

# 传统价值观成语当代语境表现分析 ——基于BCC语料库的计量研究

孙浩, 刘洋洋, 杜惠东, 刘鹏远\*, 于东, 康晨

北京语言大学 信息科学学院, 北京100083

202221198129@stu.blcu.edu.cn, 202421198094@stu.blcu.edu.cn, yzqtd@gmail.com,  
liupengyuan@pku.edu.cn, yudong.blcu@126.com, kangchen@blcu.edu.cn

## 摘要

中华优秀传统文化是提升我国新时代文化软实力的重要源泉, 将传统价值观和成语相结合, 有助于继承和弘扬我们的优秀文明。本文提出了传统价值观成语当代语境表现的研究框架, 基于BCC语料库对传统价值观成语语料数量分布和成语传统价值观偏好分布特征、在当代语境中的情感倾向及高频词分布特点、社会话题及道德特征进行计量研究, 并提出了传统价值观成语的当代社会话题及道德适应性指数, 以系统研究传统价值观成语的当代语境表现。本文为传统文化的当代计量研究提供了新的视角, 也为数字人文领域的相关研究提供了参考依据, 旨在增强中华优秀传统文化在当今新时代的影响力, 为中华文明的传承与创新作出贡献。

**关键词:** 传统价值观成语; 当代语境; 语料库计量

## An Analysis of the Contemporary Contextual Expression of Traditional Value Idioms ——A Quantitative Study Based on the BCC Corpus

Hao Sun, Yangyang Liu, Huidong Du, Pengyuan Liu\*, Dong Yu, Chen Kang

College of Information Science, Beijing Language and Culture University, Beijing 100083

202221198129@stu.blcu.edu.cn, 202421198094@stu.blcu.edu.cn, yzqtd@gmail.com,  
liupengyuan@pku.edu.cn, yudong.blcu@126.com, kangchen@blcu.edu.cn

## Abstract

Excellent traditional Chinese culture is a key source for enhancing China's cultural soft power in the new era. Integrating traditional values with idiomatic expressions helps promote and preserve cultural heritage. This paper proposes a research framework to analyze how value-laden idioms function in contemporary discourse. Using the BCC corpus, we conduct quantitative analyses on the distribution of idioms, their associated value preferences, emotional tendencies, high-frequency context words, as well as related social topics and moral dimensions. Furthermore, we introduce a topic-morality adaptability index to evaluate how traditional idioms align with modern social discourse. This study offers a structured approach for the quantitative analysis of traditional culture and provides methodological insights for digital humanities research. It aims to support the revitalization of traditional Chinese values and contribute to the inheritance and innovation of Chinese civilization.

**Keywords:** Traditional Value Idioms, Contemporary Context, Corpus-based Quantitative Analysis

\*为通讯作者

©2025 中国计算语言学大会 根据《Creative Commons Attribution 4.0 International License》许可出版

## 1 引言

“中华文化是我们提高国家文化软实力最深厚的源泉，是我们提高国家文化软实力的重要途径。要使中华民族最基本的文化基因与当代文化相适应、与现代社会相协调，把跨越时空、超越国度、富有永恒魅力、具有当代价值的文化精神弘扬起来，把继承优秀传统文化又弘扬时代精神、立足本国又面向世界的当代中国文化创新成果传播出去。<sup>1</sup>”习近平主席如是说。我们有最优秀的传统价值观和成语文化，将成语和传统价值观相结合(Sun H et al., 2024)有利于发掘我们的传统文化基因。而将传统价值观成语立足于当代语境，则更有利于我们将传统文化与时代精神相结合。当前已出现诸多有关当代社会的价值观与道德等方面的计量研究，无论是对社会主义核心价值观的探求(Liu P et al., 2022)，还是对社会道德的历时梳理(Wang H et al., 2023)，均体现出中华文明的强大生命力，语料库也为我们对当代社会现象及观念的探索(朱述承et al., 2021)提供重要支撑，这些定量研究证明，我们的时代精神是熠熠生辉的。但立足于当今复杂的国际形势，我们更要推陈出新，了解传统文化的当代语境表现有利于我们增强文化软实力，为实现中华民族伟大复兴做出贡献。基于此，本文关注以下问题：其一，传统价值观成语在BCC语料库的不同领域中的分布是否具有结构性差异；其二，传统价值观成语的语境情感表达是否与其文化内涵一致；其三，能否构建适应性指标，量化传统价值观在当代语境中的适应性特征。本文的研究创新主要包括以下三个方面：

(1) 现有社会计算与情感分析领域多聚焦于网络、新闻等现代语言资源，缺乏对汉语特有的传统文化语言单位在当代语境使用情况的计量分析。本文选取2472条传统价值观成语，基于BCC语料库，进行传统价值观成语跨领域分布、语境特征及适应性的量化建模，拓展了情感与语义分析在中文语言单元上的应用范围。

(2) 本文结合大语言模型与人工校对，完成了高频成语语境的情感标注任务，并进行名词、动词和形容词的词性和属性标注，覆盖社会话题分类与道德类型体系。该方法在中文情感及道德计算任务中具备良好的扩展性与跨任务迁移潜力。

(3) 本文引入语境适应性建模方法，结合话题、道德的分布广度与强度，提出“传统价值观成语当代适应性指数”。该指标不仅适用于成语等稳定语言单元的研究，也可以迁移至习语、典故等文化语言单元的建模任务，具有一定的通用性。

本文为传统文化的当代计量研究提供了新的视角，为数字人文研究提供了新的参考维度。

## 2 相关工作

### 2.1 成语价值观相关研究

成语作为汉语言文化的重要组成部分，承载并体现着中华民族的历史文化和传统价值观念。成语中蕴含的价值观是研究的重要方向之一。当前的研究主要包括成语价值观的定性研究与定量研究两个方面。定性研究则包括对成语中所体现的社会主义核心价值观以及儒家价值观的分析。吴云(2017)认为弘扬中华成语文化就是践行社会主义核心价值观，有利于增强我国文化软实力。李玉叶(2015)以《论语》中的成语为例，从以仁待人和以仁律己切入，阐释了“仁”的文化内涵。定量研究出现了中华成语传统价值观的知识库构建。Sun H等人(Sun H et al., 2024)从共时和历时角度以定量的手段研究了成语传统价值观的面貌及含义演变，提供了独特的文化信息。但现有研究多为成语本身的价值观分析，并未涉及到基于语料库的价值观成语的当代分布的计量研究。

### 2.2 “语境”界定及当代语境相关研究

对于“语境”的界定，刘宁生(1984)对中文的“语境”作出详细界定：“某一语言片断中各语言单位的相互制约关系”为“语境”；“与言语行为有关的知识、习惯、信仰等各种因素所构成的交际情境”为“言语环境”。虽然二者不同质，但当我们把静态的语言单位用于交际过程，“言语环境”可以转换成相应的“语境”。同时，Saeed(1997)解释了语境特征三元化：从物理语境可以估计到的；从已经说的话中可以找到的；从背景或共享知识中可以找到的。胡壮麟(2002)梳理语境界定发现，学者们对“语境”的关注从语言语境已走向多元化<sup>2</sup>。由此可见，上述学者对语境的界定已涉及到社会生活诸多领域，如“多元化”。由此我们将本文中的“语境”界定至“社会”方面，即有关社会现象观念特点的研究均属于语境的研究。

