在线视频社区知识交流效果比较测度研究*
——以哔哩哔哩网站的结构化评论功能为参照

舒科慧　过其桥　田园　程秀峰　杨金庆
（华中师范大学信息管理学院，武汉 430079）

摘要：从知识交流视角出发，围绕在线视频社区中的结构化评论功能，测量视频社区中的知识交流效果，为促进视频社区高质量知识交流提供有效参考。选取代表性在线视频社区哔哩哔哩，采集其中的视频、评论及行为数据，利用主题聚类、相关性分析等方法，测量结构化评论与一般评论在知识交流方面的效果，同时分析行为属性数据、视频文本数据及评论文本数据之间的关联关系。相较于一般评论，结构化评论的数量与视频播放量关联不大，更多取决于视频内容质量和用户的知识交流意愿。此外，结构化评论在提升知识交流深度方面具有优势，尤其对拥有特定观众群体的小众视频效果显著。因此，结构化评论功能具备增强在线视频社区知识交流的能力。

关键词：结构化评论；在线视频社区；知识交流；评论挖掘；哔哩哔哩
中图分类号：G350.7　DOI：10.3772/j.issn.1673-2286.2024.10.005
引文格式：舒科慧，过其桥，田园，等. 在线视频社区知识交流效果比较测度研究：以哔哩哔哩网站的结构化评论功能为参照[J]. 数字图书馆论坛，2024，20（10）：42-52.

作为社交网络的重要组成部分，在线视频社区中丰富的视频内容和热烈的讨论氛围极大促进了信息交流，社区功能的持续更新为信息交流带来了更加多元的选择空间和创作动力。通过社区成员分享、讨论和传播的具有精炼性、专业性和便捷性的内容是在线视频社区的重要知识形态[1]，知识交流是信息交流的一种具备特殊使用和交换价值的形式，其交流效果受到多种因素的影响，这些因素可能存在组合效应[2]。因此，在线视频社区中知识交流效果影响因素的发现、测度及评价是社交网络研究的重要方向之一。

评论功能是在线社区保持用户参与的重要形式，评论内容往往被视为知识交流的有效信息来源[3]。一般评论通常只包含自由文本或简单表情，而哔哩哔哩网站的笔记功能（以下称为结构化评论）允许用户根据网站提供的格式进行评论文本的规范化编辑，编辑内容不仅包括文本，还包括图片和时间戳等，这不仅便于用户的自我保存和分享，也使得知识的呈现更为生动、内容更为严谨。本研究采集并统计哔哩哔哩网站中2021年5月—2024年2月30余位主要视频创作者[4]的视频评论数据，发现自2022年10月起，结构化评论的发布数量增长迅速，2023年2月—2024年2月每月数量维持在5 000~7 000条，每月占比维持在4%~5%，且未来有望继续增长。这表明结构化评论提供了一种新的知识组织方式，能够扩大评论所含的信息量，对用户积极参与知识交流产生了重要影响。因此，本研究从以下两个问题切入：①与一般评论形式相比，结构化评论的形式是否更能促进在线视频社区的知识交流？②结构化评论和一般评论在知识内容之间的关联、主题的分布及演

收稿日期：2024-06-16
*本研究得到国家自然科学基金项目“基于可解释机器学习的科学知识角色转变预测研究”（编号：72304108）、中央高校基本业务费项目“基于技术要素深度语义理解的创新路径识别及引导策略研究”（编号：CCNU24ZZ140）资助。

变关系上是否存在差异？

本研究以在线视频社区哔哩哔哩的结构化评论功能为参照，对结构化评论和一般评论的知识交流效果进行比较测度，通过分析并解读实验结果，探讨知识交流的影响因素，并提出相应优化策略，为实现在线视频社区的高质量知识交流提供有效参考。

1 文献回顾

知识交流是知识主体借助某种符号系统，围绕知识开展的一系列活动[5]。在线社区中的知识交流则表现为有价值的知识内容与知识主体的资源通过一定渠道相互作用并巩固的过程[6]。知识交流效果是知识交流过程中达到预期目标或产生正面影响的效果[7]，目前知识交流效果测度有多种实现方式。如以用户主观报告（访谈、问卷）为参考标准，引入内容质量[8]、用户体验[9]、用户交互[10]等客观要素进行建模，验证或创建有解释力的理论和模型；又如基于用户行为数据，提取平台流量、访问量、用户数量、用户评论等指标，衡量知识交流效果。前者难以从客观角度捕捉交流的细节[11]，后者又会导致结果的复杂性与不一致性。

对知识交流效果进行测度需要选择合适的指标与方法[12]。对相关研究进行梳理后发现，多数研究集中在内容和用户两个方面：前者从知识内容（主要是文本）挖掘出发，提取文本、形成特征、创建指标，这些指标主要包括（知识的）权威性、可信度[13-14]、相似性[15]、类别[16]、及时性[17]、交互性及娱乐性[18]等，通过基于语料的质性分析[19-20]、主题聚类分析、情感分析[21-23]等多种自然语言处理方法可以有效地对文本内容进行挖掘；后者从用户行为挖掘出发，首先进行属性建模，而后进行指标创建，这些指标主要包括交互行为[24]、自我效能[25]、情感信任[26]、归属感[27]、报酬激励[28]、社区地位[29]、隐私披露[30]和行为关系[31]等，借助相关性分析或回归分析[32-34]等统计工具，可以将指标量化，从而测度知识交流效果。

