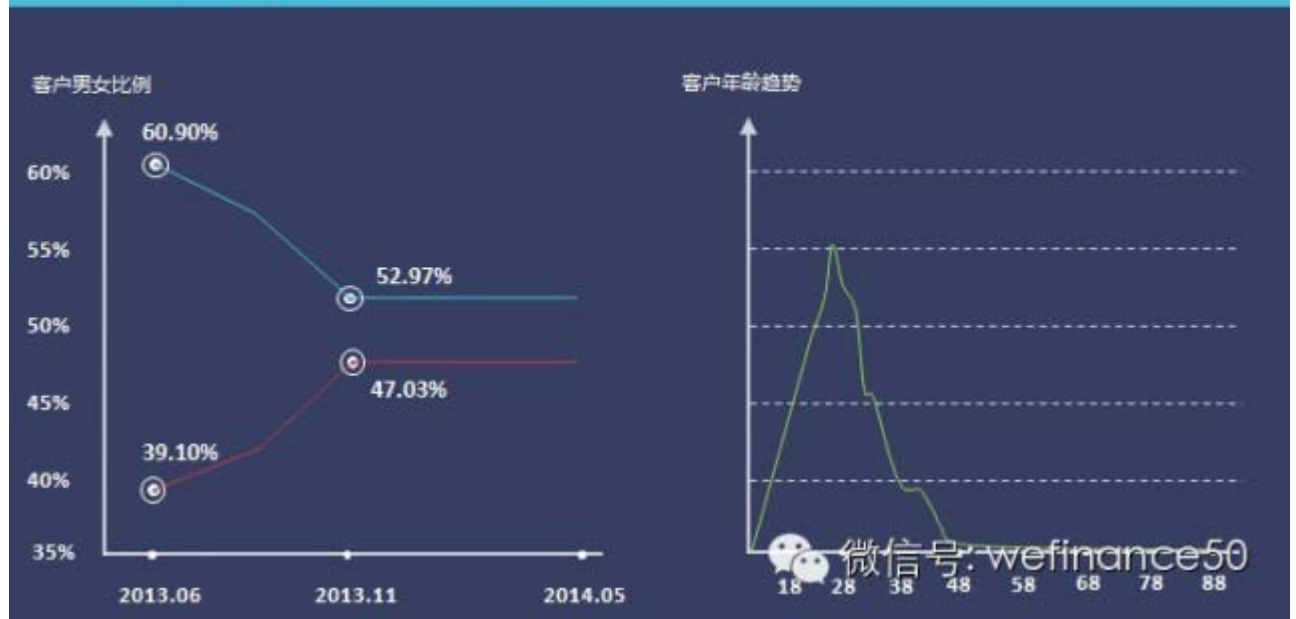


图2 余额宝与大数据：客户是谁

天弘在整个营销过程和推广过程中也通过自媒体和大家分享了一些基于数据交易额和数据挖掘的东西，比如关于前文的例子就有一个有趣的话题---趣味性挖掘结果表明去年双11新疆的男人最疼老婆，其男性账户为女性内衣支付额最高。天弘公司渠道上平均年龄客户是37岁，这只是一个平均数。利用阿里的金融云建立大数据中心之后，客户的年龄分布可以被很精确地刻画出来，这体现出对数据和业务的实时掌握。后文将举实际应用的例子。

除了解决客户是谁的问题之外，还要分析客户在做什么。在余额宝客户群体中其实无外乎把整个消费分成几大部分。其中，信用卡还款接近三分之一，剩下的主要是购物。余额宝被称为网购神器，从销售笔数来看整个客户网购消费量很大。另外就是水电煤缴费。其他许多转账应用总的来说占的笔数都很小。信用卡、购物和水电煤缴费主要几个应用占了消费中非常大的份额。余额宝的消费属性逐渐增强。其作为互联网理财的工具，同时又打通了消费购物支付，其中掺杂了金融属性，即对收益敏感的理财属性，与电商属性。余额宝一直强调在营销过程中引导客户认知其消费属性，告诉客户钱进了余额宝是能花、随时花的。反而是取钱的时候会受到一些限制，每人每天只有5万块钱的T+0的额度。

