“融”观中国

中国数据产量占全球10.5%，巨量信息催生“算法创新”

# 网络“大数据”，不只“数据大”

## ——透视“算法”系列报道之一

本报记者 **卢泽华**

在各类网络平台上，“算法”成为一项广泛应用的技术。通过数据分析和精准推荐，它改变了过去“人找信息”的搜索模式，逐步实现“信息找人”。

关于“算法”，人们观点不一。有人认为，它方便了用户，提供了更精准的网络服务；有人认为，它在“算计”用户，造成“信息偏食”等问题；还有人认为，作为一项技术，它并无好坏之分，关键在于如何用正确价值和规范去驾驭。

自今日起，本版推出“透视‘算法’”系列报道，解读“算法”带来的变化与影响。系列报道第一篇，关注大数据时代网民获取信息习惯的变迁。

▲ 在重庆市人力社保政务服务中心，工作人员在介绍人力资源大数据信息库。 新华社记者 **唐 奕**摄

▶ 贵州省贵阳市2023中国国际大数据产业博览会展馆现场。 **何晓璇**摄（**人民视觉**）

### 现状

#### 内容更丰富，平台更多元

“以前遇到问题，第一反应是打开浏览器搜一下。现在除了搜索网页，也会去微信或短视频平台找答案。”北京市民王瑛向记者展示了一段短视频：“你看，前两天下厨，想做一道鱼香肉丝，抖音一找，教学视频就出来了。”

王瑛今年36岁，从2000年左右开始接触互联网。最初10多年，她获取信息的方式主要有两种，一是在搜索引擎上检索，二是浏览门户网站。但最近几年，这种习惯开始改变。“现在网上信息太多，让人眼花缭乱，甄别起来比较费力。有时需要特定信息时，在手机软件中更容易找到结果。”她边介绍，边打开手机界面：“比如，想找购物信息，就去淘宝逛逛；想看电影，就去猫眼找找；想了解明星动态，就去微博看看；想获取专业知识，就去知乎搜搜……”

在王瑛看来，这种改变是不知不觉形成的。每天频繁使用各类手机软件，“顺手搜一下”并不费力。久而久之，就形成了用多个平台搜索信息的习惯。最重要的是，“搜”完之后，手机应用会记住自己的需求，主动推荐感兴趣的信息，省时省力。

王瑛的用网经历具有一定代表性。记者针对120名用户的搜索行为作了一项小型调研，结果显示：所有人均有使用搜索引擎的习惯，但将其作为唯一选择的仅占15%。多数人表示，自己会在各类手机应用中检索信息，并时常浏览各平台推荐的信息。比如，获取生活经验类信息，很多人都会选择刷短视频，“遇事不决，视频来学”。

行业数据也印证了这一点。截至2023年6月，中国搜索引擎用户规模已达8.41亿人。但近七成用户也会通过短视频完成搜索，微信“搜一搜”月活跃用户也已超过8亿，几乎与搜索引擎用户总量相当。另一项调查也显示，用户平均每天会在3.8个平台上搜索信息。

### 挑战

#### 信息杂芜与“平台壁垒”

虽然网上信息量更大了，获取信息的渠道更多了，但不少用户仍表示“存在使用体验不佳的情况”。

信息杂芜是原因之一。“现在网上营销信息较多，虽然大多标注了‘广告’字样，但毕竟需要甄别，影响使用效率。”王辉是一名金融分析师，经常需要在网上检索信息，且对信息质量要求较高。但他发现，现在找到一张“纯净的网页”非常困难，这在一定程度上拖慢了他的工作节奏。

王辉的情况并非个例。记者调查发现，在网上检索信息时，显示结果中很多都是广告链接。比如，搜索“培训”二字，排名最靠前的两个结果均为商业培训机构广告。

据了解，各大搜索网站也都注意到这个问题，但如何改变行之有年的商业模式，还是个尚待破解的命题。“广告是网站主要盈利来源，目前业界只能尽量优化广告信息，提升用户体验，但要彻底转型，还需一定时间探索。”一家互联网公司的销售经理告诉记者。

除了信息杂芜，互联网界还不同程度存在“平台壁垒”现象。

据了解，过去，综合类搜索网站是各类网络平台获取流量的主要入口，通过一个搜索框、一个关键词，用户可以搜到全网各个平台的信息。而随着移动互联时代来临，美食、家居、旅游、短视频等垂直平台用户规模越来越大，每个平台都形成了自己独有的“大数据库”，这个“大数据库”不仅包含海量优质信息，还囊括了用户使用习惯等数据。动辄坐拥上亿用户的网络平台不再需要其他平台“引流”，这些数据便被视为独家资源，不再对外开放。这导致用户很难通过综合类平台搜到所有需要的信息。