可对以上两类研究中的指标与方法进行选择性融合，形成多维度指标体系，从而丰富知识交流效果测度依据。相较于其他类型社区，在线视频社区鼓励用户生成内容，且用户往往对特定内容有浓厚兴趣，其交流方式的即时性和互动性强。因此，在指标选取上，选择内容的相似性、内容的交互性以及用户的交互行为3类指标来共同表征知识交流效果。其中，内容的相似性与交互性和在线视频社区中的评论回复机制相契合。相似性反映了知识交流的一致性和准确性，衡量用户接收和理解知识的程度；交互性反映了知识交流的活跃度和深度，衡量用户传播和扩散知识的效果。用户的交互行为包括评论、点赞、收藏等，这些行为指标直接反映了用户在知识交流过程中的参与度和积极性。在测度方法上，采用融合了统计分析、行为属性分析、文本相似度分析以及主题聚类分析的综合分析方法开展实证研究，对在线视频社区中的知识交流效果进行比较测度。

2 研究设计与方法

2.1 研究设计

基于在线视频社区的真实数据，根据两个研究问题设计实验，解析实验结果，并回答相应问题，具体研究框架设计如图1所示。

首先从视频网站中获取不同分区的视频和评论数据，采用描述性统计分析法，利用视频发布总数、评论总数和评论回复总数等直观反映不同视频分区下知识交流的整体效果。其次，针对问题一，采用行为属性分析，通过对行为属性数据的比较测度，探究两种评论形式中用户行为的差异；针对问题二，通过对内容相似性和互动性的分析，挖掘视频文本数据与评论文本数据之间的关联关系。最后，将评论文本数据划分为结构化评论、一般评论、结构化评论回复和一般评论回复4类，通过主题聚类横向和纵向比较各种评论类型，分析两种评论形式下知识的关联、分布及演变。

2.2 研究方法

2.2.1 行为属性分析

感知价值理论认为，知识主体在知识交流过程中的行为会受到感知价值的正向影响，包括用户有目的地选择、分析、评价、接受和利用知识的过程[35]。在线视频社区用户的播放、点赞、投币、收藏、评论等行为属性体现了用户对视频知识内容不同程度的感知[36]。在线视频社区情境中，播放量是视频受欢迎程度的最直观

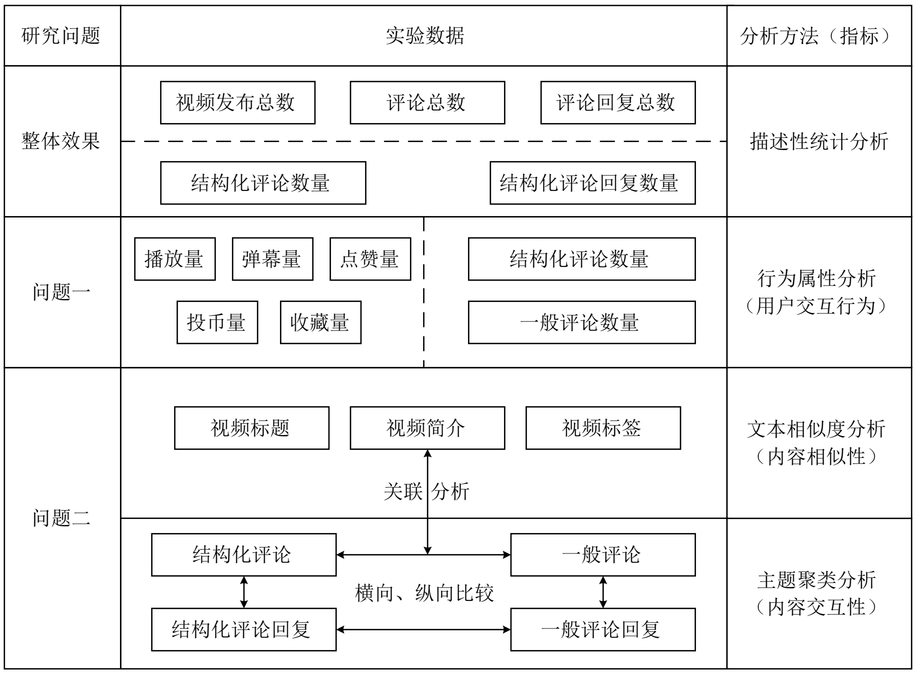

图1　研究框架

表现；弹幕量代表用户的知识交流意愿；点赞量代表用户对视频内容价值的认可程度；投币量代表用户对视频内容价值的更深层次的认可程度，这是由于哔哩哔哩中“硬币”有使用限制，用户通常只能对少量视频投币；收藏量代表用户未来的知识交流意愿，而非即时交流意愿；评论量包括结构化评论与一般评论的发布数量。通过对行为属性数据的量化分析，可以得到属性的相关性权重及排序，便于检验其影响因素。

2.2.2　文本相似度分析

主题关联是知识局部动态和演化趋势分析的前提，可以揭示主题发展和主题间交互作用的知识交流特征[37]。由于主题之间存在一定的语义关联，能以词向量的形式量化语义距离，可以通过文本主题相似度分析推导文本之间的关联关系[38]。

给定A和B两个文本集合，其中分别有m和n个文本单元，每个单元都以词向量的形式量化。$\vec{v}_{A,i}$为文本A中第i个单元的主题向量表示，$\vec{v}_{B,j}$为文本B中第j个单元的主题向量表示。两单元之间的余弦相似度计算公式如式（1）所示。

$$\text{CosineSimilarity}\left(\vec{v}_{A,i},\vec{v}_{B,j}\right)=\frac{\vec{v}_{A,i}\cdot\vec{v}_{B,j}}{\left\|\vec{v}_{A,i}\right\|\times\left\|\vec{v}_{B,j}\right\|} \qquad (1)$$

