

数据的价值与交易

@北冥乘海生

戏说互联网思维之“三个不要”

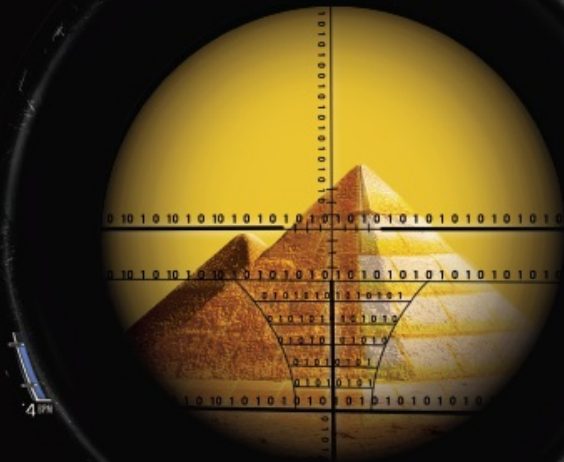
- “不要钱”
 - 免费倾销加后向变现的商业模式
 - 所有能够传播信息的商品，售价都会趋向其边际成本
- “不要脸”
 - 无底线迎合用户的产品与营销方式
- “不要命”
 - 用期权和价值观让程序猿在疯狂状态下全天候工作

COMPUTATIONAL ADVERTISING

计算广告

互联网商业变现的市场与技术

刘鹏 王超 著



中国工信出版集团

人民邮电出版社
POSTS & TELECOM PRESS

计算广告 COMPUTATIONAL ADVERTISING

在线广告市场是比较复杂的体系，它贯穿了互联网生态链的各种角色。刘鹏的著作既从商业角度介绍了在线广告，也深入到了广告的技术和算法层面，还包括对于一个大规模竞价市场在市场设计方面的相关原理和优化机会。这本书不仅是了解在线广告市场的途径，也是了解互联网商业和盈利模式设计的窗口，同时对于互联网产品设计会有很好的参考作用。

——王华 (@超凡Derek)，阿里巴巴副总裁，阿里妈妈负责人

如果说当年Yahoo!广告首席科学家Andre Broder在斯坦福开设的课程第一次把计算广告学作为一门学科，那么这本《计算广告》堪称这门新兴学科的第一本教科书，因为这是首次全面系统地阐述计算广告学的著作，涵盖了商业逻辑、产品结构、关键技术、工程实践和应用实例。在内容结构编排上，本书由浅入深，从宏观背景到技术细节，从经典的搜索广告到最新的实时竞价，既适合作为从事在线广告的商务运营人员的参考书，也适合作为一线技术开发人员的实践指导。

——余凯 (@余凯_西二旗民工)，百度研究院副院长，深度学习实验室主任

计算广告学纷繁复杂，刘鹏以一个科学的实践家的态度抽丝剥茧，全面、系统地阐述了其技术架构与产品生态，为入门者普及了概念，为从业者开阔了眼界。

——罗征，腾讯广告通总经理

在互联网深入改造传统行业的进程中，所谓“羊毛出在猪身上”的后向变现商业策略至关重要。市场上并不乏关于这一策略的推崇和讨论，不过本书从实践的角度出发，对其中关键的商业逻辑和产品结构作了全面的梳理，而这些对于商业化战略的落地至关重要。因此，我推荐所有正处在互联网化变革行业的从业者们阅读本书，结合各自的知识背景和行业问题，相信你们一定可以从本书中找到有指导意义的内容。

——陈彤 (@老沉)，小米内容运营和投资副总裁

我们的团队花了大量精力寻找和调研国内外的相关资料和文献，却一直苦于没有一套相对完整的知识体系来帮助我们的业务和技术人员加快学习、少走弯路。因此，我期待这样一本书的出现至少已经有两三年的时间。作为一门方兴未艾而发展迅速的新兴产业，我相信刘鹏的这套科学且实际的知识体系，将对数字营销领域的同仁有着非常大的指导价值。程序化营销、大数据应用是发展异常迅猛的产业，涉及大量的产品、工程以及算法，也涉及相当多的商业逻辑的深刻洞察。而刘鹏通过对产业内在逻辑的诠释和推演贯穿程序化营销产业的发展，对我们真正把握大数据给各个产业带来的深刻变化和影响，具有深刻的指导意义。

