

• 互联网+护理 •

公众对上门护理服务的态度和关注点:基于网络爬虫的文本挖掘

彭佳玲¹,周茂林²,杨青³

摘要:**目的** 基于网络爬虫获取上门服务平台评论数据,采用文本挖掘技术探讨用户对上门服务的关注焦点和情感态度。**方法** 采用 Python 代码爬取上门护理服务平台评论数据,用狄利克雷分布主题模型挖掘获取评论潜在主题,利用 ROSTCM6.0 软件对评论文本数据进行高频词统计、主题情感分析等过程,挖掘护士和患者所关注的热点、核心主题,以及评论情感倾向。**结果** 共抓取到 3 166 条评论,主题模型识别出 6 个潜在主题,分别为平台功能、服务价格和效率、服务内容和专业度、护士态度及患者体验、服务对象、平台的下载和使用;评论中用户对“护士”“服务”“态度”“方便”“平台”等词提及度较高,在一定程度上展现了用户感受和关注重点;用户积极情绪占 59.19%,中性情绪占 27.23%,消极情绪占 13.58%,用户消极情绪主要集中在服务水平存在差异、收费不合理、平台下载使用等方面。**结论** 用户对各平台整体服务感知较好,但目前发展中仍存在服务水平参差不齐,收费不规范、平台不正规等困境,建议针对发展过程中的短板精准地实施改进,推动上门护理服务良性发展。
关键词: 互联网+护理; 护士上门服务; 共享护士; 网约平台; 态度; 文本挖掘; 主题
中图分类号: R47;C931 **DOI:**10.3870/j.issn.1001-4152.2023.05.110

Public attitudes and concerns about home care services: web crawler-based text mining Peng Jialing, Zhou Maolin, Yang Qing.
School of Nursing, Chengdu Medical College, Chengdu 610500, China
Abstract: **Objective** To obtain the comment data of the on-site service platform using the web crawler, and to explore the users' focus and emotional attitude on the on-site service using text mining technology. **Methods** Python code was used to crawl the comment data of door-to-door nursing service platform, latent dirichlet allocation theme model mining was utilized to obtain the potential theme of comment, and ROSTCM6.0 software was used to calculate high-frequency term and analyze emotion theme, and mine the hotspots, core themes, and comment emotion tendencies of nurses and patients. **Results** A total of 3 166 comments were captured, and the topic model identified 6 potential topics, namely platform function, service price and efficiency, service content and professionalism, nurse attitude and patient experience, service objects, and platform download and use. Words such as nurse, service, attitude, convenience and platform were highly mentioned by users, which showed the users' feelings and focus to some extent; users' positive emotions accounted for 59.19%, neutral emotions took up 27.23%, and negative emotions accounted for 13.58%, and users' negative emotions mainly focused on differences in service levels, unreasonable charges, and platform download and use, etc. **Conclusion** Users have a good perception of the overall service of each platform, but there are still dilemmas such as uneven service levels, irregular charges, and irregular platforms. It is recommended to improve accurately for the shortcomings in the development process, so as to promote the healthy development of door-to-door nursing services.
Key words: internet plus nursing; door-to-door nursing services; nurse sharing; online appointment platform; attitude; text mining; theme

根据世界卫生组织预测,2050 年中国超过 60 岁的人口占全国人口比例将达 35%,至此将成为全球老龄化最严重的国家^[1],且随着慢性病发病率的上升,失能、失智人员的护理问题已经成为亟待解决的社会性难题。2015 年,国务院发布《关于积极推进“互联网+”行动的指导意见》^[2],鼓励推进发展基于互联网的医疗、护理、健康、养老等服务。自 2017 年底开始,全国多个城市的网约护士服务手机 App 平台陆续上线,共享护士,或称网约护士服务横空出世,切实反映了当代社会在中国老龄化和少子化人口背

作者单位:1. 成都医学院护理学院(四川 成都,610500);2. 电子科技大学护理学院;3. 四川省肿瘤医院护理部
彭佳玲:女,硕士在读,学生
通信作者:杨青,yangqingse@163.com
科研项目:四川省医学会伤口疾病(泰阁)专项科研课题(2021TG05),四川省干部保健科研课题(川干研 2020-802)
收稿:2022-10-07;修回:2022-12-12

