

# Research on News Production Strategy Under Algorithm

Sheng Wang<sup>1,\*</sup>

<sup>1</sup>Nanchong Vocational College of Culture and Tourism, Langzhong City, Sichuan Province, China  
\*1015665545@qq.com

## ABSTRACT

With the development of economy and smart mobile communication equipment, the market has higher and higher requirements for timeliness and efficiency of news production. Manual news writing is gradually unable to meet the requirements of news production due to low efficiency, high cost and poor timeliness. The machine production technology of news gradually appeared. Primary machine writing completes sentence creation by defining grammar, understanding vocabulary semantics, and word order. However, as the length of the article increases, the accuracy is poor and the sentence ambiguity problem cannot be solved. Natural language processing and other technologies, through the construction of deep learning models, extracting text features, and using massive text content for training, have significantly improved the accuracy of the language, as evidenced by the emergence of machine writing software such as Xiaomingbot and Dreamwriter. However, current language processing models still have problems such as inability to understand abstract expressions, inability to generate creative words and sentences, and rigid language expressions. The integration of media workers and artificial intelligence can truly improve the accuracy, efficiency and artistry of language understanding, analysis, expression, and generation, thereby increasing the efficiency of news production and reducing the cost of news production.

**Keywords:** news production, machine writing, man-machine integration

## 算法下的新闻生产策略探究

王胜<sup>1,\*</sup>

<sup>1</sup>南充文化旅游职业学院文化服务系 阆中市, 四川省, 中国  
\*1015665545@qq.com

## 摘要

随着经济与智能移动通讯设备的发展,市场对新闻生产的时效和效率要求越来越高,人工新闻写作因效率低、成本高、时效差等原因逐渐无法满足新闻生产的要求,新闻的机器生产技术逐渐出现。初级机器写作通过定义语法、理解词汇语义以及语序排列所完成语句创造,但随着文章篇幅增加准确性较差,且解决不了语句歧义问题。自然语言处理等技术,通过构建深度学习模型,提取文本特征,利用海量文本内容进行训练,从而使得语言的准确性得到了显著提升,Xiaomingbot和Dreamwriter等机器写作软件的出现就是很好的佐证。然而,当前的语言处理模型仍然存在无法理解抽象表达、无法生成创造性词句、语言表达刻板僵硬等问题。媒体工作者和人工智能的融合能够真正提升语言理解、分析、表达、生成的准确性、效率和艺术性,进而提升新闻生产的效率、同时降低新闻生产的成本。

**关键词:** 新闻生产, 机器写作, 人机融合

## 1. 前言

随着人工智能技术的发展,人工智能对各个领域都产生了不同程度的影响,在传媒领域,人工智能凭借其技术优势打破了媒体行业传统制作模式,对新闻行业从选题、写作、推送等方面都带了巨大的影响。探索人工智能背景下新闻生产的发展趋势与应对策略,成为业界与学术界共同关注的焦点。

## 2. 技术变革对当下媒体内容生产方式的影响

### 2.1 大数据赋能选题更加精确

随着新闻制作和发布的门槛越来越低,传统媒体不仅要面临来自上下游的挤压,也要应对自媒体的竞争。于是阅读量成为了最重要的一项指标,而选题是新闻生产工作中的第一步,选题的好坏直接影响新闻的成败。传统媒体选题一般有如下方法。方法一,深入分析用户,挖掘用户的痛点和喜好,基于用户的痛点和内容喜好制定选题;方法二,划分栏目,形成以栏目为主导下的选题大方向,比如时政要闻、社会观察、财经快报等,在指定的栏目下形成选题,缩小选题范围,降低选题难度;方法三、根据特定的节日形成选题;方法四、根据社会热点形成选题;方法五、紧跟同行。