<sup>1</sup>来源于《求是》杂志文章《加强文化遗产保护传承弘扬中华优秀传统文化》

<sup>2</sup>主要指“世界知识、集体知识、特定知识、参与者、正式程度及媒体等。”

因此，有关这一方面的当代语境研究国内多集中于：（1）中国传统与当代文化及跨文化语境下的社会观念及现象研究。刘明(2024)基于中华优秀传统文化语境，从哲学、心理学和美学角度探讨了艺术疗愈思想的应用与发展。王婉(2024)重在分析当代视觉文化语境下的传统戏曲艺术传承与创新策略。曾锡(2025)通过对电影《姥姥的外孙》的叙事研究，发现了亚洲文化语境下的传统伦理与个人主义之间的矛盾。上述研究均为定性探讨研究，缺乏文化语境下的社会主题及道德等观念的定量研究。（2）中国社会主题、道德、价值观及性别等方面的计量研究。饶高琦(2016)不仅对1946 年到2015 年的报刊语料和其中的主题历史热词进行提取研究，还做了民国《申报》主题词的历时研究(饶高琦et al., 2023)，这也是探究当代社会话语语境的较好途径和手段。Wang H等人(Wang H et al., 2023) 更是利用人民日报语料库对中国社会道德变化发展进行计量分析。同时，Liu P等人(Liu P et al., 2022)也对社会主义核心价值观行为体系进行定量研究，朱述承等人(朱述承et al., 2021)还利用语料库研究了职业性别偏见社会现象。这些研究虽为我们了解中国现当代社会提供有益视角，但未将相关社会现象有机结合，研究对象较为单一。

综上，目前针对成语价值观的研究缺少对其在语料库中的计量考察，同时更缺少对传统价值观成语的社会语境相关研究。

3 研究设计

3.1 研究框架

语料库是通过相关词汇及语境了解社会面貌的有效载体。本文通过分析传统价值观成语在BCC语料库中的语料，从成语在BCC语料库（以下简称BCC）中的“分布”和“当代语境特征及适应性”两个维度入手，设计两大分析维度和四个计量角度并最终实现四个研究目标，此举能够系统有效地观测传统价值观成语的当代语境表现。

分析维度	计量角度	研究目的
传统价值观成语 BCC分布	跨领域传统价值观	了解传统价值观成语
	成语数量分布	在语料库中的分布信息与面貌
	跨领域成语	得出成语传统价值观在
传统价值观成语 当代语境特征及适应性	传统价值观偏好分布	跨领域下的分布特点与差异
	传统价值观成语	展示传统价值观成语
	语境情感及高频词分布	当代语境情感及文化内涵倾向
	传统价值观成语语境	探究传统价值观成语语境
	社会话题及道德分布	当代社会话题和道德内容分布及适应性

Table 1: 研究框架及研究目标

3.2 研究对象及语料来源

BCC是服务语言本体研究和语言应用研究的在线大数据系统(荀恩东et al., 2016)，在“多领域、文学、报刊和对话”四个领域内广泛收录社会语料，为国际中文教育及相关社会现象的研究提供重要支撑。因此本文以Sun H等人(Sun H et al., 2024)的中华成语传统价值观知识库2472个传统价值观成语为载体，通过爬虫共获取BCC语料库相关成语语料1409276条，并对这些传统价值观成语语料进行全面细致的研究分析（相关论文与资源链接见附录A）。

4 传统价值观成语BCC分布

4.1 跨领域传统价值观成语数量分布

根据图1，报刊领域的语料主要来源于报刊文本，这表明传统价值观成语在当代报道中占据着重要位置；多领域文本则主要来源于微博等社交平台，包括相关媒体报道和网友个人观点，无论是官方媒体的报道还是个人的日常表达，都离不开传统价值观成语的使用；其次是对话领域，而文学领域的传统价值观成语使用相对较少，反映出在文学文本中传统价值观成语的使用频率较低。

由图2，在BCC中成语语料最多的传统价值观为“自强”，依次为“仁义、智、信”。最少的是“俭、孝”传统价值观。可以看出，在语料库中，有关“自强”的传统价值观成语使用频率较

