

中信项目信息收集

方楠

结论

推荐：持续跟进。

商务介入签署补充协议，确定开发成本边界，在过程中尽可能的回收占用我方人员的时间。

具体实施方案：

每2-4周上线一次，对于每次上线中出现的我方问题，给出解决的技术方案并做开发修改（驻场）。之后的无效时间不再驻场，直到下次上线并循环。

- 1、爬虫采集类需求的延期问题出现的概率很高，很难“一刀切”。**
- 2、爬虫类需求并不需要占用开发人员100%的时间。往往是碎片的和随机的。**

合同/标书条款

一句话总结：除了**验证码部分**之外，没有太大的隐患。

3.3.16 验证码要求

合作厂商应提供验证码破解技术及源码，需对以下的网站中涉及到的验证码进行破解，破解的时效需满足甲方实际的业务需求，涉及到的网站如下：

所有的**社保网站**、所有的**公积金网站**、天眼查、企查查、**国家企业信用信息公示系统**、团购类（美团网、大众点评网）、房产类（房天下、搜房网、安居客、链家网）等网站。



存在的问题：已经将所有**社保网站**开发完毕，但是客户暂时未要求进行**天眼查、国家企业信用信息公示系统**等我们做不了（开发成本极高）的验证码渠道采集。未来会根据合同条款去要求。

经过沟通，现在中信团队达成的潜在方案是。已经跟中信达成一致，先不采集除了社保网站外的其他网站，把UAT过了收到第二笔款之后，再去跟中信慢慢“磨”。能做的慢慢做，不能做的就反复沟通交换需求，一直磨到不做。

举例：这些网站部分验证码部分涉及《极验》验证码服务。

如果公司不以数据采集为主方向的话，这些技术积累意义不大。

这个条款是悬在我们头上的利剑，也是我们确定不能做的功能。

如果不能在商务上签署补充协议，那么也无法确定成本边界。

法务问题

一句话总结：违约成本很高，但是有上限的。

(1) 如果我们解除合同，后期就不再维护了。但是已开发部分的知识产权将会有争议。

(2) 关于违约金，如果解除合同，退还已付款项+合同金额的20%违约金+赔偿损失（需要对方提供证据），对方也可能主张每日0.5%的违约金，但是法律上规定违约金最高限为合同金额的30%。

(3) 我们验证码服务数个渠道无法破解，甲方可以据此拒绝UAT验收，我方需要支付违约金解除合同（违约金额度在2中说明）。至于爬虫服务的渠道破解是否构成拒收UAT的条件，跟具体行业有关。

违约是有上限的

爬虫采集类的需求对于银行比较敏感，起诉的概率很低。

利益矛盾

一句话总结：**甲方不是一个甲方，但都是我们“爸爸”**

- 1、本次上线失败，直接原因是行方的ES节点挂了一台服务器，导致数据搬运失败。
- 2、同时我们也有责任，因为3节点ES按照设计要求，其中有一台服务器挂了不应无法启动。
- 3、由于项目通过了SIT测试，即在测试环境上已无问题。只要可以成功部署在生产环境，通过UAT测试通过即可收回第二笔款。
- 4、中信生产室是我们唯一可以触达生产环境的路径，生产环境本身就对于我们来说很不友好（既有技术经验的原因，也有银行本身的原因）。势必要经历很长时间的反复上线。
- 5、下次上线我会尝试隐秘的取证，存留这个情况的证据。

未来的潜在问题

- 上线通过已经非常接近了。以两周上线一次的频率来看，在不更改需求的前提下，结合现有技术情况。我个人认为还需反复多次上线才能完成。
- 1、UAT测试完成后，**抓取渠道的反爬逻辑会持续变化**，我们的第三笔款（30%）要求上线内三个月后，第四笔款（10%）十二个月后，都要持续维护渠道稳定性。
- 2、现有代码的交接出现了断档。**只有几个核心人员**知道整体的来龙去脉。一旦出现反爬逻辑变化或者客户新增需求，修改代码将会比较困难。

重申结论

- 1、立即派遣一名有能力的商务人员，结合我们能接受的技术（成本）条件，去跟中信谈判，确定我们的开发成本边界，签订补充协议，至少要涉猎以下两条。
 - (1) 删除我们完全不能做的验证码渠道（比如极验3.0）
 - (2) 确定涉及验证码的免责开发条款。（有限维护）
- 2、开发成本边界确定后，每个以项目上线周期进行管理，持续跟进。
针对每次上限的失败的情况进行修复。