与传统媒体不同,人工智能技术在内容选题策划上可以智能分析社会热门事件、热点词汇传播的频度和热度等,并快速反馈给编辑【1】,大数据将为选题策划带来如下的改变。第一,依托大数据发掘热点与潜在热点话题。第二,基于大数据抓取关键词,以关键词为主导形成选题大方向,第三,基于节日、社会热点,抓取以往或当下数据,形成该类选题更加精确的内容关键词。在大数据的赋能之下,传媒行业的选题更加精确,提前了解受众的内容兴趣点,提前完成内容布局,让内容生产与受众阅读无缝对接。

### 2.2 AI 技术加持内容制作更加高效

新闻讲求时效性,在传统媒体的发展中,处理、分析信息数据,选择、发布新闻内容等等环节,都需要耗费大量时间【2】,环节越多,耗费的时间就越多,新闻的时效性就越差;其次,更多时间的消耗导致公司收益降低。这种传生产方万式既无法满足用户对时效的追求,也无法满足资本逐利的目的,可以兼顾效率和成本的写作机器人对于新闻行业而言是充满诱惑的。2015年9月10日,由腾讯财经研发的人工智能新闻写作机器人撰写的第一篇稿件后,引发了全国媒体的广泛关注与热烈讨论,腾讯的 Dreamwriter 和头条的 Ai 机器人张小明、新华社的“快笔小新”、《南方都市报》的小南等一批写作机器人,纷纷进入大众的视线。2017年8月8日,四川发生7.0级地震,但

不是记者编辑,而是写稿机器人最先发布了相关的地震消息。对事件信息的撰写,仅仅只用了25秒就在全球进行了首发。稿件的质量更是无可挑剔,逻辑性强,信息内容清晰、明确,对事件的描写更是确保了真实性。仅仅25秒就完成了写作、审阅和发布的一整套流程,时间短、质量高、发布快,这是人工写作需要面对的巨大挑战。

新闻写作机器人能够自动完成原始素材的收集处理,在极短的时间内完成写作与发布,新闻写作机器人的到来极大地提升了传媒行业的生产效率,而随着人工智能技术的进步,未来机器人写作的效率将会更加迅速,准确率也会大幅提升,单位时间内的新闻产量将超越人类的想象。

### 2.3 基于大数据的精准推送

报纸、杂志、广播,电视等传统媒体在内容推送方面首先是缺少反馈,传统媒介在信息传播中都是线性的、单向的,是媒介指向受众的传播,受众只能被动接受信息;其次,传统媒介是以自我为中心的推送,不对自己的受众做精细区分,不考虑受众的喜好,全部推送给受众。

智媒体的推送方式与传统媒体有本质的区别,在人工智能中,大数据是“燃料”,算法是“灵魂”【3】,算法决定着智媒体的效率和质量【4】,人工智能在算法下要实现的就是内容与用户之间的匹配,通过算法设计,机器通过对用户的行为与信息碎片进行抓取和收集,通过分析用户的行为得出用户个人肖像数据,结合 AI 算法形成精准的用户画像与客户识别,然后将用户更可能感兴趣的内容推送给用户,以此提高用户对对该媒体的依赖或商业上的转化。这是一种基于受众精细区分、知晓用户自特点、考虑受众喜好的新闻推送方式,而且是一种双向的传播,传播方将会根据用户的调整随时跟进调整,推送的内容不仅精准,而且符合受众兴趣喜好。

## 3. 智能化生产方式的带来的问题与其本缺陷

写作机器人虽然有诸多优势,但写作机器人依然面临如下挑战,一,深度内容很难自动生成。二,计算机推理能力以及自然语言理解能力等方面仍面临极高的技术难关。三,写作机器人无法做到个性化生产。

### 3.1 生产内容僵硬、刻板、重复

在逻辑推演方面, AI 受算法和程序支配,而人类则靠意识完成,这是二者信息生成机制的根本差别【5】。因为机器不具有自我意思,其根本原理是在大量数据“喂养”下形成的算法,而这种基础的算法目前存在重复、刻板、僵硬等缺点。但算法不是人工智能的短板,数据积累不足和数据的标准化问题是人工智