景下的养老需求。用户可根据自己的需求下单,护士抢单或是系统派单,护士上门提供专业的护理服务,不仅能够为居家患者提供便利,还能让护理人员充分利用自己的空闲时间,实现执业自由,提高收入。但作为一个高风险、高专业度的领域,上门服务存在一定的风险,各平台在服务内容、合作机构、人员结构等方面存在一定差异,患者和护士在享受其带来的福利的同时,也纷纷表示了对服务质量和安全的担忧。然而目前只有极少数学者关注网络平台服务质量评价^[3-4],且缺少从不同形式渠道的角度进行针对性的研究。文本挖掘技术能将难以量化的大规模文本数据整合转换为结构化数据,并抽取有价值的情报信息,已被广泛应用于商业、医疗和金融等领域^[5-6],但其在护理领域应用较少。因此,本文将基于文本挖掘,结合狄利克雷分布(Latent Dirichlet Allocation, LDA)主题模型、高频词分析和情感分析等研究方法,分析不同平台用户评论,获取用户服务需求维度,对

公众针对上门护理服务的需求和态度进行研究和探讨。帮助掌握护士和患者对平台服务质量的感知状况,为整体优化上门护理服务质量提供决策支持。

1 资料与方法

1.1 评论数据抓取 在 Android、IOS 系统应用商店中以“共享护士”“上门护士”“护士”为关键词,搜索相关 App。随即查看比较应用商店所推荐 App 的开发者、兼容性、服务项目、评分评论等基本信息,本研究通过筛选共纳入 6 个 App,分别是鸿华医疗、医护到家(患者端)、医护到家(护士端)、金牌护士(患者端)、金牌护士(护士端)、护士小鹿。以上 App 开发者均为目前较成熟互联网医疗服务机构,服务范围覆盖全国、使用率评论率较高且数据公开,均能在应用商店免费下载。网络爬虫是一种数据信息采集的常用手段,通过网页的链接地址验证超链接和 HTML 代码,根据定义好的字段筛选数据获得数据资源^[7]。截至 2022 年 4 月,我们使用爬虫代码从以上 6 个 App 抓取了患者和护士的所有评论数据,通过删除重复文本和无效评论,共获取数据 3 166 条。

1.2 数据预处理 数据预处理工作主要包括:数据清洗、中文分词、去除停用词等。数据清洗包括使用正则表达式匹配去除表情、各种符号标签、空白字段等,然后使用 Python 的三方开源工具包 jieba 进行中文分词,将文本数据构成词袋,并进行去除连词、助词、介词等停用词处理。

1.3 数据分析

1.3.1 主题分析 采用基于隐含 LDA 的主题模型^[8],从文本中挖掘出隐含的主题信息。通过 Python 中的机器学习模块 sklearn 库进行 LDA 分析,按照词频排序保留前 1 000 位的特征词项生成特征词向量空间矩阵。为寻找最优的主题数(K),使用困惑度确定主题数并采用一致性评分来做出判断。困惑度是一种对语言概率模型的效果进行评价并协助改进参数的有效方法^[9],通过计算不同主题数量对应的困惑度值,并利用 Python 的第三方工具包 Matplotlib 来绘图,以得到困惑度随主题数量变化的折线图,当困惑度最低对应的 K 值为最优。本研究中,经过分析,当 K 值为 6 时,模型的困惑度处于极小值,且主题一致性达到最优,因此,本研究主题数确定为 6,随后在 6 组主题词中出现频率在前 20 位的关键词里选择具有代表性的主题词,并根据主题词为该主题命名。

1.3.2 基于 ROSTCM6.0 软件的高频词分析和情感分析 为了将评论文本数据以直观的形式展现,以利于更好地发现数据中隐藏的特征、关系和模式。本研究引入 ROST Content Mining System6.0 (ROSTCM6.0)^[10]文本挖掘软件,将修正后的 text 文本导入 ROSTCM6.0 软件中,对关键词进行筛选处理,获得有意义的词汇及中心词,然后将评论中所产生的高频

词汇运用在线词云生成软件以词云图的形式直观展现出来。情感分析是判断在线评论文本情感类型的常用方法^[11]。本研究运用 ROSTCM6.0 中的“情感分析”功能得出情感分布结果,并从分布结果中进一步分离其中的积极情绪、中性情绪和消极情绪。最终进一步明晰用户对各平台服务感知的现状,并基于此分析当前各平台服务中的弊端并给出改进建议。