式中：$\vec{v}_{A,i}\cdot\vec{v}_{B,j}$表示两个向量的点积；$\left\|\vec{v}_{A,i}\right\|$和$\left\|\vec{v}_{B,j}\right\|$分别表示两个向量的欧几里得范数（向量的长度）。计算所有可能的两两单元组合的平均余弦相似度，如式（2）所示。

$$\text{AvgSimilarity}\left(A,B\right)=\frac{1}{m\times n}\sum_{i=1}^{m}\sum_{j=1}^{n}\text{CosineSimilarity}\left(\vec{v}_{A,i},\vec{v}_{B,j}\right) \qquad (2)$$

若两个文本集合存在先后关系，相似度较高则表明文本内容间存在较多重叠，后者是前者知识内容的延伸，知识交流呈现出一定的集中性和延续性的特征；反之则表明文本内容间差异较大，后者是前者知识内容的扩散，知识交流呈现出一定的分散性和间断性。

2.2.3　主题聚类分析

主题聚类通过生成概率模型从文本集合中梳理潜藏主题[39]。传统的机器学习方法以LDA[40]及其变体为代表，虽然在主题聚类中已经可以提供具有意义的结果，但模型的表征能力有限，对于文本背后的语义、结构、序列和上下文理解不足，这在短文本处理中表现得更为突出[41]，而BERTopic模型能有效弥补这些缺陷[42]。如图2所示，BERTopic是一种无监督模型，对原始文本进行预处理后，通过BERT预训练模型实现文本嵌入生成，经过UMAP算法降维[43]和HDBSCAN算法聚类[44]得到数个主题簇，并使用C-TF-IDF算法评估各

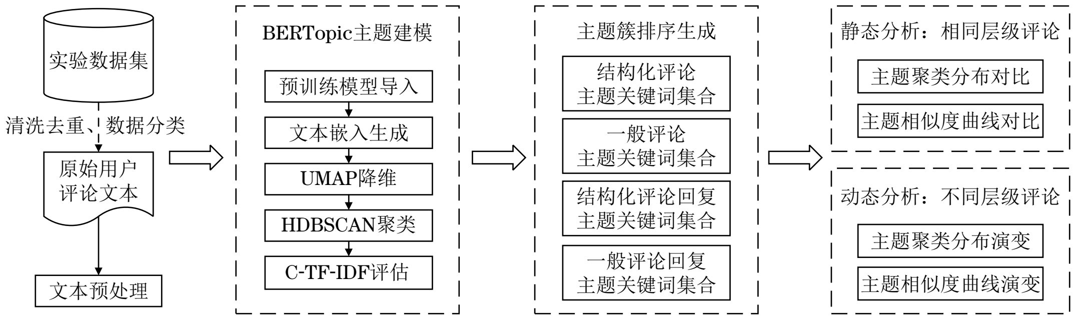

图2　BERTopic主题建模过程与研究方法

主题词重要程度，最终生成相应的主题关键词集合。

3　实验过程与结果分析

3.1　数据获取和预处理

在线视频社区哔哩哔哩月均活跃用户数达2.72亿人次，形成了围绕用户、创作者和内容的高质量社区生态系统[45]，因此以该网站为研究对象。如图3所示，首先，确定其知识类视频在网站中的具体分区类别为科技、美食、汽车、生活、游戏、知识、资讯、动物圈和其他[1]；随后，为保证视频和评论发布的数量规模，并涵盖视频热度较高的时间段，按照分区设定筛选条件，爬取2024年1月30日发布的视频在一周内的视频数据，并进行清洗、去重和分词，将其分为行为属性数据、视频文本数据和评论文本数据；最后，形成2 045组视频记录。

3.2　描述性统计分析

通过描述性统计方法分析结构化评论的整体规律。首先，以9个小类分区视频和评论的平均数量为划分依据，对9个小类分区进行分类：若视频发布总数大于215个或评论总数大于9 870条，则将该小类划分为大众视频（游戏类除外，虽然评论总数达到要求，但视频发布总数过低）；反之则划分到小众视频中，最终形成2个大类分区和9个小类分区，进而比较不同视频分区的知识交流情况。如表1所示，大众视频（美食、知识、动物圈、资讯和生活类视频）占视频发布总数的88.61%，其评论占评论总数的73.04%，其评论回复占评论回复总数的74.62%，这说明大众视频对用户具有较高的吸引力，反映了用户普遍的知识交流需求；小众视频（汽车、科技和游戏类视频）共占视频发布总数的6.41%，其评论占评论总数的22.47%，其评论回复占评论回复总数的22.65%，说明小众视频的用户互动性强、更具参与度，小众视频反映了特定用户的知识交流情况。从评论数与评论回复数占比的差值来看，评论回复数总体上多于评论本身，但在不同的视频分区中存在差异，如美食、动物圈、生活、科技和游戏分区的评论数多于评论回复数。