能最大的短板。在设计写作机器人时，第一步是大量收集写作机器人的工作成果数据，并对这些数据做ETL（数据抽取、清洗、转换、加载），然后定义数据的标准与数据的含义，随后继续定义要改进的目标，然后调整模型参数训练算法，最后让算法能直接影响到写作机器人，算法运行以后还要持续收集数据，迭代模型，修正错误，最终形成一个正循环，但完成这样一个流程，不仅技术壁垒极高，时间成本也极高。人工智能通常有“强”“弱”之分，弱人工智能是目前主攻的研究方向，强人工智能尚处于初始或停滞状态【6】。

面对技术壁垒与研发成本问题，很多企业选择了另外一种方法，基于固定产品，使变量垂直，依托语料库、素材库、语法表达库等形成新闻模板，形成半人工干预的写作机器人。以现阶段中文语义下的自然语言生成技术来说，像财经类、体育赛事类的新闻稿件，通过预定义模板的方式来实现智能写作，GPT2是目前最先进的文本生成模型，GPT2作为语言模型的一种，其基本原理依然是背诵与记忆。如果想让机器生产的文章像媒体从业者一样，且具有“原创性+通顺性”，那目前的智能写作机器人肯定实现不了。所以目前的人工智能写作，都是基于语言模型下的，背诵与记忆下的程式化的制作。

### 3.2 写作风格千篇一律

程式化写作的外在表现便是千篇一律的写作风格。“Get 智能写作”软件的界面拥有热点、素材、Ai 初稿、模板、收藏、文集、百宝箱几个大板块，软件可汇集全网的热点榜单，让写作人员不用为选题烦恼，点击初稿写作，就可以实现素材快速匹配，同时还能推荐不同的段落提纲，写作人员还可以导入其他的文章，进行修改，写完之后还可以智能检测，降低敏感词与重复率。这种写作仅仅只是机器对海量数据的深度挖掘和采集后的碎片化素材整合，其次这种写作模式依然是人机协同模式。

当前针对财经、体育等新闻报道的人工智能写作，因为涉及的数据多、研判数据对机器人而言相对简单，但不同类型的新闻报道需要软件开发者在前期，针对不同类型的新闻收集数据，并对新闻数据进行变量控制，同时准备该产品所需的语料库、素材库和语法表达库等，依托高度成熟的范文形成模板。使用者只需要输入关键的文字信息，软件便可以制作出一个简单且质量合格的文章。但由于所有的参考资料都需要人为输入，所有的行为也是要基于数据，因此人工智能写作无法跳出现有信息的桎梏，而人工智能真正无法逃出的是人设定的模板和程序。正如清华大学彭兰教授所言：如今的机器人撰稿基本上套用的使一种“人工模板”+“自动化数据填充”的模式【7】。

## 4. 智媒体背景下的内容生产策略

虽然人工智能技术在不断的进步，在数据收集、

写作模仿上也逐渐成熟，人工智能取代简单的数据分析类型与客观描述型报道已成为不争的事实，但人工智能写作完全取代人的写作尚有一段不小的距离。看到科技的力量，警惕科技的威胁，深入分析，提前布局，加快转型，有利于传媒人提高自己的生产力。

### 4.1 快速转型，技术赋能优化内容生产

大数据、人工智能等技术在各行各业的应用愈发的普遍，传统媒体应该敞开怀抱，加速转型，实现传统媒体与新型技术的融合，促进传统媒体与新兴技术的优势互补，以此发挥媒体更大的潜力和优势。

新技术的出现，受众阅读习惯的改变，传统媒体转型太慢，纸媒减版、停刊已经十分普遍，然而“报纸印刷总量的下滑趋势尽管收窄，但仍有 8.64% 的降幅”【8】。加速转型、实现破局，将人工智能技术与人自主意识相结合，让技术与内容深度融合，进而改善整个新闻生产过程。因为没有内容的加持，技术终将只是毫无意义的炫技，坚持党的方针，倡导正确的价值观，报道人民关心的问题，生产群众喜闻乐见的内容，在技术赋能的条件下，生产更多价值积极、质量上乘的内容来。