——杨伟伟 (@杨伟伟)，360副总裁，聚效广告CEO

封面设计：北京新锐设计

分类建议：计算机 / 大数据

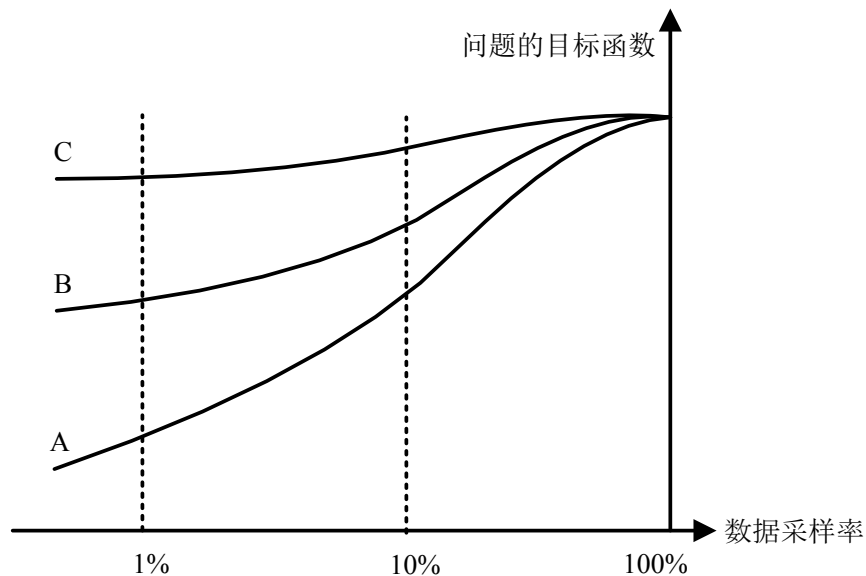
人民邮电出版社网址：www.ptpress.com.cn

ISBN 978-7-115-39249-7



9 787115 392497 >

什么是大数据问题？



A: 典型大数据问题，解决问题的效果随着采样率降低显著降低，例如计算广告、个性化推荐等

C: 一般数据分析，非大数据问题，解决问题的效果在采样率降低时变化很小，例如各种洞察、单维度统计等

B: 过渡类型问题，解决问题的效果随采样率降低温和下降，例如文本主题分析等

两类数据应用：洞察与自动化

- 洞察(Insight)
 - 全局或局部统计性的信息（统计数据）
 - 例：财务报表、人口统计、百度迁徙地图等
 - 主要用于宏观决策支持，面向领导和运营人员
- 自动化(Automation)
 - 个体的行为特征信息（行为数据）
 - 例：定向广告、个人信用、企业信息等
 - 主要用于微观业务实施，面向机器和销售人员
 - 无底线迎合用户的产品与营销方式