表2描述了结构化评论在总评论中的占比，以及结构化评论回复在总评论回复中的占比情况。可以看出，

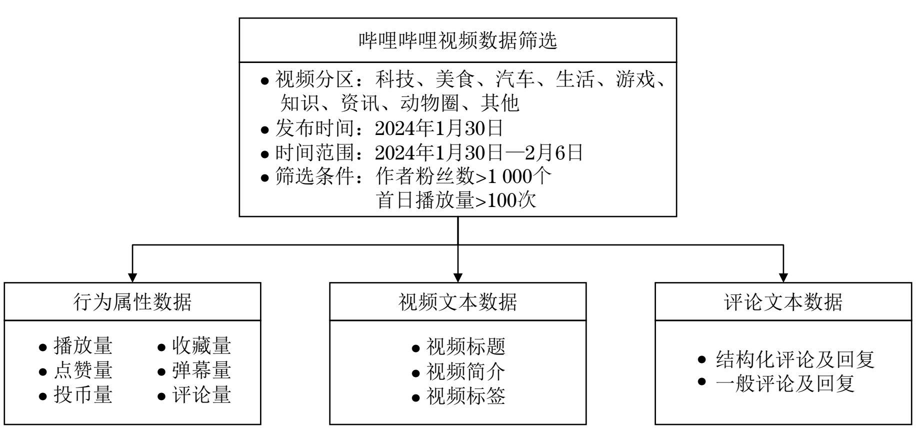

图3　数据筛选及分类过程

表1　视频发布总数、评论及其回复总数

大类分区	小类分区	视频		评论		评论回复		评论和评论回复占比的差值/%
		数量/个	占比/%	数量/条	占比/%	数量/条	占比/%	
大众视频	美食	665	32.52	20 985	22.56	25 316	16.29	6.28
	知识	437	21.37	16 488	17.73	41 133	26.46	−8.74
	动物圈	358	17.51	6 512	7.00	8 155	5.25	1.76
	资讯	210	10.27	13 783	14.82	26 678	17.16	−2.34
	生活	142	6.94	10 164	10.93	14 715	9.47	1.46
	总计	1 812	88.61	67 932	73.04	115 997	74.62	−1.59
小众视频	汽车	83	4.06	5 707	6.14	14 947	9.62	−3.48
	科技	35	1.71	2 490	2.68	4 089	2.63	0.05
	游戏	13	0.64	12 704	13.66	16 179	10.41	3.25
	总计	131	6.41	20 901	22.47	35 215	22.65	−0.18
其他分区	其他	102	4.99	4 177	4.49	4 229	2.72	1.77

目前结构化评论在各视频分区中的占比较少，但针对结构化评论的回复则较多。从结构化评论与其回复占比的差值来看，除动物圈（0.25%）和资讯（0.20%）两个分区外，结构化评论占比均小于其回复的占比，说明结构化评论一定程度上激发了用户的评论意愿，具体体现在评论回复的参与度上。在有效的知识交流过程中，知识主体较高的参与度有助于形成活跃的互动氛围，形成良性的知识循环，从而提升知识交流的深度与广度。

3.3　行为属性分析

为测度用户的交互行为指标，对结构化评论与一般评论的数量与行为属性数据进行斯皮尔曼相关性分析。在视频样本中筛选出具有结构化评论的468个视频，依次统计各视频的播放量、弹幕量、点赞量、投币量、收藏量、结构化评论数量和一般评论数量7种属性，斯皮尔曼相关性计算结果如表3所示。

表2　结构化评论及回复在总评论及回复中的数量占比

大类分区	小类分区	结构化评论		结构化评论回复		结构化评论和评论回复占比的差值/%
		数量/条	占比/%	数量/条	占比/%	
大众视频	美食	539	2.57	1 276	5.04	−2.47
	知识	332	2.01	837	2.03	−0.02
	动物圈	423	6.50	509	6.24	0.25
	资讯	46	0.33	36	0.13	0.20
	生活	323	3.18	730	4.96	−1.78
小众视频	汽车	148	2.59	459	3.07	−0.48
	科技	35	1.41	141	3.45	−2.04
	游戏	793	6.24	1 668	10.31	−4.07
其他分区	其他	127	3.04	173	4.09	−1.05

表3　行为属性数据的斯皮尔曼相关性矩阵

变量	播放量	弹幕量	点赞量	投币量	收藏量	结构化评论数量	一般评论数量
播放量	1.000						
弹幕量	0.803	1.000					
点赞量	0.928	0.767	1.000				
投币量	0.779	0.808	0.842	1.000			
收藏量	0.859	0.725	0.906	0.839	1.000		
结构化评论数量	0.603	0.635	0.611	0.652	0.582	1.000	
一般评论数量	0.843	0.828	0.822	0.794	0.750	0.659	1.000

注：全部在0.01级别（双尾）相关性显著。

整体上看，结构化评论数量与行为属性数据的相关性排序为投币量（0.652）、弹幕量（0.635）、点赞量（0.611）、播放量（0.603）和收藏量（0.582）；一般评论数量与行为属性数据的相关性排序为播放量（0.843）、弹幕量（0.828）、点赞量（0.822）、投币量（0.794）和收藏量（0.750），均在0.01级别（双尾）相关性显著。通过比较相关系数，可以看出一般评论数量与行为属性数据的相关性总体上大于结构化评论，说明其对用户行为的影响强于结构化评论，原因可能是结构化评论数量和占比较低。

具体到与各个行为属性的相关性，两类评论与不同行为属性的关联揭示了知识交流效果在不同层面的特点。投币量（0.652）与结构化评论数量最为相关，说明内容价值受到用户高度认可的视频普遍有更高的结构化评论数量，反映出用户在认可知识内容价值时，更倾向于选择便于表达明确主题和观点的结构化评论，进而增强信息价值，提高知识交流的专业性和深度。播放量（0.843）则与一般评论数量最为相关，说明更受欢迎的视频普遍有更多的一般评论，而非结构化评论，这表明一般评论相对而言更随意与普遍，用户倾向于选择便于情感表达与简单互动的功能。

3.4　文本相似度分析

通过文本相似度分析探究视频内容与用户评论的关联。首先将筛选数据集中的评论文本数据分为结构化评论、一般评论、结构化评论回复和一般评论回复4类，通过BERTopic模型生成文本向量表示后，利用式（1）与式（2）按分区依次计算视频文本数据与4类评论文本的相似度，结果如图4所示。