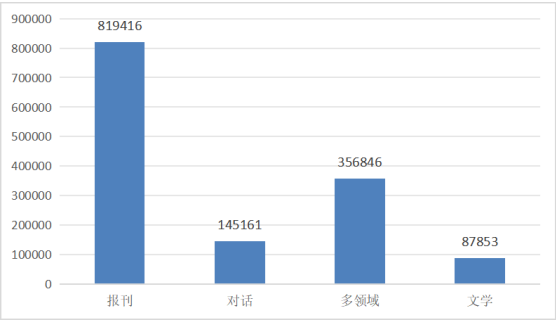


Figure 1: 成语跨领域数量分布图

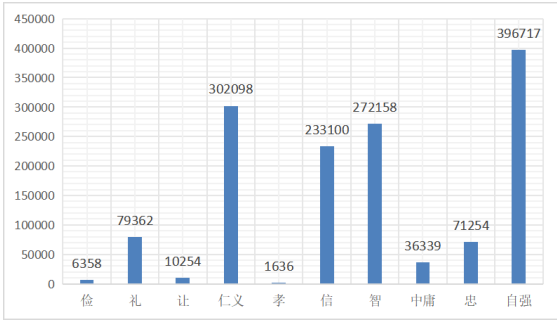


Figure 2: 成语传统价值观数量分布图

高，关于“俭、孝”的传统价值观成语使用较少，可以看到传统价值观如“自强、仁义、智”在当今社会的流行度与社会整体的价值观讨论偏好。

#### 4.2 跨领域成语传统价值观偏好分布

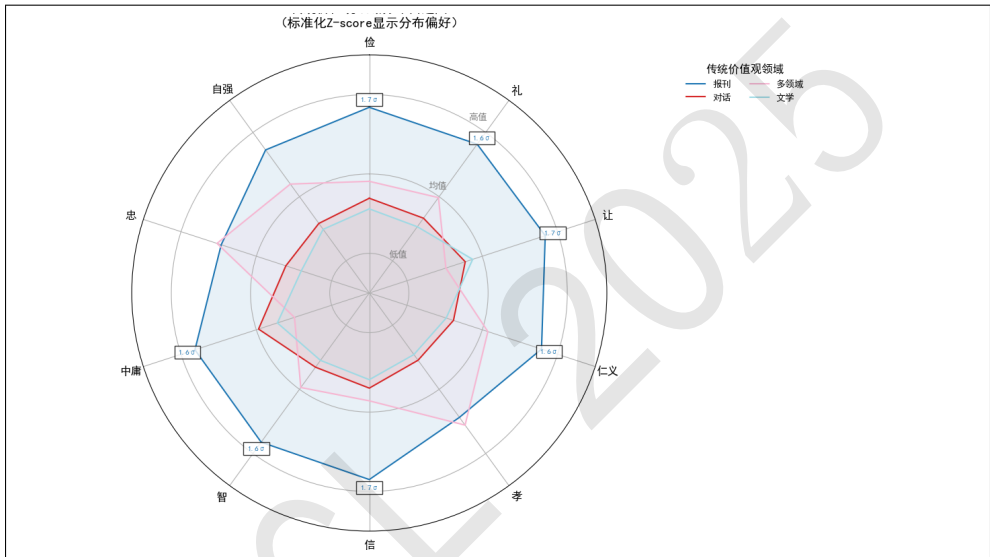


Figure 3: 跨领域成语传统价值观偏好分布图

我们对不同传统价值观成语在语料库中跨领域的语料分布进行Z-Score标准化计量，得出偏好分布（详见附录B）。如图，**报刊领域**在“仁义、信和智”等传统价值观维度上的Z分数较高。这一偏好分布与报刊的社会功能息息相关。根据哈尔马斯的公共领域理论，社会媒体是价值观念传播与建构的核心平台。在强化主流意识形态的过程中，媒体通过选取具有高度认同感与道德规范作用的成语来强化社会规范与集体意识，所以偏好“仁义、信和智”这类道德高度集中的传统价值观。这体现了报刊媒体在现代语境中对‘道德合法性’的持续追求；在日常对话语境中，交际行为受限于言语行为中的礼貌原则，‘信’与‘让’等成语往往用于维护社会秩序与人际和谐，符合中国传统高语境文化的交流特征。‘俭’和‘中庸’价值观反映出节制和均衡在个体表达中的文化内化过程。这样的交际特点使得**对话领域**在“中庸、信、让与俭”这四个传统价值观方面存在一定的偏好，其Z分数相对较高；由于社交媒体的自由公开属性，人们可能更注重人与人之间在爱情、友情、亲情等方方面面的忠诚，这一特点使得**多领域**在“忠、孝和自强”传统价值观维度上呈现出较强的偏好，尤其是在“忠”这一传统价值观维度上表现得尤为突出；文学人物形象中不乏挑战传统伦理观念的角色，以此展现出复杂多样的社会和个人心理。这种艺术特性使得**文学领域**在“仁义和忠”等传统价值观上的数值较低，偏好表现较弱。通过对这些领域在成语传统价值观方面的偏好差异进行统计研究，不同领域具有不同的价值定位：报刊着重关注社会价值，对话则注重理解和节俭，多领域强调自我成长，文学体现出独特的价值内涵。



## 5 传统价值观成语当代语境特征及适应性

### 5.1 传统价值观成语当代语境情感及高频词特征

### 5.1.1 传统价值观成语当代语境情感特征

Sun H等人(Sun H et al., 2024)在中华成语传统价值观知识库中将传统价值观成语按照正向符合传统价值观与负向违反传统价值观进行分类, 同时由于BCC中多领域的语料更贴近当代语境, 因此我们使用大模型(DeepSeek-R1)标注了多领域语料数量分布最多的五条正向—符合传统价值观成语与五条负向—违反传统价值观成语, 每条成语各抽取1000条语料(共1000条)进行情感倾向标注。传统价值观成语标注流程及信息详见附录C。

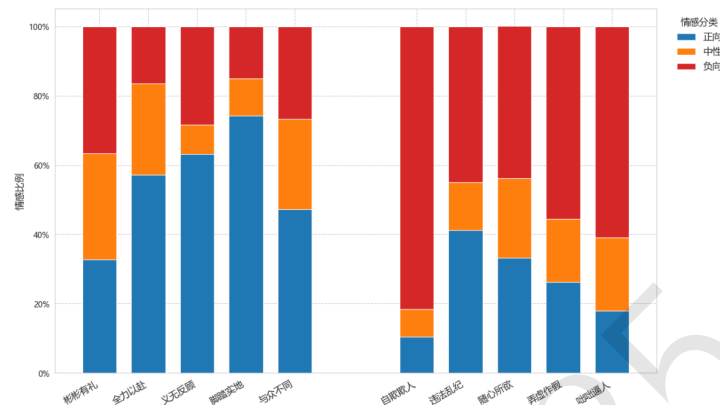


Figure 4: 多领域—高频传统价值观成语语料情感倾向分布统计图

如图4左侧，正向类成语的正向情感倾向远超负向与中性倾向。“全力以赴、义无反顾、脚踏实地、与众不同”这四个成语的语境则多为正向情感，其中前三个成语的正向情感倾向呈递增态势，反映出当代人不断奋进的追求。而传统价值观成语“彬彬有礼”的语境负向情感倾向高于正向，可以看出人们运用该成语的嘲讽心态。图4右侧的负向类成语的语境情感倾向占比与本身感情色彩相符。“自欺欺人、违法乱纪”等成语明确表达人们对不良行为的否定态度。传统价值观成语示例语料详见附录D。

### 5.1.2 传统价值观成语当代语境高频词特征

本文通过对两大类即正向—符合与负向—违反的传统价值观成语在多领域中的语料句子进行分词,由此得到涵盖名词、动词、形容词和副词的语境词汇。这些词汇分别反映了具体传统价值观成语所关联的语境内容、行为方式及价值态度。我们据此对每个传统价值观成语的前20个高频语境词汇进行热力分析(如图5和图6),由此可以看到不同传统价值观成语在当代语境中的特点。