### 4.2 深耕细分领域，做小众定制内容

随着人工智能技术的成熟与传统媒体、自媒体人对人工智能技术的高度接纳，业界必然会迎来内容量产的大爆发，但这种大爆发并非原创的大爆发，而是基于人工智能的粘贴复制、伪装洗文、批量生产，内容同质化现象将会愈演愈烈，整个媒体产业的内容生产将会面临重复、泛滥、枯燥等问题，看似令人眼花缭乱的内容市场，却让大众陷入同质化的审美疲劳。人工智能对模板的套用，也会逐渐在部分领域中实现内卷，标准化的文体与优质的内容写作，将会让一批新闻写作者无所适从。

移动互联网时代，更多元的文化造就了多元的用户追求，创作者必须深挖富有个性的小众用户，向更精准地用户提供更具有价值的内容，满足小众群体的内容与审美需求，提高用户粘度。目前这类媒体流量虽小，但足够精准、足够有特点，以此为拓展根基继续深耕，今后便会有更多的该类用户涌入。

深耕细分领域，做小众定制内容不仅在内容的生产上与人工智能展开差异化的竞争，也在竞争中让生产者对给领域做到了深入垂直，深入的了解了用户的审美、喜好、特点与消费习惯，这样一来也更容易实现商业上的转化，从而保证媒体生产的良性运转。

### 4.3 做好深度报道

人工智能实现的量产是基于大数据下的组合与拼贴，新素材的来源依然需要人来完成。在信息爆炸的当下，有深度、有温度、有态度的文章依然备受读者追捧。深度报道以其极强的真实性、紧凑的故事叙

述、跌宕起伏的情节发展且具有巨大的社会影响而备受读者喜爱。而对事件进行跟踪、深入调查、对素材进行收集和整理与阐释等工作，只能依靠人来完成。

2019年6月中旬，一则名为一孕妇在泰国被丈夫推下悬崖的消息震惊了国人，事件发生之后各大媒体都对该事件进行了报道。2020年9月6日我在南京见到了当事人王女士，经过十多个小时的相处，收集音频资料5个多小时，拿到所有素材以后，我开始剪辑与文字编辑，最终以37分钟的声音纪录片+图片+文字的形式，在澎湃新闻首发，虽然1年前已有国内外多家媒体对该事件进行了报道，绝大多数的受众已经了解了该事件的来龙去脉，这则新闻已经失去了新鲜度，其意义和价值已经不大，但这篇名为《声音纪录片|绝地重生，泰国坠崖孕妇案当事人口述》的文章在推送之后，在发出短短几天内，便有了40多万的阅读量。通过这件事情可以说明，即便在机器人写稿时代，人类记者对事实真相勇往直前的深挖与报道，依然是未来新闻报道的根基【9】。更加真实、更加全面、更多细节的深度报道，依然是媒体的核心竞争力。

#### 4.4 深耕评论性质的文章

人工智能的本质是建立一个或多个模型，而这种模型则是一种复杂函数，当输入一个自变量后，函数可以生成一个符合人脑思维结果的因变量。当前智能写作通常需要人类事先准备好训练语料或写作逻辑，算法容易做到让写作结果符合语法，但很难做到让写作结果符合人类常识或特定知识。序列生成算法依然是最重要和最有潜力的模型思路，但常用的序列生成算法需要大量平行语料才能完成训练任务，这与智能写作面对的发散性的需求并不相符。其次机器基于充分的数据训练，训练模型才拥有了创作能力，然后根据人类的指令，产出符合特定格式要求的创作结果，然而模型产出的结果是无法事先预期的，所以人工智能并不能完全代替创作者一次性就能成型一篇，观点鲜明、论据充分、情感充沛且打动人心且的好文章。因为当前的人工智无法构建完成这样一种复杂的模型，而构建这杨一种复杂的模型并用数据验证模型的合理性，且让这类模型适用于各类新闻选题的写作还有很长的路要走，而这其中最难的一个部分就是自我观点的形成，人工智能作为人造物，是否会具有自我意识，本文不做讨论，假设人工智能不具备自我意识，那人工智能的写作将还是基于人工干预下的人机协同写作。然而社评、社论等对个体意识和个体表达要求极高的工作，人工智能更是无法完成。而在这个部分，个体价值的体现则愈发的明显。作为内容的生产者，客观公正的评价一个事件、从不同的角度和立场发表对事件的看法，极其考验新闻工作者的知识储备、行文逻辑与理论水平。专业的传媒从业者在长期积累的新闻工作实践中形成的情感感知能力、价值判断能力和写作技巧能力，是传媒产业得以存续和兴旺的生命力【10】，在人工智能写作机器到来的今天，提高