数据变现基本原理



流量价值

10000

+



6000

+



6000

||

数据价值

2000

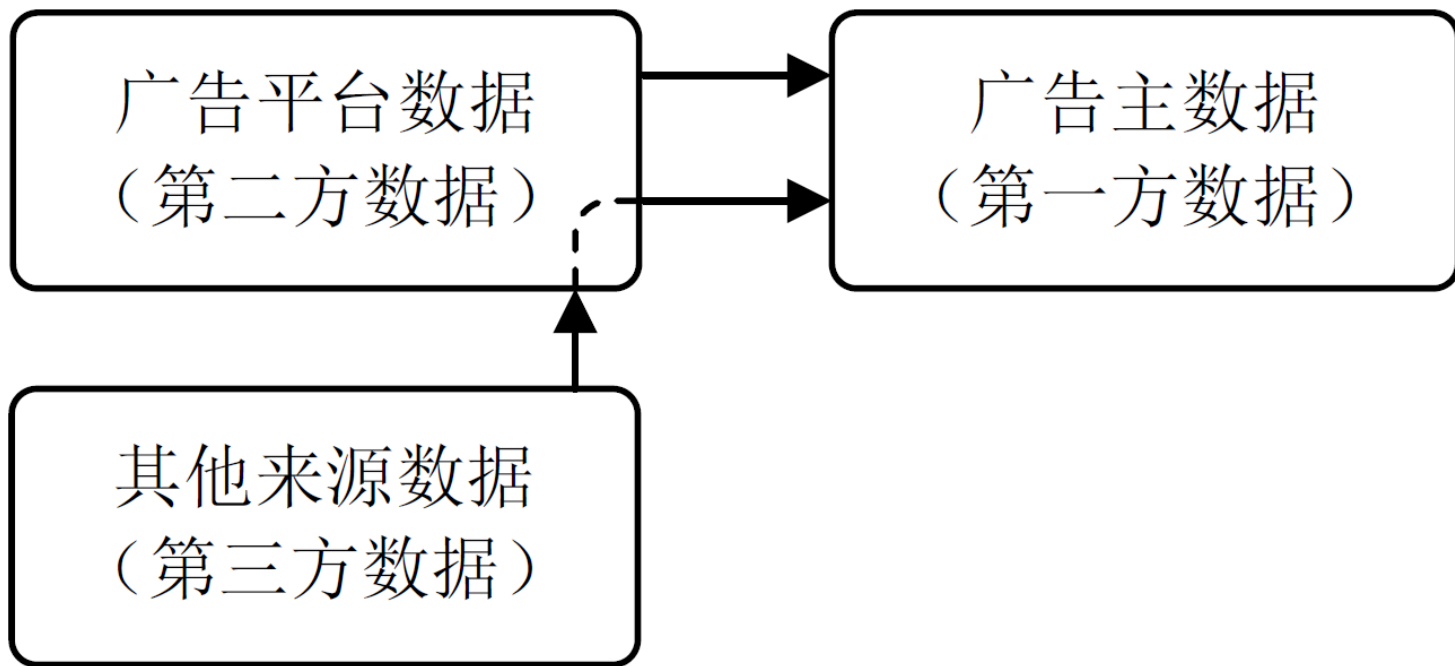
=

12000

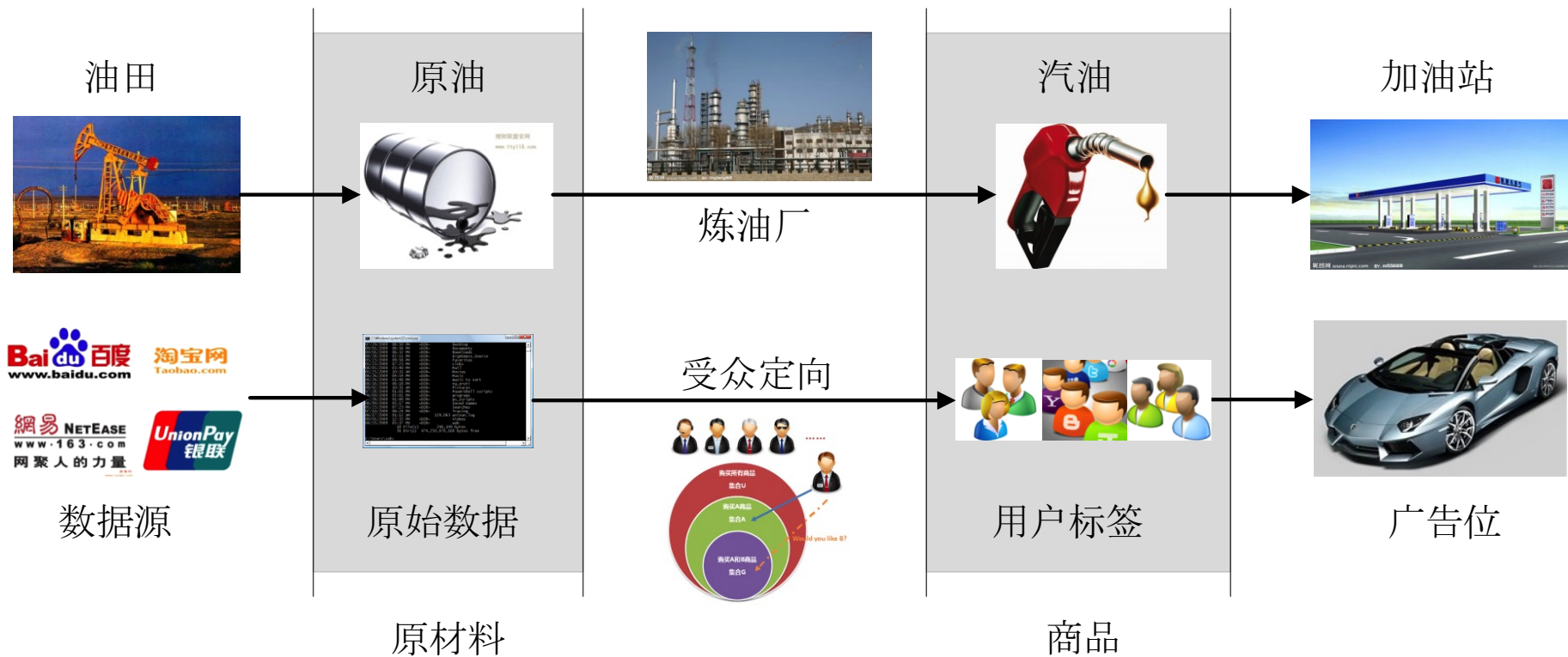
行为数据对广告市场的支撑



三方数据与数据交易



数据的价值与地位



数据交易所(Data Exchange)

- 目的：
 - 聚合多种原始行为数据，加工成统一标签后，在公开市场上售卖
 - 聚合多种加工后标签数据，在公开市场上售卖
- 主要特征：
 - 主要负责第三方数据的收集、加工和流转，不一定直接从事广告交易
 - 多于广告交易平台、SSP等合作进行数据变现