余额宝与大数据：客户在做什么



图3 余额宝与大数据：客户在做什么

整个上线以来，客户消费比例不负众望一直处于缓慢提升的趋势，这也是证监会和央行领导们关心的余额宝业务规模这么大流动性管理如何做背后的天弘信心满满的原因。今年6月份，行业所有货币基金面临非常大的赎回压力，很多公司遭遇流动性困难。无数人关心余额宝有没有遇到比较大的赎回，余额宝业务对客户的引导一直以消费属性、电商属性为主，而不宣传收益率的属性，所以当时遭遇到的还是比较平稳的态势。这和一些竞争对手产生非常大的区别，当时几家在微信平台上包括交易所场内都遇到比较大的流动性的问题，单日赎回额都在200亿左右，实际上当时有些银行不同意提前支取存款，外部比较平静，行业内部面临比较大的压力。3月以来余额宝也面临很大的来自证监会和央行的监管压力，但余额宝的赎回却一直保持平稳的态势。

二、大数据支撑下的互联网金融产品管理

数据分析的应用很广泛。

第一个例子是流动性的管理预测，这是从余额宝上线的第二周就开始上线运行的，是天弘跟阿里数据的团队一起建立的共同分析的机制。这个分析机制的开始是因为双方约定必须要有一个最晚的清算时间。涉及到不同的银行，涉及到交易所场内账户，而很多交易所场内账户受终端本身清算能力和清算效益的影响，经常晚上才能清算。双方根据实时数据分别对下一个交易日和下一个月的申购情况和赎回情况做出预估。每天误差三千万左右，从现在的体量来看，几百亿基金，每天头寸估差在三千万的水平，其误差比例是非常小的。开始的时候阿里的数据分析团队非常强，而天弘基金人才缺乏，启用量化分析师做流动性的管理预测，结果不合适。首先在数据分析应用的方法和对业务的及时响应都存在问题，后来请了有互联网背景的周卫国来做数据分析，从此天弘的大数据真正到了正规化的道路。开始的第一个月阿里给的预测比天弘误差小。第二个月阿里给的误差大概介于六千万到1.2亿。天弘模型优化后误差稳定在三千万，经过几次磨合，从以阿里的预估数据为头寸基准到了双方共同商量再到三个月之后变成了阿里认可天弘的数据分析和预估能力。持续误差在整个三个月中一直保持着三千万的水平，而基金规模已经上千亿。每天头寸误差，即第二天要进多少钱和出多少钱，仍然保持平均三千万的量级上。后来完全转化成以天弘自己的独立分析能力应用分析了。



图4 余额宝与大数据：我们该做什么

第二个例子是应用场景拓展。传统银行通过形成一个基础账户的属性把客户留住：客户偿还房贷、车贷需要用钱，离不开基础账户。类似的，天弘强调可用和电商属性，从而做了应用场景的拓展，包括定期转出的功能的推出。看起来，做这种定期转出，是把自己的客户的钱往外投。其实，根据实际调研，有几百万用户每月有金额较大的固定转出，很可能是偿还房贷和车贷的需求。只有让客户实际的需求，不用离开余额宝账户，才能抓住客户黏性。余额宝的各种支付场景，包括和定期转出的场景都是这样的思路下逐渐拓展的。