2 结果

2.1 主题分析结果 LDA 模型将 3 166 条评论数据分为 6 个主题,各主题及主题词见表 1。

表 1 LDA 主题模型划分的上门护理服务 App 评价文本的 6 个主题及对应主题词

主题	主题词
平台功能	陪诊、医护、平台、问题、解决、预约、服务、挂号、客服、定制
服务价格和效率	医院、排队、时间、医护、预约、比较、价格、软件、放心、提前
服务内容和专业度	护士、上门、打针、输液、专业、服务、看病、预约、不疼、技术
护士态度和患者体验	满意、专业、态度、服务专业、热情、好评、体验、很好、准时、支持、到位
服务对象	医院、护士、老人、服务、软件、医疗机构、朋友、家人
平台的下载和使用	软件、平台、专业、App、使用、下载、好用、检查、下单

2.2 基于 ROSTCM6.0 软件的高频词分析 如图 1 所示,由词云图形式对词频排列在前 100 的词语进行展示,可以直观看到在上门服务平台的评论数据中提及到的高频词汇。词云中字体大小代表词频大小,字体越大,表示词频越高,词频的高低能直接反映用户感知的高低,在一定程度上展现了用户针对上门护理服务关注的重点。其中“护士”888 次,“服务”828 次,“上门”680 次,反映了护士是上门护理服务的主力军,护士上门开展护理服务已经为大众知晓并接纳;“方便”515 次,说明方便是上门护理服务带给用户最直接真实的感受;“态度”445 次,“医院”434 次,“平台”175 次,“技术”114 次等则反映出了用户在使用服务时关注的重点内容;在项目上,“预约”288 次,“打针”283 次,“输液”193 次,“挂号”162 次,说明这些服务项目使用频率较高,受到广大患者的欢迎;在服务评价方面,以称赞居多,如“热情”278 次,“准时”186 次,“耐心”111 次,“满意”112 次等,展示了患者对上门护理服务整体形象感知较好,以积极正面情绪为主。词云图能很容易看出护士与患者的评论所涉及的方面,从而能从他们的角度总结出关于上门护理服务开展的一些启示和建议,便于在以后提高支持度和关注度,从而实现患者、护士、平台三方和谐共赢。

2.3 基于 ROSTCM6.0 软件的情感分析结果 针对 3 166 条关于上门护理服务平台评论的所有评论情感得分分布见表 2。总体来看,用户对各平台服务整体认可度较高,如表 2 所示积极情绪占比最高,59.19%的用户对上门护理服务持积极态度,主要表现在以下方面:①为广大患者带来便利。评论主要为

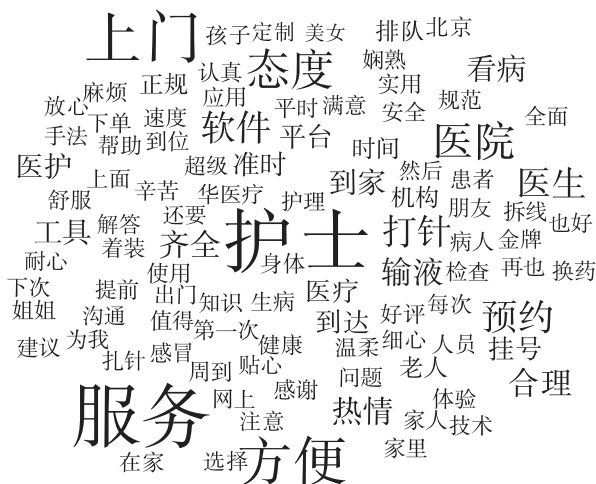


图 1 高频词词云图

“足不出户享受到了医疗服务”“感谢护士为我们解决了大问题”“再也不用来回跑医院”等,即上门护理服务的开展实实在在解决了一部分患者看病难的问题。②护士态度和患者体验。评论主要为“服务很到位”“护士态度和蔼,服务非常周到”“沟通非常愉快,强烈推荐”“服务很人性化、专业”即护士的服务态度和专业服务为患者带来了良好的体验感。③平台功能强大。评论主要为:“个性化的定制陪诊,太棒了”“客服在线实时咨询,功能强大”“能轻松学到医学科普知识”“告别拥挤挂号难”等,说明各大平台功能齐全强大,满足了患者的多层次、多方位医疗需求。④给护

表2 文本情感分析($n=3166$)

情绪分类	分析结果		一般		中度		高度	
	数量(条)	百分比(%)	数量(条)	百分比(%)	数量(条)	百分比(%)	数量(条)	百分比(%)
积极情绪	1874	59.19	1106	34.93	513	16.20	255	8.05
中性情绪	862	27.23						
消极情绪	430	13.58	296	9.35	125	3.95	9	0.28