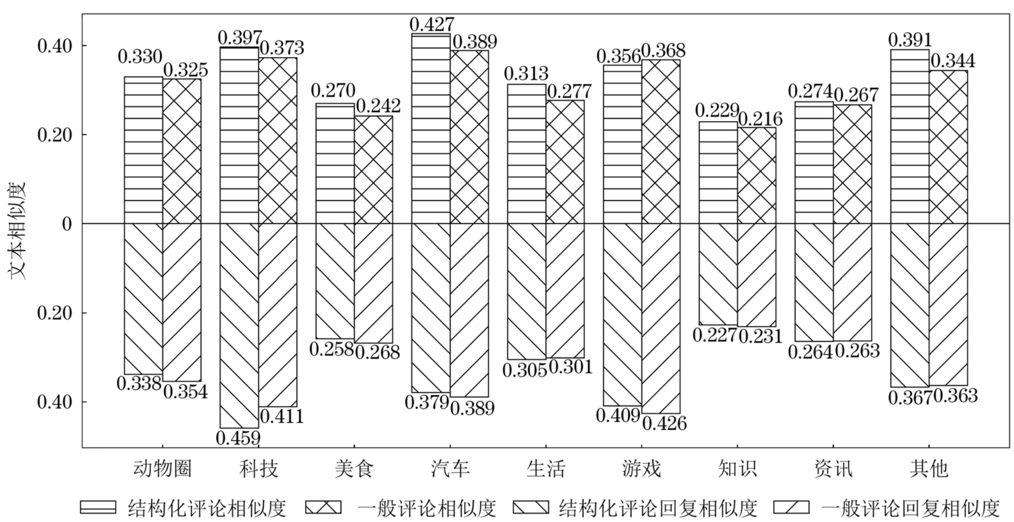

图4　视频文本数据和评论文本数据的相似度

科技、汽车和游戏区的评论文本与视频文本数据的4类文本相似度平均值排名前3，分别是0.410、0.396和0.390。科技、汽车和游戏均属于小众视频，这说明随着结构化评论功能的加入，小众视频的评论内容与视频内容更为贴近，可以看作视频知识内容的延伸。这种高度相似性侧面表现出该类分区中的知识交流呈现集中性和延续性的特点，即用户在小众视频中的交流更聚焦于原始知识内容本身。

除游戏区外，结构化评论文本与视频文本数据的相似度高于一般评论，且游戏区中两者数值十分接近。这说明结构化评论对视频知识交流的促进作用强于一般评论，通过提供更具针对性和深度的内容，在一定程度上弥补了一般评论分散性和间断性的不足。而结构化评论回复和一般评论回复与视频文本数据的相似度与视频分区更相关，这说明结构化评论对视频知识交流的影响主要体现在评论本身，而非评论回复。这可能是由于结构化评论回复在形式上受到限制，与一般评论类似，只能使用自由文本和简单表情，难以继承结构化评论的优势。

科技、游戏和汽车区中，结构化评论回复相似度与结构化评论相似度差值的绝对值排名前3，分别是0.063、0.053、0.048，说明评论文本呈主题集中或发散趋势的可能性同时存在，且在小众视频中该趋势更为明显。而除资讯区外，一般评论回复与视频文本数据的相

似度均高于一般评论，说明虽然一般评论与视频内容的关联性较弱，但其评论回复与视频内容的关联性增强，一般评论回复有利于缩小一般评论在知识表征方面与结构化评论的差距。

3.5 主题聚类分析

通过BERTopic进行主题聚类，从评论文本的主题分布及演变的角度探究其对知识交流的影响。将评论文本划分为结构化评论文本、一般评论文本、结构化评论回复文本和一般评论回复文本4类。考虑到数据集规模不同，选取4类评论文本中频次排名前16的主题，并按照分区主题总占比降序排列，从而直观展现各视频分区的主题分布情况，以及各类评论文本主题的分区占比排名，如表4~表5所示。

3.5.1 结构化评论及其回复的主题聚类分析

结构化评论的主题涵盖了全部的9个视频分区，其回复的主题也涵盖了8个分区，说明结构化评论的高频主题的分区覆盖率高。相较于结构化评论，其评论回复在知识和游戏两个视频分区的占比相对增加，而在生活分区的占比相对减少，这一变化可能反映了用户在参与结构化评论回复时，更倾向于深入探讨特定领域的专业知识。从主题内容的深度和具体化程度来看，知识分区主题从较为宽泛的事件解析、文学作品和视频剪辑深化为自我价值、股市行情和交友关系；游戏分区主题从较为宽泛的游戏攻略具化为角色强度、版本环境和角色培养。这种主题具体化的知识交流有助于用户对细分领域的选择。上述数据表明，结构化评论虽然在一定程度上可以维持知识交流规模，但侧重于促进小众视频主题内容的具体化和专业化，而非大众视频主题，这种趋势可能是由结构化评论鼓励针对性和深入知识交流的效果决定的。

3.5.2 一般评论及其回复的主题聚类分析

一般评论的主题涵盖了8个视频分区，其回复的主题涵盖了7个分区，说明一般评论的高频主题的分区覆盖率相对结构化评论较低。相较于一般评论，其评论回复在美食和生活两个视频分区中的占比有所提升，而科技分区的占比排序有所下降。在一般评论的知识交流过程中，用户更倾向于参与日常生活和美食等大众话题的讨论，而不是科技等可能需要特定知识的领域。从主题内容的深度和具体化程度来看，美食分区主题从寿司起源变为美食料理和西式快餐；生活分区主题从住房问题、家用电器和旅行攻略变为学习生活、旅行攻略和春节生活。尽管美食和生活主题占比有所提高，但主