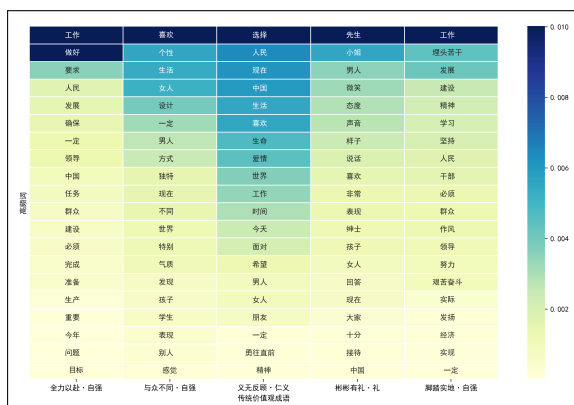


Figure 5: 多领域——正向高频传统价值观  
成语语境高频词信息图

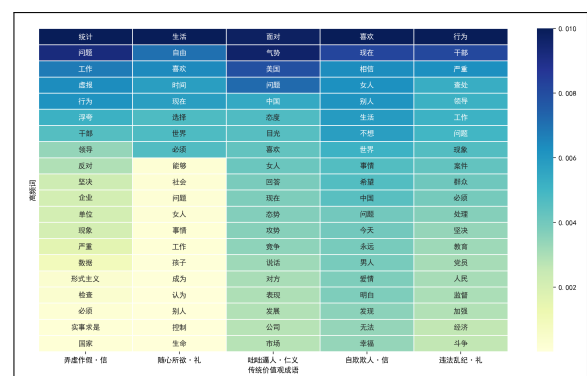


Figure 6: 多领域——负向高频传统价值观  
成语语境高频词信息图

由图5可以看出成语“全力以赴·自强”在语境中多伴随“一定”、“必须”、“重要”等具有强烈动员性主观色彩的副词以及“做好”、“要求”、“完成”等执行性行为，这些态度与行为均与诸如“工作”、“任务”、“建设”、“发展”等国家发展、经济建设及集体任务相关。这一成语语境强调任务导向和实干精神，凸显了“自强”价值观在国家与社会发展中的重要性；“与众不同·自强”的整体语境明显偏重主体意识与个性表达。语境中“生活”、“气质”、“感受”等凸显的主体意识的名词，搭配“喜欢”、“发现”、“表现”、“感觉”等注重心理活动的动词，外加进一步突出其个性化的形容词如“独特”、“不同”、“特别”等；“义无反顾·仁义”则侧重表现责任、信念与情感选择。它与“人民”、“生命”、“朋友”等高度责任化的名词相关联，并积极“选择”、“面对”、“勇往直前”，体现了高度的道德水准，同时语境中诸如“希望”、“一定”的副词则展现出强烈的积极性情感倾向；“彬彬有礼·礼”更多出现于人际交往与树立形象的语境中。面对诸如“小姐”、“孩子”等典型交际对象，往往礼貌“微笑”、“接待”、“回答”，“非常”、“十分”等副词符合言语会话中的礼貌原则，更加满足人们的礼貌期待；“脚踏实地·自强”类成语所建构的语境则突出现实主义与实践导向精神。“精神”、“群众”、“作风”、“实际”、“经济”等名词与“埋头苦干”、“坚持”、“努力”等动词体现着坚定的奋斗路径，同时“一定”、“必须”等副词则体现着人们坚定不移的态度。这体现出“务实奋斗”价值的社会诉求。如图6所示，“弄虚作假·信”类成语偏重出现在如“统计”、“问题”、“工作”、“行为”、“干部”、“领导”等高度严肃的话语语境中。诸如“虚报”、“查处”、“检查”、“反对”、“坚决”等强硬的行为动词常与“严重”、“形式主义”、“实事求是”内容有关。可以看到，这一传统价值观成语多出现在讨论揭发并处理组织制度中的违反纪律行为的语境中，鲜明反映出公众对造假及形式主义等不合理现象的高度抨击与舆论批判；而“随心所欲·礼”的语境则与“自由”、“选择”、“问题”、“生命”、“社会”等观念行为及领域相关，诸如“能”、“喜欢”、“认为”、“控制”、“成为”等主观性动词具有强烈个人主义色彩。可以看到当代人“反传统”的多元价值取向；从“咄咄逼人·仁义”语境中的如“气势”、“问题”、“态度”、“对方”等较为严肃性名词，如“回答”、“表现”、“发展”、“面对”、“竞争”等互动性与对抗性行为动词可以看出整体语境多呈现对抗性和对立性特点，常表达公共话语中的压迫感与侵略性；“自欺欺人·信”出现在“相信”、“别人”、“世界”、“希望”、“今天”、“爱情”、“明白”等个人情感泛滥的语境，凸显出理想与现实之间的落差状态，语境中涉及“女人”、“男人”主体，更多涉及到“爱情”、“生活”等话题。此外，“无法”则表明这一语境也含否定性情感，展现出人们对谎言与幻想的痛苦情感；“违法乱纪·礼”所处语境最具现实特点，如“党员”、“干部”、“案件”、“人民”等政治性名词、“查处”、“处理”、“监督”等政治类动词。反映出鲜明的政治纪律与社会治理话语语境，可以看出公众对腐败、违法行为的强烈谴责以及对制度监督完善的期待。

综上所述，十类高频传统价值观成语在当代语料中呈现出各自独特的语境面貌。我们在保留了传统价值的核心概念的基础上，通过具体语境词汇得出传统价值观成语在现实社会结构中的当代表达。

## 5.2 传统价值观成语当代语境社会话题及道德特征

由于报刊语料为当代报道，因此本节中我们对BCC中传统价值观成语报刊语料与多领域语料使用哈工大停用词词表进行分词，并得到词频> 500的词汇；进一步经过人工筛选，得到390个报刊领域有效高频词汇以及136个多领域有效高频词汇。同时我们又对这些词汇结合人工使用大模型（DeepSeek-R1）进行名词、动词、形容词词性标注；又对名词进行四类社会话题（政治、经济、文化教育与伦理规范、公共服务<sup>3</sup>）划分标注；对动词与形容词进行道德属性判断，即符合道德与违反道德；进而再对符合道德与违反道德词进行四种道德类型（社会公德、职业道德、家庭美德、个人品德）与对具体道德类型下面的词汇道德小类<sup>4</sup>进行判断标注。具体标注信息见附录E。

