



中國人民大學
RENMIN UNIVERSITY OF CHINA

计算传播理论与实务

2019-2020秋季学期

第六讲

数据伦理

授课教师：范举副教授、塔娜讲师
时间：2019年12月23日

讨论

- 经过简单调研，你们认为什么是需要关注的数据伦理问题？

数据伦理问题

- Quarks and cells neither mind when we discover their secrets nor protest if we alter their environments during the discovery process.
- 计算传播学能否把算法当做“黑箱”？





数据伦理

- ✓ 信息茧房
- ✓ 算法歧视

第6.1节

信息茧房：两种观点

“伪命题”说

- 媒介技术是社会发展的基本动力，每一种新媒介的产生，都开创了人类感知和认识世界的新方式。——麦克卢汉
- 基于算法推荐的内容分发机制，不仅可以满足不同受众的个性化信息需求，还能大大节省用户获取内容的时间成本，使其能够最为效率地获得需要的信息

web 1.0时代

- 由信息私人订制所造成的信息偏食引起众多学者和评论家的担忧：个性化信息的盛行将导致民众只关心他们感兴趣的内容，从而侵蚀社会的共识基础，危害民主社会发展
- “信息茧房”（Information Cocoons）这一概念就在此时兴起，用以描述信息偏食所造成的风险
- 经过长时间的使用，这个概念似乎已经成为一种定理，当算法时代来临，它也自然而然地被挪用在算法推荐的内容分发机制之上，并逐渐成为后者的“原罪”

“信息茧房” 的前世今生

- 1995年，比尔·盖茨曾预言：在未来世界，量身定制的信息将自然增加，每个人都可以编排一份完全符合自己兴趣的“日报”
- 次年出版的《数字化生存》中，尼葛洛庞帝也预言了个性化定制的“我的日报”（The Daily Me）即将出现
- “信息茧房” 最为人熟知的来源是美国法学教授凯斯·桑斯坦的《信息乌托邦》@2006年，当时信息的个人化、定制化已经成为一大趋势
- 桑斯坦非常认同“我的日报”，并在此基础上进一步提出“信息茧房”的概念。他指出，在互联网时代，伴随网络技术的发达以及信息量的剧增，每个人都能随意选择关注的话题，并可以根据自己的喜好打造一份个人日报，但这种信息选择行为将会导致“信息茧房”的形成

桑斯坦的“信息茧房”

- 是指传播体系个人化所导致的信息封闭的后果。
- 当个体只关注自我选择的或能够愉悦自身的內容，而減少对其他信息的接触，久而久之，便会像蚕一样逐渐桎梏于自我编织的“茧房”之中。
- 桑斯坦认为，这将导致视野偏狭和思想的封闭甚至极化，进而会加强偏见并制造出非理性的极端主义，直至侵害政治民主。

算法推荐时代

- 在传统媒体时代，媒体机构面向大众生产新闻，并依赖人工编辑的方式进行分发，并无明确的受众细分意识。而基于算法推荐的内容分发，以用户的个人特征为标准进行信息筛选，并向用户推荐与其兴趣和价值观高度匹配的个人化信息，由此形成“千人千面”的内容消费形态。
- 从积极的方面来看，这种分发方式使信息传播更加扁平化，各类信息内容的显现度与曝光量由算法规则所决定，受众可以自由选择信息，削弱了专业媒体的内容把关能力和议程设置能力
- 算法推荐机制通过个性化的名义背离多元性的信息传播方式：它是否会过分迎合用户偏好，让我们只能看到自己想看到的东西，只听到自己认同的观点，进而导致信息封闭？
- 在这种情况下，“信息茧房”概念又被重新提起，用以批评算法分发所导致信息接收窄化、信息结构失衡问题。无论是用户还是内容供应商，都开始认同这一概念。更多人将其与网络场域的意见分裂、极端化以及社会整体价值观离散化联系在一起，以证明“信息茧房”对社会共识基础的侵害。

怎样论证“信息茧房”的存在？

- 1.从信息消费维度，证明人们只通过有限的媒介渠道（这些媒介还必须是算法分发型）接触信息
- 2.从受众心理维度，论证有限的信息接触，导致了受众观念的单一化和极端化
- **2016年Seth Flaxman等学者进行了一项很有代表性的实验**
 - 研究人员选择了5万名参与者，要求他们报告自己最近阅读、观看或收听的新闻媒体，同时通过电子手段直接监测和记录他们的实际新闻消费行为，包括网页浏览历史等。
 - 通过两项数据的对比，研究最终发现人们实际的媒体消费比他们自己想象中更具有多样性，也就是说，人们实际并没有陷入“信息茧房”中，但他们可能会装作自己陷入了其中。