表4 结构化评论及其回复的主题聚类结果

结构化评论				结构化评论回复			
分区	主题	占比/%	排序	分区	主题	占比/%	排序
生活	工作生活	6.11	2	知识	自我价值	4.31	2
	旅行攻略	3.28	5		股市行情	4.31	4
	服饰装备	1.09	15		交友关系	2.19	9
美食	美食料理	8.08	1	美食	美食料理	4.67	1
	寿司起源	1.86	10		西式快餐	4.31	3
知识	事件解析	5.24	3	游戏	角色强度	2.48	8
	文学作品	2.18	8		版本环境	2.12	10
	视频剪辑	2.07	9		角色培养	1.61	12
动物圈	宠物驯养	2.62	7	动物圈	宠物饲养	4.01	5
	宠物性格	1.53	12		流浪动物	1.31	14
汽车	汽车测评	3.60	4	生活	居家购物	2.70	7
游戏	游戏攻略	1.75	11		生活体验	1.17	15
	抽卡扭蛋	1.20	14		日常分享	1.17	16
资讯	国际局势	2.62	6	汽车	汽车测评	3.28	6
其他	魔术表演	1.31	13	其他	社区平台	2.12	11
科技	脑机接口	1.09	16	科技	脑机接口	1.53	13

表5　一般评论及其回复的主题聚类结果

一般评论				一般评论回复			
分区	主题	占比/%	排序	分区	主题	占比/%	排序
资讯	国际局势	4.64	1	资讯	国际局势	3.40	2
	社会与法	0.49	12		社会与法	0.86	9
知识	股市行情	2.58	2		婚姻话题	0.51	16
	名校课程	1.01	5	美食	美食料理	3.94	1
	心理健康	0.86	6		西式快餐	0.59	14
	理性思维	0.47	13	生活	学习生活	1.86	5
科技	脑机接口	1.54	3		旅行攻略	0.96	7
	小米品牌	0.55	9		春节生活	0.58	15
生活	住房问题	0.69	7	知识	股市行情	2.01	4
	家用电器	0.53	10		文学作品	0.64	11
	旅行攻略	0.43	14		农民问题	0.60	13
动物圈	流浪动物	1.53	4	动物圈	流浪动物	2.64	3
游戏	角色强度	0.43	15	科技	小米品牌	0.93	8
	抽卡扭蛋	0.41	16		星链计划	0.73	10
美食	寿司起源	0.68	8		脑机接口	0.60	12
汽车	汽车测评	0.52	11	汽车	汽车测评	1.27	6

题内容的宽泛性没有明显变化。上述数据表明，一般评论虽然可以保持或扩大知识交流规模，但更倾向于围绕大众视频主题展开，而非小众视频主题，这可能是因为一般评论通常更加随意，它们更多地反映用户的即时感受和浅层次的意见，而非深入的专业知识。

3.5.3　结构化评论与一般评论的比较分析

通过计算两两主题间的余弦相似度，得到4个子数据集主题的共60个相似度矩阵，按主题数升序依次计算其平均相似度，可以得到评论文本内容的集中或分散趋势，如图5所示。

结构化评论的主题相似度远低于一般评论，且变动更明显。但随着主题数增加，二者的相似度将趋于一致。这说明结构化评论便于激发知识交流中的思考与创新，使得知识交流的发散效果更明显，容易产生新的知识内容；而一般评论中知识交流的集中程度更高，需要通过更多的评论数达到类似结构化评论的效果。对于评论回复，两者主题相似度均呈现出先增长后平稳的趋势，但结构化评论回复的变动相对明显，且随着主题数增加，其相似度低于一般评论回复。这说明结构化评论对知识交流效果的影响主要体现在评论本身，而非评论回复，结构化评论回复可能因为形式上的限制未能继承原评论的优势。

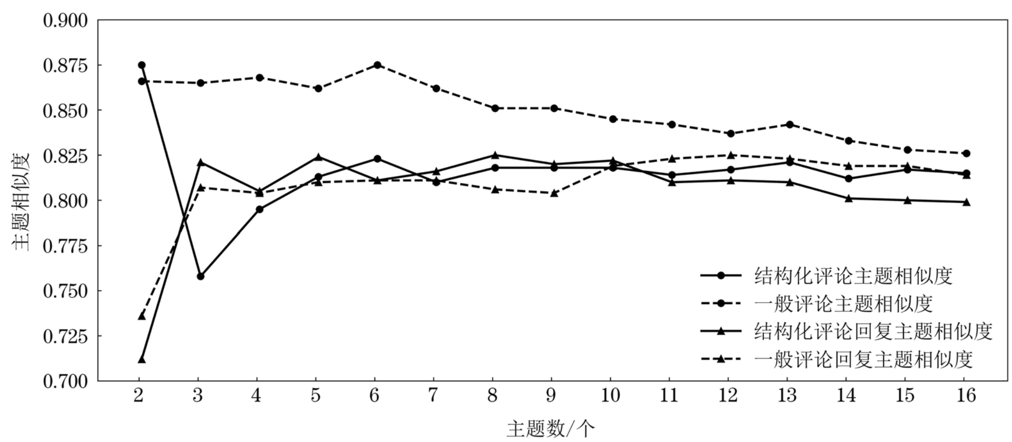

图5　评论文本主题相似度曲线

结构化评论回复与评论本身的主题相似度间有一定差值。这说明随着评论回复增加，结构化评论可以促进知识内容的发散，丰富知识交流的主题，这是因为回复中的新观点和讨论进一步拓展了原评论的讨论范围。而一般评论回复与评论本身的主题相似度差值更大，说明相较于结构化评论，一般评论本身对知识交流的影响较小，但可以通过评论回复的数量优势缩小差距，进而覆盖更多的话题和角度，丰富知识交流的主题。