<sup>3</sup> 公共服务主要为“家庭与人际关系、社会保障与福利、法律与秩序、城市与社区生活、环境与生态、媒体与传播、医疗与健康”较细话题。

<sup>4</sup> 道德类型与道德小类均来源于山东省公民基本道德行为四十则。

5.2.1 传统价值观成语当代语境社会话题特征

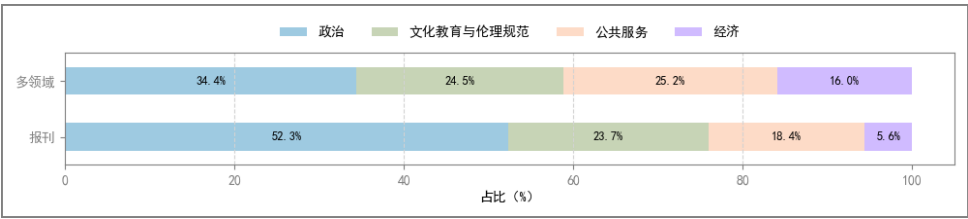


Figure 7: 跨领域传统价值观成语社会话题对比图

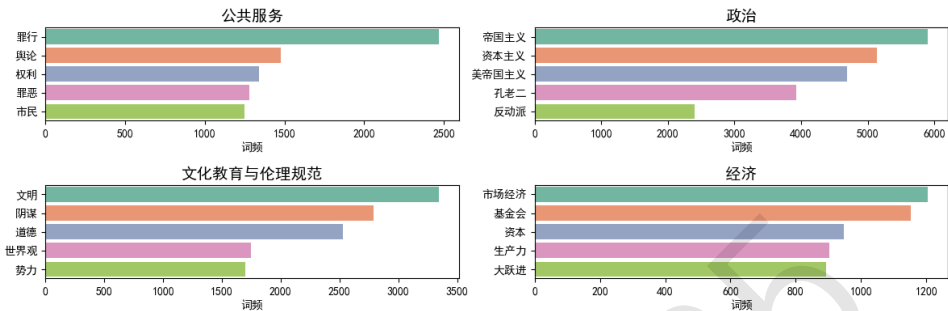


Figure 8: 报刊领域传统价值观成语高频社会话题图

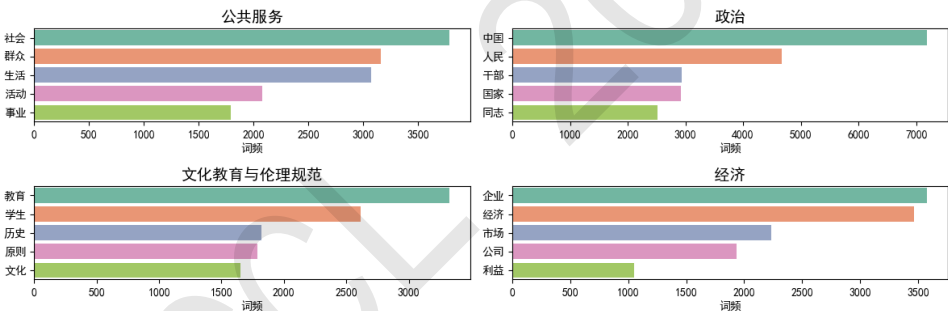


Figure 9: 多领域传统价值观成语高频社会话题图

我们对BCC不同领域下的不同社会话题语料频次进行统计分析，得到了有关社会话题的跨领域分布结果（如图7），同时又对不同领域下不同社会话题的前5个占比较高的主题词进行展示，有助于我们详细了解传统价值观成语在当代话语语境中的主题特征。如图7，传统价值观成语在报刊领域及多领域的社会话题分布多集中于“政治”，“经济”类话题占比较少。但多领域中的“文化教育与伦理规范”及“公共服务”话题分布较为均匀。如图8与图9，传统价值观成语语境在当代报刊话语中的“公共服务”中更关注“罪行、舆论、权利”等实际性的词汇；而在多领域中则会关注“社会、群众”等；同时在这两个领域中的“政治”话题也不同，报刊多为历史性的事物，范围为国际，而多领域则更注重当代国内；在“文化”话题方面，报刊中则会更注重更为普适的概念，多领域则会注重具体生活；在经济领域则差别不大。

5.2.2 传统价值观成语当代语境道德特征

我们对传统价值观成语语料中的道德词进行道德属性标注（符合道德与违反道德），但我们认为无论符合还是违反均是有关道德探讨。还统计了BCC不同领域的道德属性（图10）及道德类型（图11）的分布。如图10，多领域中传统价值观成语语境的符合道德词频次远高于违反道德词频次；而在报刊中，虽然符合道德频次高于违反道德频次，但违反道德词频次则略高。

这说明，由于多领域语料来源于社交媒体等平台，当代人们的交流多显礼貌，人与人之间比较友爱和谐；而官方报道则需要承担更多责任，起到社会宣传的作用，由此可以看到我国官方对社会的管理与把控是极为严格的。

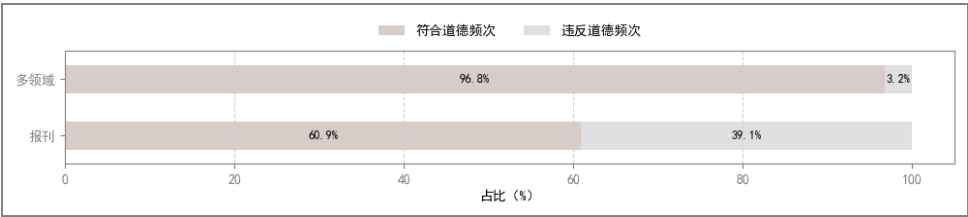


Figure 10: 跨领域传统价值观成语语境正负向道德词频次对比图

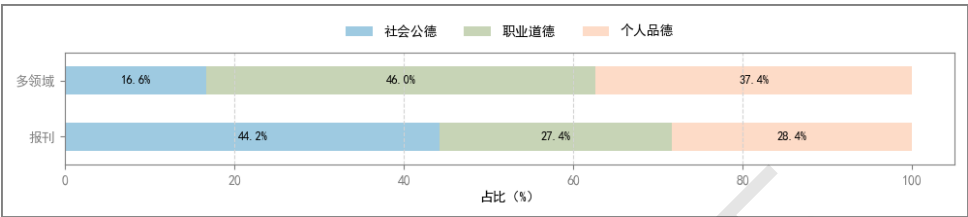


Figure 11: 跨领域传统价值观成语语境道德类型对比图

如图11，公民四类道德类型，传统价值观成语语境中只涉及三种，“家庭美德”没有出现。这可能与我们的传统价值观成语注重“仁义、礼、智”等如何修身、如何做人如何利他等相关理念，因此几乎不出现在探讨家庭关系的语境当中。值得关注的是，报刊领域代表着当代中国官方话语，传统价值观成语在此领域中则常出现在“社会公德”道德语境之中，更关注社会运行中人们之间的秩序问题。传统价值观成语在此领域中的职业道德与个人品德语境分布则较为均衡。同理，由于多领域的自由个人性特点，传统价值观成语则常常出现于职业道德、个人品德当中。这一分布差异反映了福柯提出的‘自我规训技术’理念。报刊作为主流媒体，力求通过道德话语构建‘国家-社会-个体’的权力关系，强调社会责任；而社交平台更多是个体道德的实践空间，个体可以选择性地表达职业理想与自我期许。