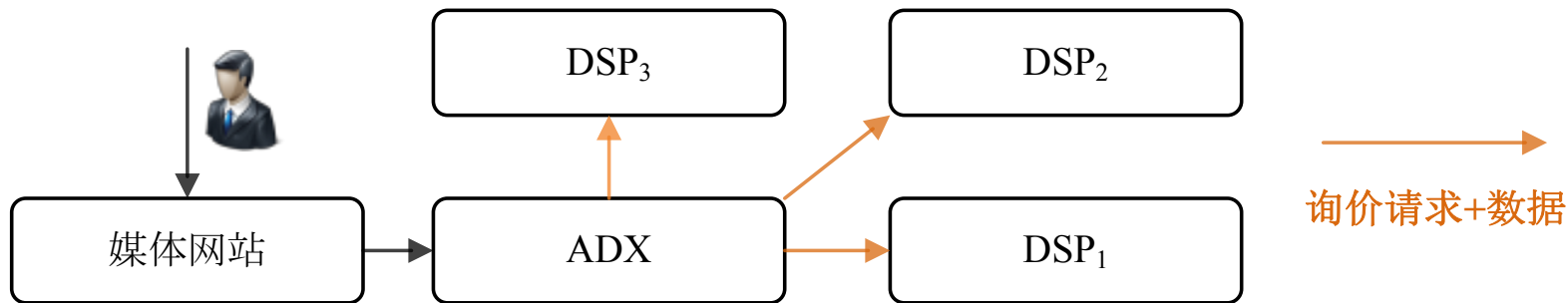
行为数据交易三定律

- 第一定律
 - 行为数据只能交易，不能交换或共享
- 第二定律
 - 只有按效果而非购买量付费，才有足够的需求
- 第三定律
 - 同一数据被越多人使用，价值就变得越低

关于第一定律：为什么数据不能共享？

- 疑问：数据交换似乎在发生啊？
 - 那往往是因为有更高层次的交换，即投资关系。
- 为什么大公司不把数据共享出来？
 - 你见过大公司把钱共享出来么？
 - 短时的贴补性共享是可行的
- 政府数据是可以共享的，这本质上是转移支付

关于第二定律：数据交易该怎么做？



- 数据传输附着在实时竞价过程中，无额外开销
- 需求方可以自由地选择需要的部分人群数据，并且按照实际的广告展示付费

关于第三定律：如何给数据定价？

- 市场化的定价方式是唯一的选择
- 目前数据的价值是被低估的
 - 上页的交易方式并未限制数据供给次数
 - 这间接地抬高了流量价格，而低估了数据价格
- 能否采用竞价的交易方式？
 - 不限量供应的商品，是无法竞价的
 - 数据的限量供应怎么做？

大数据隐私问题严重么？

- 隐私安全基本原则
 - A29：欧盟负责隐私保护条例制定的委员会
 - A29原则
 - Personal Identifiable Information (PII) 不能使用
 - 用户可以要求系统停止记录和使用自己的行为数据
 - 不能长期保存和使用用户的行为数据
- Quasi-identifier与K-anonymity
 - Quasi-identifier: 朝阳区，35岁，在360上班
 - K-anonymity: 北京市，30-40岁，互联网行业

Ad Choices

OVERSTOCK CLEARANCE


18 Piece Set Android 4GB Tablet PC

Price **\$59** Save **76%**

Dawgs Women's 9" Cozy Boots

Price **\$19** Save **76%**



Why Did I Get This Ad?

During your online experience for consumers, some of the ads you may receive from AOL Advertising and its networks are tailored based on your haviors/visits on this device.

Who Provided This Ad?

This ad has been brought to you by [AOL Advertising](#)

Where Can I Learn More About AOL Advertising?

The practices of AOL Advertising, including advertising on our network of websites, as well as across other sites, products, and services that are offered by AOL Advertising are detailed in our [AOL Privacy Policy](#).

How to Opt-Out



Clicking the Opt-Out button above will opt you out of behavioral advertising delivered by the AOL Advertising Network.

Please note that you will continue to receive ads; however, we will no longer tailor ads to your preferences and usage.

大数据隐私问题比想象的更严重

- 稀疏行为数据的新挑战
 - 从一个人观影或购物记录，能否反推他是谁？
 - 实际案例：Netflix推荐大赛，有人从数据集里发现了自己的同事是同性恋
 - 理论研究：Robust De-anonymization of Large Sparse Datasets
- 深度个性化系统也有隐私安全风险！
 - 相关研究课题是差分隐私(Differential Privacy)
- 隐私是大数据头上的达摩克里斯之剑

需求方数据安全问题