4　结语

本研究通过统计分析、行为属性分析、文本相似度分析以及主题聚类分析的综合分析方法，对在线视频社区结构化评论和一般评论的知识交流效果进行测度，揭示了其中的规律。研究认为，结构化评论对于在线视频社区的知识交流发挥了重要作用。从在线视频社区中知识交流主体的角度出发，为充分发挥结构化评论的优势，本研究提出以下建议：首先，在线视频社区须完善结构化评论功能，优化结构化评论在评论区中的排序机制，进而提升用户的感知价值，促进用户发布结构化评论，扩大其对知识交流的影响力；其次，视频创作者须注重粉丝经营，积极打造专业形象，重视对知识质量的打磨，获得用户对视频内容价值的高度认可；最后，用户须重视结构化评论发布在知识交流中的积极作用，合理利用文本编辑功能和图片插入等特色功能，丰富评论内容表现形式。

本研究仍存在一些局限和不足之处：①主要以行为属性数据、视频文本数据和评论文本数据3类客观数据为数据来源，并未分析视频本身和结构化评论的图片部分，未来可结合两者进行深入分析；②评论挖掘方法以主题聚类为主，对于文本风格、情感表现等其他特征的分析相对欠缺，未来可以进一步提升内容分析深度，进而挖掘更深层次的知识交流规律。

参考文献

[1] 王春刚，张许歌，王俭. 在线健康社区与短视频融合的微知识共享模式研究[J]. 情报科学，2023，41（6）：79-83，93.

[2] 张瑞，胡丽娜，黄炜. 虚拟学术社区知识交流多智能体仿真演化研究[J]. 现代情报，2023，43（10）：16-26.

[3] KIM A J，JOHNSON K K P. Power of consumers using social media：examining the influences of brand-related user-generated content on Facebook[J]. Computers in Human Behavior，2016，58：98-108.

[4] BILIBILI POWER UP 2023 百大UP主盛典[EB/OL]. [2024-06-16]. https://www.bilibili.com/BPU2023#/stage2.

[5] 宓浩，刘迅，黄纯元. 图书馆学原理[M]. 上海：华东师范大学出版社，1988.

[6] SHAN W，QIAO T，ZHANG M L. Getting more resources for better performance：the effect of user-owned resources on the value of user-generated content[J]. Technological Forecasting and Social Change，2020，161：120318.

[7] 杨瑞仙，黄书瑞，于政杰. 基于三阶段DEA模型的学术虚拟社区知识交流效率评价研究[J]. 情报学报，2021，40（3）：256-266.

[8] JIN X L，YIN M J，ZHOU Z Y，et al. The differential effects of trusting beliefs on social media users' willingness to adopt and share health knowledge[J]. Information Processing & Management，2021，58（1）：102413.

[9] ULLAH A，ZHANG Q Y，AHMED M. The influence of intellectual property rights protection on contribution efforts of participants in online crowdsourcing contests[J]. Computers in Human Behavior，2021，123：106869.

[10] GUAN T，WANG L，JIN J H，et al. Knowledge contribution behavior in online Q & A communities：an empirical investigation[J]. Computers in Human Behavior，2018，81：137-147.

[11] 程豪杰，张宁，袁勤俭. 学术虚拟社区知识交流的研究进展[J]. 现代情报，2018，38（5）：158-163.

[12] 王瑞，李思豫，袁勤俭. 学术社交网络用户特征对知识交流效果的影响：以南京大学ResearchGate用户为例[J]. 图书情报知识，2020，37（4）：97-105，132.

[13] JIN J H，YAN X B，LI Y J，et al. How users adopt healthcare information：an empirical study of an online Q & A community[J]. International Journal of Medical Informatics，2016，86：91-103.

[14] 李月琳，王姗姗，阮妹. 跨源健康信息搜寻的动机、信息源选择及行为路径[J]. 情报学报，2021，40（1）：77-87.

[15] XIONG J，FENG X D，TANG Z W. Understanding user-to-user interaction on government microblogs：an exponential random graph model with the homophily and emotional effect[J]. Information Processing & Management，2020，57（4）：102229.

[16] 袁毅，杨莉. 问答社区用户生成资源行为及影响因素分析：以百度知道为例[J]. 图书情报工作，2017，61（22）：20-26.

[17] 袁勤俭，毛春蕾. 学术虚拟社区特征对知识交流效果影响的研究[J]. 现代情报，2021，41（6）：3-12.

[18] 邓胜利，管弦. 基于问答平台的用户健康信息获取意愿影响因素研究[J]. 情报科学，2016，34（11）：53-59.

[19] 王若佳，张璐，王继民. 基于扎根理论的在线问诊用户满意度影响因素研究[J]. 情报理论与实践，2019，42（10）：117-123.

[20] 周阳，谭春辉，朱宸良，等. 基于扎根理论的虚拟学术社区用户参与行为研究：以小木虫为例[J]. 情报科学，2022，40（1）：176-183.

[21] XU X，WANG X Q，LI Y B，et al. Business intelligence in online customer textual reviews：understanding consumer perceptions and influential factors[J]. International Journal of Information Management，2017，37（6）：673-683.

[22] 熊回香，向瀛泓，孟璇，等. 在线健康社区中基于用户评论挖掘的医生服务评价研究[J/OL]. 情报科学：1-19[2024-06-16]. https://kns.cnki.net/kcms/detail/22.1264.G2.20240126.1758.004.html.

[23] 叶艳，吴鹏，周知，等. 基于LDA-BiLSTM模型的在线医疗服务质量识别研究[J]. 情报理论与实践，2022，45（8）：178-183，168.

[24] RYAN R M，DECI E L. Brick by brick：the origins，development，and future of self-determination theory[J]. Advances in Motivation Science，2019，6：111-156.