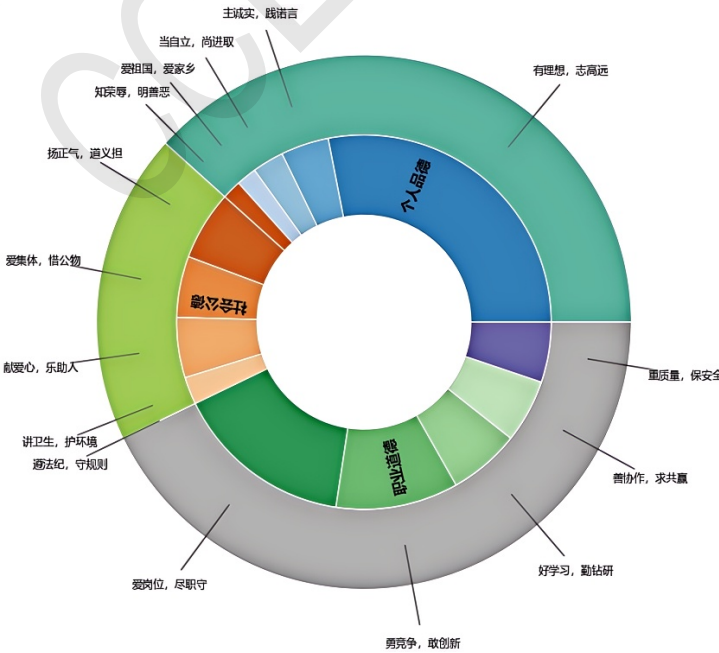


Figure 12: 多领域传统价值观成语语境道德类型高频道德小类分布图



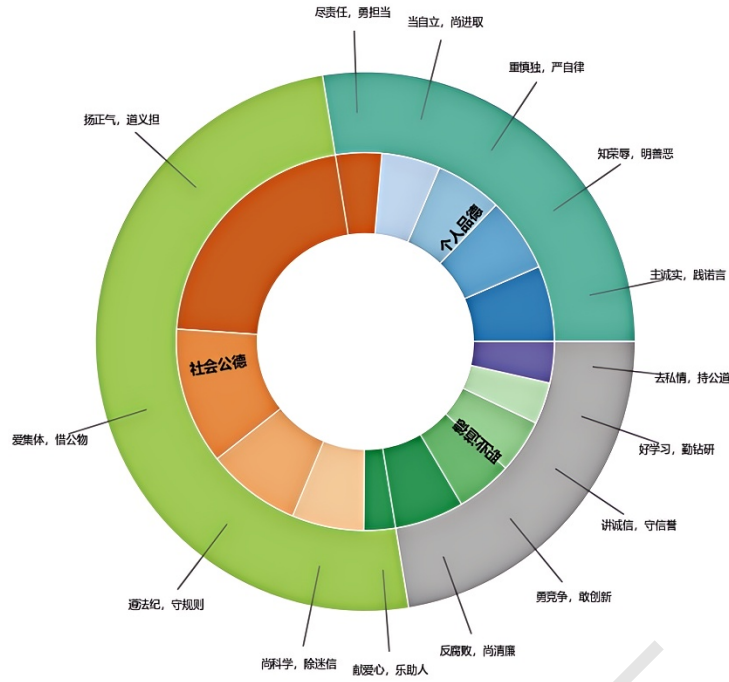


Figure 13: 报刊传统价值观成语语境道德类型高频道德小类分布图

我们仅展示报刊及多领域中不同道德类型下的5个高频道德小类分布，如图12和13（全部道德小类分布见附录F）。图12中，成语在多领域社会公德中所出现的小类最多的为“扬正气”与“爱集体”；在报刊中“扬正气”则占据很大比重。传统价值观成语在报刊语境中的“个人品德”与“职业道德”各小类占比较为均匀，而在多领域语境中的“个人品德”则主要为“存理想”以及在“职业道德”中则主要为“爱岗位”。可以看出传统传统价值观成语在报刊语境中的道德小类分布较为均匀，讨论较为全面，而在多领域中则较为集中。

### 5.2.3 传统价值观成语当代语境社会话题及道德适应性

#### (1) 社会话题及道德适应性计算方法

为评估传统价值观成语在当代语境社会话题、道德类型及其道德小类的适应性特征，本文结合成语语境社会话题和道德类型及道德小类在BCC跨领域中的分布频率与强度提出了综合适应性指数 $A$ （Adaptability Index），这一指数包含两个维度：

- **分布广度  $G$** ：表示该道德小类及社会话题所涉及的BCC领域（报刊与多领域）数量，范围为0（未涉及）、1（一种领域）或2（两种领域）；
- **分布强度  $S$** ：表示该道德小类及社会话题在BCC所有领域出现的总频次，定义为

$$S = f_{\text{多领域}} + f_{\text{报刊}} \quad (1)$$

为了统一计量，分别对 $G$ 和 $S$ 的极差进行标准化计算，则有：

$$G' = \frac{G - G_{\min}}{G_{\max} - G_{\min}}, \quad S' = \frac{S - S_{\min}}{S_{\max} - S_{\min}} \quad (2)$$

将最终的社会话题及道德小类适应性指数确定为：

$$A = \alpha \cdot G' + \beta \cdot S', \quad \text{其中} \alpha = 0.4, \beta = 0.6 \quad (3)$$

而对于道德类型适应性指数，我们根据道德小类分布确定如下：

$$A_{\text{道德类型}} = \frac{1}{n} \sum_{i=1}^n A_{\text{小类}}(i) \quad (4)$$

其中:

- $n$  为该道德类型下的道德小类种类 (最多10 个小类) 。
- $A_{\text{小类}}(i)$  为第 $i$  个小类的适应性指数。

(2) BCC传统价值观成语语境社会话题及道德适应性分析

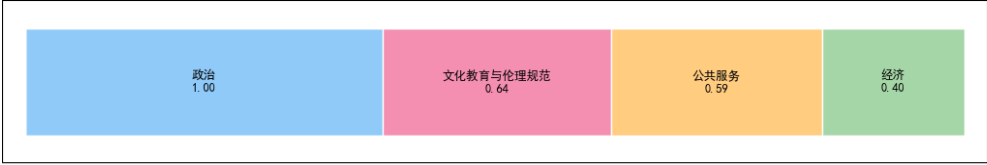


Figure 14: 传统价值观成语社会话题适应性

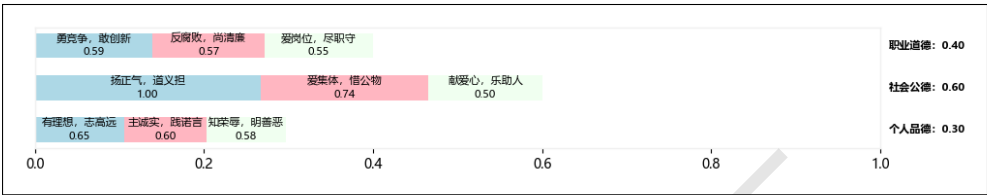


Figure 15: 传统价值观成语各道德类型及Top3小类适应性 (比例结构+实值标注)

如图14, 传统价值观成语的社会话题当代适应性排序为“政治>文化>公共服务>经济”。在“政治”与“文化”社会话题方面适应性较强, 话语语境较为宏大, 更关心政治与“文化”, 这与中国传统价值观的特点息息相关, “修身、齐家、治国、平天下”, 传统价值观成语在当代依旧有其独特面貌。如图15, 我们将道德小类适应性指数归一化, 每个道德类型仅展示适应性最强的前三个道德小类。成语在当代语境道德类型的适应性排序为社会公德>职业道德>个人品德; 在道德小类适应性中, 传统价值观成语对社会公德中的“扬正气”、职业道德中的“创新”、个人品德中的“理想”这三类适应性最强, 可以看出与我们当代的社会主义核心价值观融合度极高。无论是传统文化亦或是当代文化都实现了二者的有机结合, 生生不息地继续繁荣下去!