### 趋势

#### 向智能化与精准化发展

如今，中国已经成为数据量最大、数据类型最丰富的国家之一。国家数据局日前发布的信息显示：我国数据产量已占全球数据总量的10.5%。更有研究机构预测，到2025年，中国数据总量或占全球近30%。

“数据作为新型生产要素，已快速融入生产、分配、流通、消费和社会服务管理等各环节，但目前在发展过程中也存在数据应用潜力释放不够、数据供给质量不高、流通机制不畅等问题。”国家数据局副局长沈竹林表示。

在这一背景下，仅仅“数据大”已经无法满足社会需要，还要进一步实现“数据精”“数据准”，从而有效降低用户获取信息成本，提供更加个性化、精细化的内容服务。在这种需求驱动下，“算法”技术应运而生。

“算法”是人工智能技术与大数据技术的关联节点。中国科学院科技战略咨询研究院院长潘教峰认为，人工智能正在科研、贸易、交通、医疗等诸多场景中改变人们沟通的方式。其“活字典”、移动的“大百科全书”的特性，加上越来越接近人类水平的语言理解和生产能力，让其“答复”既有常识性，又呈现出集成创新性。

因此，在互联网海量数据基础上，“算法”可以通过人工智能技术对用户兴趣偏好和需求进行分析归纳，描绘出用户画像，打上特性“标签”，进而以此为依据，进行精准化、个性化的定制推荐。

在“算法”技术的调配下，用户每次打开手机应用，都可以收到针对个人需求的信息推送。同时，平台通过不间断收集信息，会变得越来越“懂”用户——打开购物网站，近期想要购买的产品已被推荐到首页；进入短视频平台，接踵而至的都是自己喜欢的视频内容；浏览旅游软件，想去的地方一一列在眼前，连机票购买链接都放好了……这一切，正是凭借日渐成熟的算法技术才得以实现。

然而，有人担心，“算法”在大数据时代精准调配网络信息、提升服务水准的同时，也带来了“信息偏食”、网络沉迷、数据泄露、技术滥用等隐患。这些问题多大程度存在？又有哪些破解之道？我们将在“透视‘算法’”系列报道第二篇进行解读，请大家持续关注。

新媒视点

# 用正确价值导向驾驭『算法』

周培源

自上世纪90年代以来，搜索引擎是当之无愧的互联网第一入口，其广泛普及重构了人与信息的关系，实现了从被动接收信息到主动寻找信息的转变。

随着移动互联网的兴起，用户获取信息的习惯也发生了一定变化。从过去网上搜索的单一渠道，向各类社交平台与垂直社区延伸。这反映了用户行为和偏好的改变，也对信息传播模式产生较大影响。

如果说，传统网站更关注内容与内容的连接，那么移动互联网则更关注人与人的关联，后者增强了用户获取信息行为的个性化、体验感和信任度——他们更青睐个性化、智能化的搜索体验。

如今，这种体验因为“算法”的出现，又获得了质的提升。各类手机软件普遍使用“算法”为用户画像，通过对用户行为的长期跟踪与分析，提供更具个性化的信息服务，“不仅懂问题，更懂问问题的人”。

其实，过去近30年来，互联网界一直都在优化“算法”，以提高信息传播的效率和价值。但“算法”普遍进入公众视野，还是在人工智能技术大规模应用后形成的。近年来，生成式人工智能的兴起，迅速掀起全球范围的大模型技术发展热潮。

生成式人工智能，是指基于算法、模型、规则生成文本、图片、声音、视频、代码等内容的技术。具有上千万甚至百万亿参数的深度学习或机器学习模型——大模型，是生成式人工智能的“灵魂”。过去一年，中国相继推出了一系列大模型工具，中国科技部新一代人工智能发展研究中心发布的《中国人工智能大模型地图研究报告》显示，中国研发的大模型数量排名全球第二，10亿参数规模以上的大模型已发布79个。

在这一背景下，生成式人工智能为人们提供了全新的信息获取模式——它以同聊天机器人“对话”的形式，直接提供答案，并在不断“学习”中提升答案的质量与效率。

值得注意的是，“算法”技术是把双刃剑。它日益渗入日常生活、赋能各行各业的同时，也存在一些隐忧。比如，不当使用可能带来欺诈、虚假信息传播等不良影响；大模型依靠从互联网抓取的海量数据训练，有可能损害数据隐私、侵犯版权等；它还可能导致信息泡沫、信息“偏食”等问题。越来越多的研究显示，“算法”提供给用户的信息更多趋向于表层阅读而非深度思考，从长远来看，这不利于用户在获取信息过程中提高处理问题的能力。

