标题：你的App为什么知道你想买什么？

引入：案例2~3个

1. 你刚和朋友在微信上和朋友说要去日本玩，就在朋友圈看到了机票广告
2. 你妈妈刚给你打电话让你多喝牛奶，你就在淘宝上看到了牛奶推荐

提出问题：这些广告都是怎么来的，为什么这些App都知道你想买什么

当然，这些App并不会窥探你的生活，而是你的行为数据让App和其中的广告系统意识到了你需要这些

。

信息化时代，使用一个App必然要注册账号，你需要输入自己的名字、手机号、有时候还会补充性别、所在地等等，但这些只是微不足道的基础数据，更重要的是你的消费记录、打车频率、关注的公众号、玩过的游戏、理财习惯、有没有车贷房贷、有没有买过保险、是否发过的红包，这些行为最后会变成几千个事实标签成为你行为数据的一部分，从多维度将你描绘出来。

分题：事实标签是怎么来的？——你的App对你的隐私做了什么

当你注册微信的时候，你需要阅读一份8300多字的隐私协议，必须勾选“我已阅读并同意上述条款”，才能进行下一步，但我们都知道，其实你并没有读，也不知道你究竟同意了什么条款，理论上，只要你同意，你在 App里填写的住址、电话号码、聊天记录、手机里的通讯录、照片、短信、通话记录都可以被App们自由地采集和分析，它们对你的隐私数据到底做了什么？

首先我们要知道，App究竟能采集到你的哪些信息，在Android提供的开发者文档中，我们可以把App可以获取的权限分为普通权限和危险权限，普通权限有140项，不需要经过用户同意就可以获取，比如控制你的振动传感器和日历数据；危险权限分为9组，比如录制音频访问短信等等，而这些授权都是一次性的，比如你需要通过App发送语音打开了麦克风权限，那么这个App在理论上是可以一直在你的使用过程中录音的，这时我们就需要分清App权限和App可以采集的数据，在大多数App的隐私协议里，都会说明它们不会偷听和偷看，但你使用麦克风时就不好说了，比如这份报道：苹果会对1%的用户采集他们和Siri几秒的对话录音用于改善Siri的识别准确率，但除了这种用于改进产品的情况，目前并没没有证据可以证明App会偷偷上传分析你的录音或者照片，你的数据更常见的用途是生成前面提到的事实标签给数据分析师和广告主的用户画像，也正是这些用户画像模型，使得App知道你要买什么。

接下来想和大家聊一聊隐私安全的问题，互联网公司可以这么分析数据的前提，是你同意了这一切虽然你很可能没看过，但它们要守住的底线是——不能通过大数据反推出你的个人信息，举个例子：假设我们有100个关注者，在后台我们可以看到百分之五十来自上海，百分之四十来自陕西，百分之十来自湖北，这时如果我们移除一个关注者甲，剩下99人的地域比例就变成了百分之50.5，百分之40.5和百分之9，这样甲的隐私就暴露了，我们可以轻易地推算出他来自湖北，但实际情况中，工程师会使用如差分隐私等方案在统计数据库查询时，最大化数据查询的准确性同时最大限度减少识别其记录的机会，核心思想就是对数据加入噪声，让有没有甲都输出差不多的结果。对于当下更普遍的隐私泄漏问题，互联网公司们的做法通常是“考证”，其中最重要的是ISO/IEC 27001信息安全标准，作为全球公认的最高信息安全标准，它包含了信息安全、人力安全、物理安全在内的14个模块（信息安全策略/信息安全组织/人力资源安全/资产管理/访问控制/密码控制/物理与环境安全/运行安全/通信安全/系统获取、开发和维护/信息安全事件管理/业务连续性管理的信息安全方面/符合性）。2014年ISO在27001的基础上提出了27018，一项在隐私问题上更严格的扩展标准。以网盘为例，像百度网盘这样的企业就需要搭建“数据保护权限系统”，“数据脱敏处理算法”等更复杂的隐私保护体系，这也意味着你的App为了保护你的隐私得花更多钱。在中国也只有BAT这样大厂旗下的产品比如百度云、阿里云等服务通过了这两项认证，这样我们才可以说我们的数据是相对安全的。那作为个体，我们对我们的隐私安全可以做什么努力呢？这是几条建议

1. 不要随意下载注册App。
2. 通过手机的权限管理关闭软件不需要的权限，如通讯录、短信记录、通话记录、麦克风、位置信息、上网记录等权限都尽量不要开放给APP，或仅允许App在前台使用时拥有上述权限，最小化的防止APP来采集个人数据。
3. 开启不追踪功能（以ios和chrome浏览器为例），关闭广告功能或使用广告屏蔽软件（以小米、Ad Guard为例），定期清理浏览器cookies、浏览记录