如何戳破“信息茧房”这个概念

- 1.信息偏食是人的本性，在传统媒体时代就开始有了
- 2.网络并非无菌室，而是芜杂交融的信息环境
 - 互联网的低准入门槛降低了信息获取的成本，不同类型的内容自由流通，人们能接触到比过往更广泛的信息。从这个角度出发，互联网只会打破信息封闭，降低“信息茧房”出现的概率。
 - 互联网也并不仅仅是算法支配的
- 3.信息偏食的后果真的有那么严重吗？

为何相信“信息茧房”的存在？

- 算法不够成熟?
 - 成熟的算法能够为人们提供各种类型的有益信息，在提高生产、生活效率的同时，又能促进人们视野开阔、思想进步，这才是技术向善的体现。
 - 面对算法分发可能存在的信息偏向问题，信息供应商也开始采取各种手段进行纠正，比如新闻APP“跨越分歧的阅读”（Read Across the Aisle）、Chrome浏览器“逃离泡泡”插件，以及微信的“看一看”功能也在以社交推荐的方式对信息偏向进行技术纠偏。
- 桑斯坦也在《信息乌托邦》中指出：“新的传播技术正在使事情变得更好而不是更糟”

“真命题”说

- “信息爆炸” → “信息孤岛”
- 主体局限性：活动领域的边界=>茧房
- 离出发点越远，各个群体间的分化就越深刻，意识形态间的异质程度就越强。而基于正常心理的正向错觉，会让人们高估自身的正确性和重要性，在相互争抢话语权的过程中产生公地悲剧，进而导致社会撕裂。
- 信息茧讨论的不是外部信息如何控制、如何封闭，而是主体本身用碎片式的信息编织了一个结。不同的群体打不同的结，这样异质的社会或许不容易和谐。

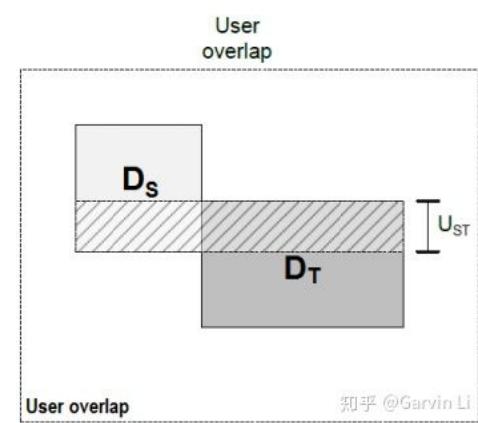


程序猿说？

- 跨域推荐cross domain
- 用户A，经常浏览同一个类型的新闻，比如体育新闻，如何找到A喜欢的其他类别新闻呢？

- **Attribute level** (*Comedy* \leftrightarrow *Thriller*)
 - same type of items, different values of certain attribute
- **Type level** (*Movies* \leftrightarrow *Books*)
 - similar types, sharing some attributes
- **Item level** (*Movies* \leftrightarrow *Restaurants*)
 - distinct types, differing in most, if not all attributes