如今，有关部门正在加大力度整治“算法”带来的价值导向失范和不良内容多发等乱象，督促各类平台强化内容审核制度，全面规范“算法”技术的使用。各大网络平台也在着力解决算法价值导向存在偏差、优质内容呈现不足等问题，不断优化流量分配机制，防止“重指标轻质量”、重经济效益轻社会效益。我们相信，只要切实做到用正确价值驾驭“算法”，就一定能更好满足人民群众精神文化需求，推动构建更加清朗的网络空间，促进互联网行业健康有序发展。

（**作者系南开大学融媒体研究中心研究员**）

前沿动态▶▶▶

## 广东“数字人大”上线

**本报广州电**（记者**姜晓丹**）1月21日，在广东省第十四届人大二次会议前夕，广东“数字人大”上线发布会举行，“粤当家”小程序正式亮相。打开“粤当家”小程序，全省21个地市人大代表联络站地图一目了然，群众选择距离自己最近的代表联络站，即可联系代表、反映问题。

据悉，“数字人大”通过打造民主民意平台，建设面向社会公众的“粤当家”小程序，面向代表的“粤政易”移动端，面向代表联络站、立法联系点的基层站点工作端，面向各级人大机关的民主民意管理端，形成民主民意数据收集、流转、处理、利用的全流程线上闭环。

“‘数字人大’建设以‘好用、实用、管用’为导向，实现四级人大全贯通、五级代表全覆盖、智慧履职全方位，标志着广东省人大工作正式迈入了数字时代。”广东省人大常委会主任黄楚平表示。

## 内蒙古举行工业互联网开发者大会

**本报鄂尔多斯电**（记者**张枨**）以“筑人工智能平台，聚数字产业生态”为主题的内蒙古鄂尔多斯工业互联网开发者大会日前在该市东胜区举行，发布内蒙古首个以行业AI（人工智能）大模型和工业应用商城为核心的工业互联网平台——鄂尔多斯工业互联网平台。本次大会由鄂尔多斯市人民政府、国家能源投资集团有限责任公司主办。

据悉，该平台由鄂尔多斯市创新投资集团与华为公司共同打造，包含基础算力、AI算力、AI平台、工业大模型、应用开发平台、应用商城等多个模块，可向应用开发者和工业企业提供开发、运营、销售、咨询等一站式服务。该平台中的AI模块包含了先进的工业大模型，工业企业无需自建数据中心，可以充分利用相关AI和应用为其生产服务，支持低门槛、高效率的AI工业化开发，解决当前AI开发的门槛高、周期长、精度低、泛化性差等问题。“到2025年，鄂尔多斯重点产业工业互联网平台覆盖率将达100%，带动数字经济核心产业增长值达到200亿元。”鄂尔多斯市委副书记、市长杜汇良表示。

岛妹聊事儿

# 这事能推广吗？电影票开场还能退，

袁子涵

眼瞅着，这春节就要来了，春节档的影片大战也提上日程。最近，岛妹发现了一个变化：有的电影院，电影票开场也能退了！

从今年元旦开始，湖南长沙的部分影院摆上了自助退票机。电影开场30分钟内，观众觉得影院环境不好，或者挑了个烂片，这票钱花得不值了，就能在机器上自助退票。服务一出，讨论度很高。有人感叹，终于告别给烂片充钱的日子了；但也有人担心，这会不会有漏洞？

花上两位数甚至三位数，空出几个小时的时间，坐在电影院里，谁都不愿意碰上烂片。但很多时候，看电影就像开盲盒，开场之后才能知道好坏。这些年我们有了评分网站，买票之前，会习惯性地上网看眼评分、刷下评价再买票。可亲眼见的，毕竟还是和网上听来的有差距，辛苦攻略再踩了大雷的情况也不少。这时，影院出了“后悔药”，让消费者用选择投票，确实能倒逼行业提升质量。

而从实际来看，这个规则也有问题。对电影好坏的判断，实际上是个很主观的事。放给观众自主选择，其实就是“无理由退票”。这会不会招来别有用心的人仗着规则恶意退票？还有观点觉得，电影作为产品有特殊性。一般商品，无理由退换不影响二次销售。可电影票一旦开场才退，座位就浪费了。

现在，“电影票开场还能退”这个服务，因为争议较大已经暂时下线。但它推出后，通过网上的反应就能看出：观众对电影产业，显然还有更高的期待。说到底，一个健康繁荣的市场，离不开消费者持续、积极的反馈。

请扫描二维码观看视频