## 6 总结与讨论

基于语料库去观测传统价值观成语的当代语境表现是重要且有意义的。本文爬取BCC中2472个传统价值观成语语料共140余万条, 统计计算了跨领域传统价值观语料分布数量和成语传统价值观跨领域分布偏好特征, 同时我们标注了高频传统价值观成语在多领域中的语境情感倾向、得到了相应语境高频词, 更将报刊及多领域中的传统价值观成语语料高频词进行词性标注, 得到相应的名词与动词及形容词, 我们又对名词进行四类社会话题标注, 对动词与形容词进行道德类型及道德小类标注, 提出了传统价值观成语的当代适应性指数计量方法, 由此得到了精确有趣的传统价值观成语的当代语境表现特点。

我们发现, BCC不同领域对成语的传统价值观有着不同的偏好, 这与不同领域的语料特点有关, 更与其所偏好的传统价值观是吻合的; 通过对多领域高频成语语境情感倾向的标注发现, 传统价值观成语的语境情感与其本身的情感是较为一致的, 同时也得到了相应的当代语境信息, 发现诸如“与众不同”、“义无反顾”等传统价值观成语均体现着强烈的当代中国社会特点; 通过对报刊及多领域的传统价值观成语语料进行社会话题与道德适应性计量发现, 社会话题当代适应性排序为“政治>文化>公共服务>经济”, 当代道德类型适应性排序为“社会公德>职业道德>个人品德”, 我们又分别对不同道德类型中的前三个适应性极强道德小类进行展示。由此看出, 传统价值观成语在当代社会的话语语境偏宏大叙事, 更关注社会状况。可以看出, 本次研究是对传统文化的时代精神面貌的一次有益探索。

当然, 当前研究覆盖率还不够高, 对传统价值观成语的当代社会适应性挖掘还不够细致深入, 未来计划将立足于更全面的语料, 做更全面的计量研究, 甚至可以按社会不同领域 (如社交媒体、新闻报道、文学书籍等) 分别进行细致的研究与探讨。

## 参考文献

- Liu P, Zhang S, Yu D, et al. 2022. *CoreValue: 面向价值观计算的中文核心价值-行为体系及知识库(CoreValue: Chinese Core Value-Behavior Frame and Knowledge Base for Value Computing)*. [C]//Proceedings of the 21st Chinese National Conference on Computational Linguistics. 2022: 417-430.
- Saeed J. L. 1997. *Semantics*. Oxford: Blackwell.
- Sun H, Du H, Wang Y, et al. 2024. *Construction and Analysis of a Diachronic Knowledge Base of Traditional Values in Chinese Idioms*. 2024 International Conference on Asian Language Processing (IALP). IEEE, 2024: 139-150.
- Wang H, Yu D, Liu P, et al. 2023. 中国社会道德变化模型与发展动因探究——基于70年《人民日报》的计量与分析(*The Model of Moral Change and Motivation in Chinese Society—The Vocabulary Analysis of the 70-year "People's Daily"*). [C]//Proceedings of the 22nd Chinese National Conference on Computational Linguistics. 2023: 289-299.
- 胡壮麟. 2002. 语境研究的多元化. 外语教学与研究, 2002, (03): 161-166+239.
- 李玉叶. 2015. 从《论语》中的成语看儒家文化之“仁”. 常州工学院学报(社科版). 2015, 33 (02).
- 刘宁生. 1984. 语境分析. 南京师大学报(社会科学版), 1984, (01): 76-82.
- 刘明. 2024. 中华传统文化语境下艺术疗愈思想的研究. 新传奇, 2024, (43): 35-37.
- 饶高琦. 2016. 时代精神: 基于1946年到2015年报刊语料和隐含主题模型的历史热词提取. 语言规划学研究, 2016.
- 饶高琦, 胡星雨. 2023. 自晚清到民国《申报》主题词的历时演变和分期. 南京师范大学文学院学报, 2023, (03): 10-22.
- 王婉. 2024. 当代视觉文化语境下传统戏曲艺术的传承与创新研究. 新传奇, 2024, (40): 37-39.
- 吴云. 2017. 中华成语文化与社会主义核心价值观教育探析. 教育现代化, 2017, 4(07): 178-180.
- 荀恩东, 饶高琦, 肖晓悦, 臧娇娇. 2016. 大数据背景下BCC语料库的研制. 语料库语言学, 2016(1).
- 曾锡. 2025. 代际冲突与家庭纽带在亚洲文化语境下的体现——以幽默温情电影《姥姥的外孙》为例. 喜剧世界(下半月), 2025, (03): 67-69.
- 朱述承, 苏祺, 刘鹏远. 2021. 基于语料库的我国职业性别无意识偏见共时历时研究. 中文信息学报, 2021, 35(5): 130-140.

## 附录A. 相关论文与资源链接

详见链接: <https://github.com/sunhaoqfbx/IBTVR25>

## 附录B. 跨领域成语传统价值观偏好分布计算方法

$$\mu_i = \frac{1}{n} \sum_{j=1}^n x_{ij}, \quad \sigma_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \mu_i)^2} \quad (1)$$

$$z_i = \frac{x_{ij} - \mu_i}{\sigma_i} \quad (2)$$

如式(1)所示,  $x_{ij}$  表示表2中第*i*个领域(如报刊、对话、多领域或文学)中第*j*类传统价值观成语的语料数量。 $\mu_i$  表示第*i*个领域中所有语料数量的均值, 即该领域中所有传统价值观成语语料数量的平均值。 $\sigma_i$  表示第*i*个领域中所有语料数量的标准差, 即衡量该领域内数量分布的离散程度。

如式(2)所示,  $z_i$  是**Z-Score**标准化值, 通过该值可以衡量每个传统价值观的语料数量相对于该领域均值的偏离程度, 并以标准差为单位。Z-Score为正表示该数据点大于均值, 为负则表示该数据点小于均值。

	俭	礼	让	仁义	孝	信	智	中庸	忠	自强
报刊	2487	38232	3972	167605	650	143775	162904	22010	56992	220789
对话	1262	12299	2183	30788	159	27957	28351	8618	3903	29641
多领域	1490	19519	1750	84250	717	44079	64798	1058	7194	131991
文学	1119	9312	2349	19455	110	17289	16105	4653	3165	14296

Table 2: 跨领域传统价值观成语语料分布数量表

附录C.BCC传统价值观成语语境标注信息及流程

正向成语	传统价值观	语料	标注
全力以赴	自强	4700	1000
彬彬有礼	礼	1300	1000
义无反顾	仁义	2700	1000
脚踏实地	自强	3058	1000
与众不同	自强	3175	1000

Table 3: 正向价值观成语标注数量信息表

负向成语	传统价值观	语料	标注
自欺欺人	信	1500	1000
违法乱纪	礼	1500	1000
随心所欲	礼	1751	1000
弄虚作假	信	3500	1000
咄咄逼人	仁义	1700	1000

Table 4: 负向价值观成语标注数量信息表

标注模型	Prompt	准确率	一致率
DeepSeek-R1	”请标注以下句子的情感倾向: {essay_text}。情感倾向为正向时标注为1, 负向为2, 中性为0, 按如下方式回复: 情感倾向: 1—2—0”	87%	95%

Table 5: 标注流程表。如上表, 准确率为模型标注准确率, 即三位标注人员中两位及两位以上标注人员均认可模型标注结果的概率; 一致率为人工修改一致率, 即三位标注人员针对模型标注结果不正确的语料进行二次标注的一致率, 以两位及两位以上标注人员的一致为标准。