知乎 @Garvin Li



作为终端用户，如何应对？

- 主动出击，用行为投票
 - 两个短视频APP的较量
- 增加信息获取渠道
 - 兼听则明
 - 一些严肃的信息来源
- 生活在别处
 - 控制屏幕时间

第6.2节

算法歧视

大数据杀熟？





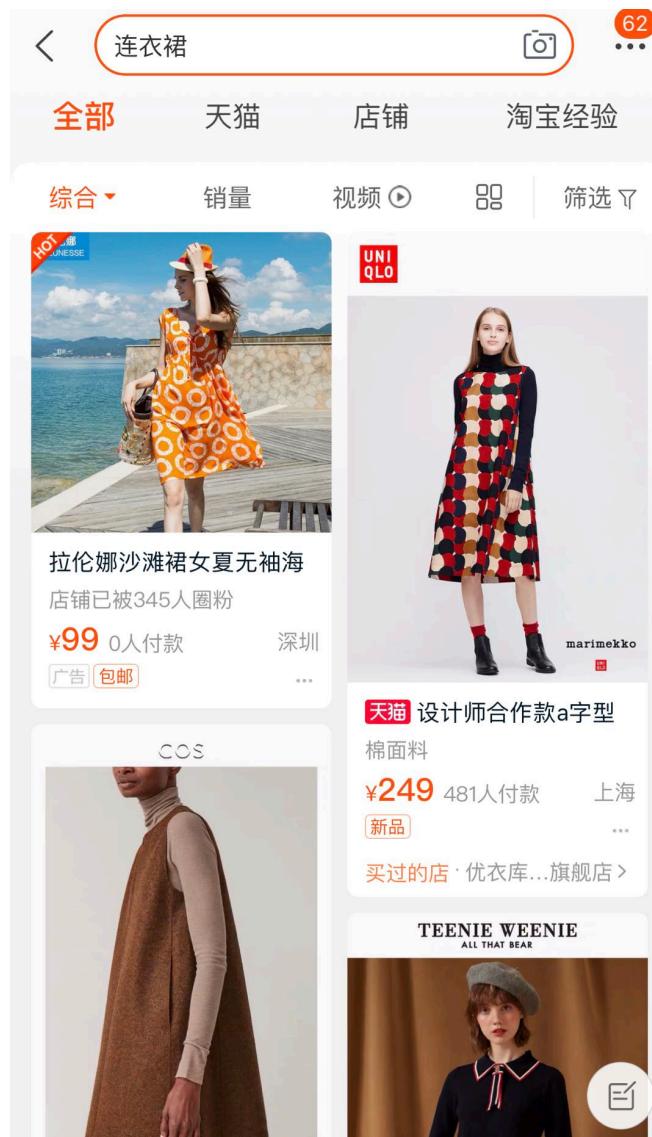
对比

- 一听快乐水
 - 超市: ¥ 2
 - 五星酒店: ¥ 20
- 同一航班机票
 - 价格敏感客户: ¥ 400
 - 价格不敏感客户: ¥ 1000
- 有什么不同?

算法歧视：一些案例

- 谷歌
 - 错误地将黑人程序员上传的自拍照打上“大猩猩”的标签
 - 谷歌搜索中，输入黑人姓名更易出现暗示犯罪历史的广告
 - 谷歌广告服务中，男性会比女性有更多的机会看到高薪招聘消息
- Flickr，错将黑人照片标记成“猿猴”
- 微软，AI聊天机器人Tay，在和网民聊天时被“教坏”，被灌输了许多脏话甚至是种族歧视思想，上线不到一天就紧急下线了
- 亚马逊
 - 购物推荐系统，偏袒自家及合作方的商品
 - 在比价服务中，隐瞒自家商品与合作方的运费，以致于消费者得到的比价结果有失公允

淘宝搜索“连衣裙”？



淘宝搜索“连衣裙”？



风中的厂长 V 95

2月24日 16:50

很多朋友不理解连衣裙客单价低于128的含义，其实答案很扎心。

这个消费区间的人群，被淘宝系统打上了低价人群的标签，经过分析各种行为后，做成了大数据。

低价群体中有较高比例的人会投入更多时间在讨价还价，售后问题，退换返现上，这无形中增加了商家的运营成本。

简单来说就是穷人的时间不值钱，可以为了几块钱和你耗一天。

这样说我心里也很难受，但这是一个社会现实。

最后我想说，莫欺少年穷，我们努力提升自己的客单价吧💪 收起全文 ^

谣言！

算法价格歧视

- 经营者运用算法手段，通过跟踪、分析消费者的消费记录或其他上网记录，绘制出消费者的偏好、习惯等画像。这样使得经营者可以进行个性化的尝试、验证，挖掘出结果数据，获得消费者的最大支付意愿信息后，针对不同的消费者确定不同的价格。
- 更严重的问题是什么？
 - 不公平的技术和信息获取：新的“知沟”

算法歧视的几层含义

- 算法的可获得性 (affordability)
 - 抢票软件、支付软件、排队软件.....
- 算法自身的预设逻辑
 - 目前的广泛共识：黑箱，不好！
 - 警惕：推荐头部流量 (e.g.,头部流量占比30%)
- 算法使用的数据
 - 公说公有理婆说婆有理（微信的报告vs微博的报告）

应对算法歧视

- 基础设施建设（软、硬件）
 - 例：浙江某市的图书馆提供的老年人上网讲座
- 技术层面
 - 改进算法
 - 瓦解算法中存在的性别歧视。训练算法忽略单词间的某些特定的关联，同时保留它们需要的关键信息，如“前台” vs “女性”
 - 可解释性
- 法律法规

法律法规

- 2019年10月9日，文化和旅游部公示《在线旅游经营服务管理暂行规定（征求意见稿）》
- 针对最受关注的“大数据杀熟”问题，《暂行规定》明确规定，在线旅游经营者不得利用大数据等技术手段，针对不同消费特征的旅游者，对同一产品或服务在相同条件下设置差异化的价格
- 大数据杀熟还将面临最高五十万元的处罚。《暂行规定》第三十五条规定，违反该条规定的，由县级以上文化和旅游行政部门依照《中华人民共和国电子商务法》第七十七条的规定处罚。具体来说，由市场监督管理部门责令限期改正，没收违法所得，可以并处五万元以上二十万元以下的罚款；情节严重的，并处二十万元以上五十万元以下的罚款

推荐系统的算法歧视

- 推荐算法的偏差来源
 - 观测偏差：推荐某个项目，此项目得到反馈，进入后续推荐 vs 从未曝光的项目，完全没有反馈
 - 不平衡的数据：数据中存在的系统偏差。例如：有关老人的数据很少，男女搜索习惯数据，
- 如何改进
 - 定义公平：什么是公平的算法？仍然是个开放问题
 - 基于“公平”，改进算法

内容总结

- 信息茧房
- 算法歧视



Thanks!