注：中性情绪不做分段。

3 讨论

在互联网时代,越来越多的用户愿意通过在线评论的方式分享自身使用体验,网络信息来源广泛且数据量庞大,与传统问卷调查数据相比,网络文本能够更加真实、广泛地反映公众对某一事物的感知^[12]且已经在社会科学中广泛应用^[13],提示在线评论数据可以对传统研究数据来源进行良好的补充,值得深入探究。但其在护理研究领域尚缺乏关注。

3.1 平台应用概况

现有研究缺乏对服务平台质量的关注,目前仅发现一篇文献对共享护士 App 进行了应用评价^[14],其余未见与本研究相似内容的报道。本研究共纳入 6 个 App,这些 App 的功能强大,不仅包含基础护理项目,还有 PICC 护理、造口护理等专科护理项目,有的还有提供针对特色人群的服务,如母婴护理、临终关怀、安宁照护等,以及各类辅

士带来额外收入。评论有:“这个平台解决了护士收入低,人员不稳定的问题,很推荐”“作为一名护士,能在业余时间赚一些零花钱非常开心”“很好的副业平台,能运用自己的专业做个副业”等,表明上门护理服务的开展为护士带来就业机会,实现了技能变现,具有潜力巨大的发展前景。

消极情绪和中性情绪也占有一定比重,需要进一步改进。主要表现在:①服务水平存在差异。评论主要为患者提到的“操作不严谨,事后不负责”“服务不专业,纱布连伤口都没盖全,不推荐”“护士不会拆线弄得很难”“可能是自己倒霉,碰到了没有专业素养,态度又恶劣的陪诊护士”等。即表明上门护理服务存在服务水平参差不齐的现状。②收费不合理。如患者提到的“补差价被说的名正言顺”“拆线七针 200,另外的加收 10 块一个,取消订单还扣 50”“价格客服没有提前说,私自扣款”等。护士评论“无法提现,手续费还那么高”“平台分成太低,不公平”等,即上门护理服务收费和分成存在缺乏公开性和透明性、护士平台利益分配不明确的问题。③App 下载使用。如护士提到的“感觉就是骗取医护人员证件信息”“认证和证件都通过了一直还提示要认证”“信息审核速度堪忧”等,以及患者提到的“下载软件后频繁收到骚扰短信”“打开速度慢”等评论,表明现有 App 发展存在打开速度慢、审核时间过长以及短信骚扰、信息泄露等问题。因此上门护理服务可以以此为突破口进行完善与改进。

助功能,如挂号、陪诊、在线咨询等。服务范围覆盖全国,可通过定位进行城市切换,表明上门护理服务的出现的确能够解决医疗资源紧缺带来的看病难问题。然而在分析评论数据过程中,发现平台之间存在服务水平参差不齐、收费标准不统一,护士资格审查不严、信息泄露带来的安全隐患问题以及 App 自身的下载、运行不畅等问题。相关部门应在相关制度完善、风险防控、价格和支付政策等方面进行持续改善。

3.2 护士针对上门护理服务的感知状况 就护士而言,上门护理服务的出现帮护士实现了技能变现,增加了额外收入。但随之而来的执业安全风险、信息泄露、收费分配不合理等问题,导致不少护士对上门服务浅尝辄止。有研究表明,脱离医院工作环境,会导致护士因自身安全考虑而减少参与度^[15]。实名认证、医疗信息共享也带来了信息泄露的安全隐患,此

种担心不仅来源于护士和患者,提供服务的医疗机构与第三方平台亦有此方面的担心和顾虑^[16]。以及有护士提到自己由于在使用部分 App 过程中存在短信骚扰、无法通过审核、平台利益分成不公等问题,选择放弃使用该平台。因此若无法解决上述问题,则可能限制“互联网+护理服务”的长远发展。

护士是上门护理服务的主力军,保障上门服务护士的人身安全是顺利开展“互联网+护理服务”的关键。有研究指出,作为服务提供者的护理人员的合法权益特别是安全问题应得到更好的制度保障^[17],另外已有不少学者提出给护士配备一键报警设备、强化患者信息真实性监管、纳入男护士陪同上门等方式^[18],值得学习借鉴。信息共享是保障护理安全的重要决定因素,但也必须采取相应措施规避其带来的信息泄露问题。有研究指出,采用信息加密技术,从技术和法律层面为医疗机构、上门护士和患者三方构建监管体系实施全面管理,能够有效平衡信息共享与隐私保护,维护个人信息安全^[19]。此外,我国“互联网+护理服务”平台数量虽已达 20 余家,但只有极少数平台依托具有《医疗机构执业许可证》的实体医疗机构^[20],仍有实际并未运营,界面无法下单甚至为骗取手续费的虚假 App 存在,不少平台还存在抽取提成过高和乱收手续费的问题,平台开发者应在满足用户需求的前提下,开发出更正规专业的 App,相关部门应加强对平台的监管,规范上门护士注册流程,按照统一标准收取相关费用。此外应用商店应在今后加强对 App 上架的审核,帮助上门护理服务走向良性的可持续发展道路。