[25] ZHANG Y，ZHANG M L，LUO N，et al. Understanding the formation mechanism of high-quality knowledge in social question and answer communities：a knowledge co-creation perspective[J]. International Journal of Information Management，2019，48：72-84.

[26] 王子喜，杜荣. 人际信任和自我效能对虚拟社区知识共享和参与水平的影响研究[J]. 情报理论与实践，2011，34（10）：71-74，92.

[27] SAURA J R，PALACIOS-MARQUÉS D，RIBEIRO-SORIANO D. Exploring the boundaries of open innovation：evidence from social media mining[J]. Technovation，2023，119：102447.

[28] ZHAO C Y，ZHU Q H. Effects of extrinsic and intrinsic motivation on participation in crowdsourcing contest：a perspective of self-determination theory[J]. Online Information Review，2014，38（7）：896-917.

[29] DONG L F，HUANG L Q，HOU J H，et al. Continuous content contribution in virtual community：the role of status-standing on motivational mechanisms[J]. Decision Support Systems，2020，132：113283.

[30] PU J C，CHEN Y，QIU L F，et al. Does identity disclosure help or hurt user content generation? social presence，inhibition，and displacement effects[J]. Information Systems Research，2020，31（2）：297-322.

[31] 陈小卉，胡平，马宇彤. 问答社区知识传播者网络声望、行动相似度与知识传播效果[J]. 管理学报，2021，18（5）：741-750.

[32] YANG M C，REN Y Q，ADOMAVICIUS G. Understanding user-generated content and customer engagement on Facebook business pages[J]. Information Systems Research，2019，30（3）：839-855.

[33] 韩玺，蒋佩瑶，韩文婷，等. 医生在线评价信息特征的影响因素研究：社会资本和社会交换理论的视角[J]. 信息资源管理学报，2023，13（1）：78-90.

[34] 沓钰淇，傅虹桥. 网络口碑对患者就医选择的影响：基于在线医生评论的实证研究[J]. 管理评论，2021，33（11）：185-198.

[35] 宋雪雁，王萍. 用户信息行为研究述评[J]. 情报科学，2010，28（4）：625-629，636.

[36] YE L L. User perceived value and continued participation in barrage video[C]//Proceedings of the International Conference on Education，Economics and Information Management（ICEEIM 2019）. Paris：Atlantis Press，2020：150-153.

[37] 黄菡，王晓光，何静，等. 基于矩阵相似度的主题演化路径判别研究[J]. 情报学报，2023，42（11）：1265-1275.

[38] 张瑞，何禄鑫，杨艳妮. 考虑语义距离的领域科学知识主题关联与演化研究[J]. 情报杂志，2022，41（10）：121-129.

[39] 黄佳佳，李鹏伟，彭敏，等. 基于深度学习的主题模型研究[J]. 计算机学报，2020，43（5）：827-855.

[40] BLEI D M，NG A Y，JORDAN M I. Latent Dirichlet allocation[J]. Journal of Machine Learning Research，2003，3（4/5）：993-1022.

[41] EGGER R，YU J. A topic modeling comparison between LDA，NMF，Top2Vec，and BERTopic to demystify twitter posts[J]. Frontiers in Sociology，2022，7：886498.

[42] GROOTENDORST M. BERTopic：neural topic modeling with a class-based TF-IDF procedure[EB/OL]. [2024-06-16]. https://arxiv.org/abs/2203.05794.

[43] MCINNES L，HEALY J，MELVILLE J. UMAP：uniform manifold approximation and projection for dimension reduction[EB/OL]. [2024-06-16]. https://arxiv.org/abs/1802.03426.

[44] MCINNES L，HEALY J，ASTELS S. HDBSCAN：hierarchical density based clustering[J]. The Journal of Open Source Software，2017，2（11）：205.

[45] 关于我们-哔哩哔哩[EB/OL]. [2024-06-16]. https://www.bilibili.com/blackboard/aboutUs. html.

作者简介

舒科慧，男，硕士研究生，研究方向：信息组织与信息检索，E-mail：rhodonite_shu@mails.ccnu.edu.cn。
过其桥，男，硕士研究生，研究方向：信息组织与信息检索。
田园，女，硕士研究生，研究方向：信息组织与信息检索。
程秀峰，男，博士，教授，硕士生导师，研究方向：信息组织与信息检索。
杨金庆，男，博士，特任副教授，研究方向：信息组织与信息检索。

Comparative Measurement of Knowledge Exchange Effects in Online Video Communities: Referring to the Structured Comment Function of Bilibili

SHU KeHui GUO QiQiao TIAN Yuan CHENG XiuFeng YANG JinQing
(School of Information Management, Central China Normal University, Wuhan 430079, P. R. China)

Abstract: From the perspective of knowledge exchange, this study focuses on the structured comment function within online video communities to assess the effectiveness of knowledge exchange, aiming to provide valuable insights for promoting high-quality knowledge exchange in these communities. By collecting data on videos, comments, and user behavior from a representative online video community, Bilibili, and employing methods such as topic clustering and correlation analysis, the research evaluates the impact of structured comments versus general comments in the knowledge exchange process. Furthermore, the study examines the interconnections between behavioral attribute data, video content data, and comment text data. The findings suggest that the quantity of structured comments is not strongly correlated with video playback numbers but is more influenced by the quality of video content and users' willingness to engage in knowledge exchange. Additionally, structured comments show a greater ability to enhance the depth of knowledge exchange, particularly for niche videos with a dedicated audience. Therefore, the structured comment function is demonstrated to have the potential to reinforce knowledge exchange in online video communities.

Keywords: Structured Comment; Online Video Community; Knowledge Exchange; Comment Mining; Bilibili

（责任编辑：管清潆）