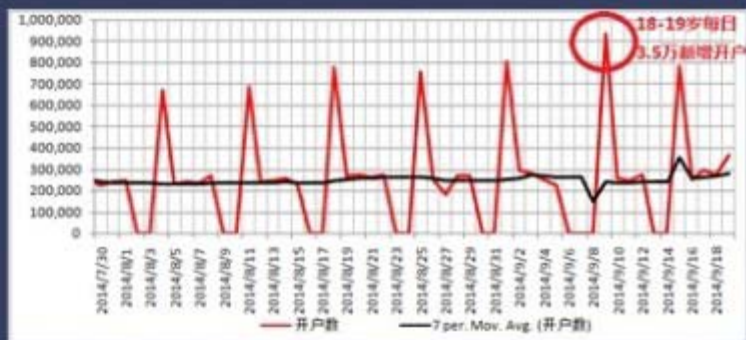
第三个例子是指导运营。天弘每天有数据的监控，每周有生产数据的分析会。数据中心是一个基础平台，但数据中心只是在诸如天弘周年庆等重要场合做综合的客户画像和各种分析，不可能实时提供很大的分析和数据。因此天弘要求从客服到产品到运营，每个小部门都要有一个业务 BI，每天看实时交易数据。9月初天弘产品的业务 BI 提出开户数据有异动，每天增加三万到四万。通过请求大数据中心精准分析数据后发现增量开户数据中有明显的年龄分布特征，异常数据 90%是来自 18—19 岁。而 9 月初正是大学生开学季，18—19 岁很可能是大一新生。很多大学新生入学前在家没有自己管过钱。第一次离开父母，踏入社会第一次自己掌管钱。余额宝在大学用户中普及率比较高，通过交互迅速有感知效应，有可能带来大量的客户增长。天弘针对精准的营销活动，迅速地联系校园渠道，找了大多数学校的学生会，一些定点的社团，进行新生的宣讲。一共筹备三天，在发现数据异常情况下，第四天开始新生宣讲，大约是持续了半个月热度。这是快速反应和机构设计层面的成果。没有动态监控，没有每个业务部门对数据本身有敏感性的观察者和定期的交流机制，就不可能有基于数据驱动的精准的运营和营销活动。

大数据支撑下的互联网金融产品管理

刻画客户需求

提升投资管理

指导业务运营



微信号: wefinance50

图5 大数据支撑下的互联网金融产品管理

三、互联网金融业务的数据化生存

在数据应用方面，有三个步骤。第一就是要做一个数据采集和整合。天弘的交易系统利用阿里金融云的环境，数据实际上分在很多组的交易服务器上，在收盘的时候才会把数据再归总起来。开始有几个月都看不到归总的数据，在运营等方面像是盲人在打仗。因此第一个工作重点是要解决数据的采集与整合问题，把基于云的交易数据实时采集进来。数据中心当时提的目标就是每五分钟精准数据出来，基于这五分钟数据有异常监控，年龄、地域知道什么出现了异常。对数据的分析是以去年双11为先导。去年双11是天弘第一次经历电商大的节日。第一次和阿里BI讨论是在8月底，基金一共有500亿的量，阿里提出要保证双11的时候，有200多亿支付的可能性。天弘成立了基于数据分析的小组，发现数据服务各个方面，所有座席按照当时业务量比例，满足不了。有赖于自己在互联网思维上面的熏陶，最后天弘开始启动云小二，就是客户帮助客户。传统基金公司有自己的客户、劳务派遣和外包三层结构。天弘招募了云客服，其本身就是客户，能帮助客户解答问题，有问题可以求助其他更高级客户。启动云客服之后，咨询量里面90%是云客服承担。采用网上自动排班机制，把一周之内客户响应可接答的比例放大十倍。从招募到排班仅用一周的时间。利用阿里云的金融的IT的响应是一个道理。为什么上云？当时考虑在云上，比如用400台，高峰期可以再多用200台，扛过高峰期就不再续租多的200台了。

在数据分析上以双11的响应，包括投资、客服准备上的分析为先导，天弘跟阿里逐渐碰撞，不断壮大自己的数据对。去年对余额宝在大数据的交易分析，天弘建立了一个历史客单价的历史参与比例，每个人的客单价是多少，去年这些用户有没有在一年一内有没有消费，有消费的定为活跃用户，参与大数据的比例有多少。基于这个链条评估结果，最后跟实际双11的60多亿的支付结果非常相近，差了四个多亿。应用数据这块建立了生产数据分析的机制，要求所有部门领导和业务BI都要参加。若时间有冲突，二者必须有其一参与到业务分析会上。