社论、社评等专业技能，将会让新闻工作者立于不败之地。

## 5. 结论

人工智能算法并不能完全代替人的创作，但人工智能写作通过对传媒行业的增添助力，可以让传媒人从简单枯燥且重复的基础写作中解脱出来。放眼未来，只有站在技术的肩膀上，把握时代的新机遇，加快传统媒体再转型，实现传统媒体与人工智能写作的快速融合，有效提升新闻产品的采制速度与质量，才能实现传统媒体的新生。再强大的模型也无法创造出从未见过的句子，人工智能写作依然需要人的干预，而媒体和人工智能写作将成为命运共同体，共存共生。

## REFERENCES

- [1] Xu Fang. Reconstruction of the publishing industry in the era of artificial intelligence and big data [J]. Journal of Beijing Institute of Graphic Communication. 2018, 26(6): 40~44.
- [2] Lin Xiaolin. Analysis of the integration strategy of traditional media and new media in the all-media era [J]. Media Forum, 2019, 2(17): 79-80.
- [3] Guo Quanzhong. Three New Trends in the Development of Smart Media [J]. News Front. 2019, (23)
- [4] Guo Quanzhong. Smart Media: Media Convergence greatly [J]. Southern Media Research. 2019, (03)
- [5] Fu Xiaoguang, Wu Yutong. On the Logical Features of AI News Writing——Based on the Comparative Analysis of Dreamwriter Reporting and Manual Reporting [J]. Modern Publishing. 2021, (01)
- [6] Luo Zhenglin. The three changes of artificial intelligence and news communication ecology [J]. Journalism Lover. 2020, (06)
- [7] Peng Lan. The media has penetrated into everything—a new round of technology-driven pan-medialization trend [J]. Editorial Friend, 2016(03): 5-10.
- [8] In 2018, the total print volume of newspapers nationwide was 78.1 billion pairs, and the opening decline narrowed [EB/OL]. People's Daily Online, [2019-04-25]. <http://media.people.com.cn/n1/2019/0425/c426303-31050291.html>.
- [9] Zhang Linhe. An Analysis of Reporters' Coping Strategies in the New Style of Robot Writing [J]. Journalism Lover. 2019, (03)
- [10] Hu Jialin. The value and transformation of traditional media in the era of smart media [J].

Youth Journalist, 2017(29): 25-26

- [11] State of the News Media 2015 — Journalism & Media, Pew Research Center, 2015.
- [12] Thurman, N. , Drr, K. , & Kunert, J.( 2017).When Reporters Get Hands—on with Robo—Writing : Professionals consider automated journalism’s capabilities and consequences. Digital Journalism, 2017.
- [13] Latar N L, Nordfors D.The Future of Journalism: Artificial Intelligence And Digital Identities[D]. Stanford Center for Innovation and Communication Stanford University, 2011.
- [14] Dong L, Smith R G, Buchanan B G.NewsFinder: Automating an Artificial Intelligence News Service[C].IAAI, 2011.
- [15] Bucher, T. (2016).‘Machines don’t have instincts’: Articulating the computational in journalism. New Media and Society.