3.3 患者针对上门护理服务的感知状况 上门护理服务的出现为广大“看病难”患者带来了福音,让患者足不出户就能享受优质的医疗服务。有研究表明,服务态度是影响患者服务体验感以及评价的相关因素之一^[21]。本研究结果显示,患者对上门护士的服务态度总体比较满意,公众对上门护理服务总体表现出积极情感,这与邱欣等^[22]调查结果一致。证明上门护理服务确实给患者带来了良好的服务体验,受到了广大患者的欢迎,具备良好的发展前景。但由于我国上门护理服务起步晚,各平台入驻护士资质良莠不齐、平台管控不严等原因,导致服务水平参差不齐,此外由于院内外操作不具备同质化,导致护理质量存在差异,使患者对上门护士的工作能力存疑^[19]。而随着上门护理服务范围的拓宽和服务对象需求多样化,对护士也提出了更高要求,只有不断提升护士能力,才能确保服务质量。目前上门护理服务缺乏统一收费标准,同一护理服务在不同 App 间价格区间较大^[23],且在收费上大量使用“×元起”“×元以上”等词语,缺乏公开性和透明性。相关调查发现,82.52% 的患者认为上门服务的价格过高,且目前支付形式以

自费为主^[24],这在一定程度上制约了上门服务的开展。以及本研究中有患者反映有平台存在私自扣款、加收费用的问题,容易增加患者对上门护理服务的不满情绪。

尽管国家卫健委已对上门护士的资质作出了明确规定^[25],但根据研究显示,仅有 33% 的 App 介绍了上门服务护士的职业资质信息,仍有部分平台存在对护士资质范围控制不严的情况,导致目前上门护士群体出现资质混乱的局面^[14]。因此,明确上门护士的资质,完善准入机制,通过统一的培训和考核来保障护理质量是不可或缺的,平台更应该在护士资格审查过程中严格把控。此外有学者指出,搭建统一培训中心、构建统一培训方案,组建专家库来承担培训工作并对护士进行选拔,组建医护团队、增加多学科合作等措施来提高上门护士执业能力,为上门服务筑建安全防线^[19],均值得借鉴。自“互联网+护理”服务开展以来,服务项目的收费规范,一直是公众关心的重要问题之一。未来,建议政府加强政策支持,进一步统一规划和正确引导,统筹“互联网+护理服务”平台中上门护理服务项目的分类和收费标准,并合理纳入社会医疗保险覆盖范围,并加强对 App 的监管。除此之外,在服务前有必要与患者签订知情同意书,注明收费标准及相关事项,在患者知情并同意签字情况下执行相关操作及费用收取,避免不必要的纠纷^[26]。

4 小结

网约护士的出现为民众带来了福利,也为护士提供了更多实现价值的机会,但上门护理服务发展尚在起步阶段,各平台发展缺乏系统的评价分析,而网络信息时代为上门护理服务研究提供了更多的数据来源。本研究对网约护士平台评论数据进行了文本挖掘分析,解析了用户真实需求与平台发展中所遇困境,有利于帮助平台掌握护士和患者对平台服务质量的感知状况,为各大上门护理平台优化自身服务质量提供决策支持。本研究的局限:首先,由于部分平台信息不对外公开或评价信息较少,样本来源只有 6 个平台,结果可能不能完全代表国内所有网约平台和用户,此外评论数据是不断更新的,具有一定时效性,在后续的研究中,将继续扩大样本量进行更深入研究。

参考文献:

[1] 张登利. 中国居家养老服务的支持性政策研究[D]. 武汉:华中科技大学,2019.
[2] 中华人民共和国国务院. 国务院关于积极推进“互联网+”行动的指导意见[N]. 经济日报,2015-07-05(005).
[3] 曹峻玮,张贵花,路瑶,等. 我国移动医疗 APP 服务质量与用户满意度关系研究[J]. 中国卫生质量管理,2022,29(1):51-56.
[4] 陈婷. 基于文本分析的移动医疗 APP 用户服务满意度研究[D]. 济南:山东财经大学,2019.