其实数据分析或者是以数据思维为先导的过程，整个服务不只是上文提到的客服多少座席，是否起用云客服，云客服要多少业务量的问题。云推广，举了大学生的例子。在线上做活动，财务要如何支持也是通过数据分析提前给出方案。每个运营活动要有运营效果，除去正常干扰趋势因素之外，都会经历导入、平台、衰减期。运营活动在导入期的时候截止，钱就花完了，再申请钱，这是新的导入期。这就白白浪费了导入期的效果，但是也不一定，每一种活动可能是导入期非常短，平台期非常长，衰减期非常短。基于效果评估中，管财务要多少钱，从财务角度就是每个月给钱。

十一之后富达请笔者做现金流管理，他们评估得到一个交易员在现金管理市场上，这里交易员包括基金经理和分析师，整个团队人员和他们资产管理比例差不多是150亿/每人。货币基金每天的交易量需要多少人管理跟数据的可视化程度有很大的关系。天弘的基金经理通过手机就可以看到未来一个月头寸的变化，预估出来是多少，实际上每天到期有多

少现金，会不会有趋势性变化；直接可视化的显示现在头寸是多少，预估有多少头寸需求，实际数据和预估数据差距有多少，最大交易量有多少。假如说一天交易量超过 350 亿，四个交易员忙不过来，或者交易员有休假，系统可以自动提示交易负荷较大，建议提高每笔业务的交易金额。



图 6 数据开源

天弘做业务数据标准化采集，越来越体现效率，还可以观察很多有意思的现象。去年 11 月金融机构给天弘报价出现推晚十天的情况，本来上午给一个报价，下午一点多如果觉得天弘还需要钱就开始提价，争取单子。去年 11 月开始他们提高报价的时间先推晚半个小时，具体哪个省级分行报价先晚，哪个后晚，持续有错后的现象，每次都是谁先谁后。最终是什么结果？大家是一个博弈，金融机构知道天弘每日交易量很大。上午试探性询价，报一个稍微低点的价，遇到真正钱紧，下午报高的价格。这造成了天弘金融机构交易伙伴托管行操作风险极大增加。整个清算压缩在最后一个小时、半个小时，肯定清算风险上升。天弘通过数据分析很早就发现这个，第一时间先等带头人来做这个事的人来谈谈。开始要讲究策略，跟他合作中找前三个开始实施的分支机构，找下个月到期集中的分支机构做重点的沟通，这个月不和该分支机构合作了。大家可以理解当一个分支机构行长看到支行同业存款一个月少了几百亿，尤其是遇到考核时点，将面临很大的考核压力。进行重点突破之后，跟随者一看领头者都把报价时间提前了，也就自然提前了。其他包括公关等等许多方面都是做了数据分析。

天弘是通过包括客户结构数据、客户交易行为数据和投资管理数据的企业内部数据，以及包括电商平台关联数据、社交网络数据和宏观经济数据的企业外部数据进行数据开源。天弘数据中心定义人员架构分为四层：数据科学家、数据探索者、数据架构师和数据业务员。发现规律的人、整合数据和搭建数据平台的人以及实施数据交互的人，这三个角色在团队中是非常重要的。通过整个构建，天弘现在是 IT、产品、运营、客服四个团队基本上都初步建立以数据分析为先导的工作的方式和工作的状态。希望将来产生更多的在数据引导下的成果。

互联网金融业务的数据化生存：数据团队



数据科学家



数据探索者



数据架构师



数据业务员



微信号: wefinance50

图 7 数据团队

（根据微金融 50 人论坛闭门会议整理而成，本内容分享者为天弘基金互联网金融产品部副总经理 李骏）