附录D.BCC多领域传统价值观成语不同情感倾向语料示例

	彬彬有礼	全力以赴	与众不同	义无反顾	脚踏实地
正向	举止言谈如同传说中的王子那样文雅而~。	为了爱努力, 就要~!	后现代潮流气氛~, 不一样的视觉享受。	她为了梦想不屈不挠, 为了音乐~, 值得学习!	追求梦想一直是那么的~, 信你支持你
中性	~、态度和蔼的贝尔便向陪审团提出异议。	做到平时保证重点, 紧急情况下~。	曾有过大量的阅读, 野莽要求自己作品能~。	教师应该~为学生全面提供阅读时机、素材。	任何事情都必须从“一”做起, ~。
负向	~、卑躬屈膝的人, 竟是我的死敌!	没有~的人, 怎么会成功、快乐呢?	你的一切都是假的, 只有我才是真的~	~地去为李洪志实现他的政治目的当工具。	欺世盗名者往往比~的提升得更快

Table 6: 正向—符合传统价值观成语不同情感倾向语料示例



	弄虚作假	违法乱纪	咄咄逼人	自欺欺人	随心所欲
正向	统计部门坚决 坚持依法统计， 反对～。	不管谁～，都 要按照党纪国 法查处。	这款外套并 不～，非常 明亮活泼。	走自己的路， 不能虚假， 更不能～。	想有一个人在 我的世界里～
中性	把“～”作为 一种违纪类型 列入党纪条规。	如果不受法律 约束，一有漏 洞就会～。	～的笑声表 示情势均衡， 叫旁人放心。	如果仅考虑 反战立场是 ～的。	物境的壮伟是 由心境而然， 形式却是～
负向	不是我～，而是 残酷的现实是 我憎恨不过的。	那些严重～的 人，根本不配 做共产党员！	你们有基本 的宽容吗？ 一开始就～。	采用形式化 的考核办法 都是在～！	现实是不能～， 只能无限的 忍受着……

Table 7: 负向—违反传统价值观成语不同情感倾向语料示例

附录E.BCC报刊及多领域传统价值观成语语料高频词词性、社会话题及道德标注信息

词性、社会话题及道德类型标注	DeepSeek—R1准确率	人工核改一致率
词性 ①名词②动词③形容词	99%	100%
道德属性 ①符合道德②违反道德	99%	100%
社会话题 ①政治②经济③公共服务④文化教育与伦理规范	85%	96%
道德类型 ①社会公德②职业道德 ③家庭美德④个人品德	90%	97%

Table 8: 高频词词性、社会话题及道德标注信息表

道德小类标注	DeepSeek—R1准确率	人工核改一致率
社会公德 讲文明，树新风。懂礼貌，礼让谦。 献爱心，乐助人。济贫困，救危难。 爱集体，惜公物。讲卫生，护环境。 重低碳，节能源。尚科学，除迷信。 遵法纪，守规则。扬正气，道义担。	92%	100%
职业道德 爱岗位，尽职守。好学习，勤钻研。 讲诚信，守信誉。重质量，保安全。 勇竞争，敢创新。善协作，求共赢。 去私情，持公道。反腐败，尚清廉。 人为本，优服务。任劳怨，乐奉献。	94%	100%
个人品德 爱祖国，爱家乡。尽责任，勇担当。 当自立，尚进取。有理想，志高远。 重情义，知感恩。主诚实，践诺言。 知荣辱，明善恶。重慎独，严自律。 衣得体，品貌端。惜生命，身心安。	90%	100%

Table 9: 道德小类标注信息表

附录F.传统价值观成语语境道德小类分布

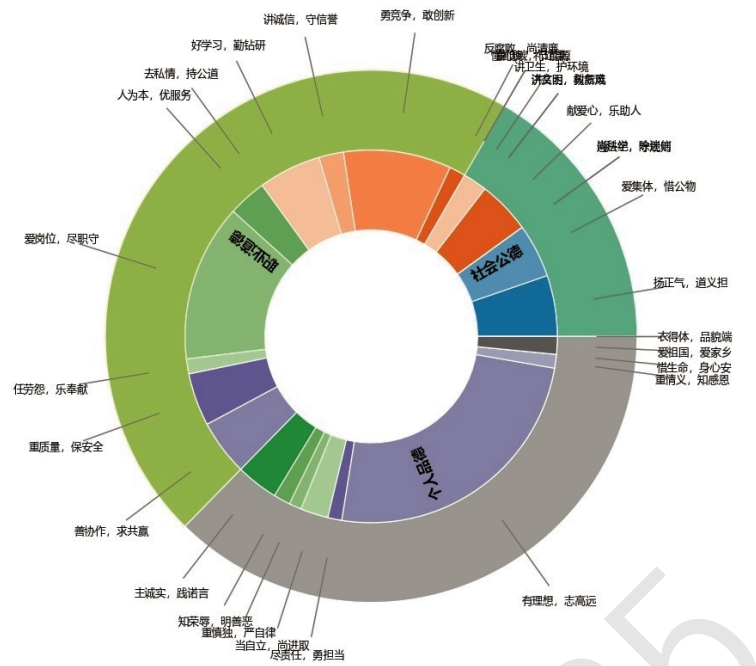


Figure 16: 多领域传统价值观成语语境道德小类分布图

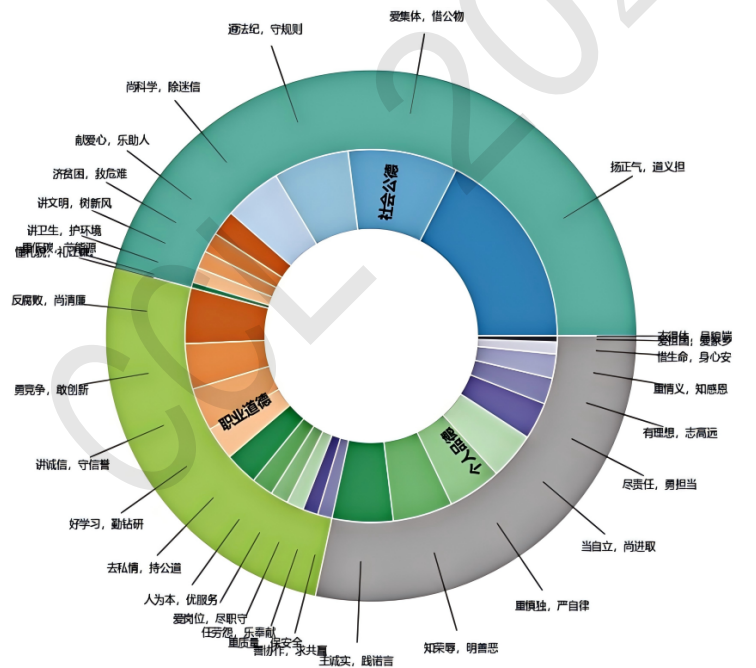


Figure 17: 报刊传统价值观成语语境道德小类分布图