

关于疫情的文本研究： 一个在线内容分析法的思考

张荣显，曹文鸳，刘源

摘要

新冠肺炎疫情期间产生了大量的文本资料，如新闻报道，社交媒体上的网络讨论、政策文本等等，为学术研究提供了丰富的数据素材，产生了不少研究论文成果。本研究集中盘点内容分析法在疫情研究的应用，尤其关注在线内容分析法的运用，使用文献分析，理解传统和在线内容分析法在方法论上的相关理论及操作要点异同。以“疫情+内容分析法”为主题词在 CNKI 数据库检索，结果发现，大部分的研究采用小样本数据，使用大数据研究的项目较少。在内容分析法的操作执行上欠缺规范，例如大部分论文缺乏对信度的汇报，对数据样本的采集方式说明不足等。部分研究采用计算机或大数据技术辅助（38%论文有汇报所采用的数据处理工具），其中的部分仅是使用某种软件来管理数据，采用技术来辅助数据挖掘或分析的研究相对较少，且形式较为单一，应用程度较浅。

内容分析法作为一种强调科学性和系统性的研究方法，经过长时间的发展和运用，已在社会科学中占有一席之地。随技术革新，大数据技术在疫情研究中应用具有一定优势，能够在不同方面改良文本数据的处理和挖掘，例如，采用科学的方法检索数据和清理数据，建构具有代表性和符合研究目标的样本库，利用网络挖掘和大数据技术快速挖掘初步的数据结果，观察数据概貌。从研究理论出发，建构测量指标，即是编码类目的设置依据。根据类目的不同，决定采用人工编码或算法编码，通过不同的质量控制体系（编码员信度，人机结果对比等），最终获得统计结果。大数据技术辅助在线内容分析能够帮助解决信息的覆盖度问题、数据的测量问题和海量信息的分析结果的解释度问题，解读数据意义，其意义挖掘的深度由浅层逐渐转向了深层。

关键词

疫情研究，内容分析法，大数据，研究方法

作者简介

张荣显，澳门互联网研究学会会长，电子邮件：angus@e-research-solutions.com

曹文鸳，澳门互联网研究学会，电子邮件：wennycao@boyidata.cn

刘源，澳门互联网研究学会，电子邮件：liuyuan@boyidata.cn

一、研究背景

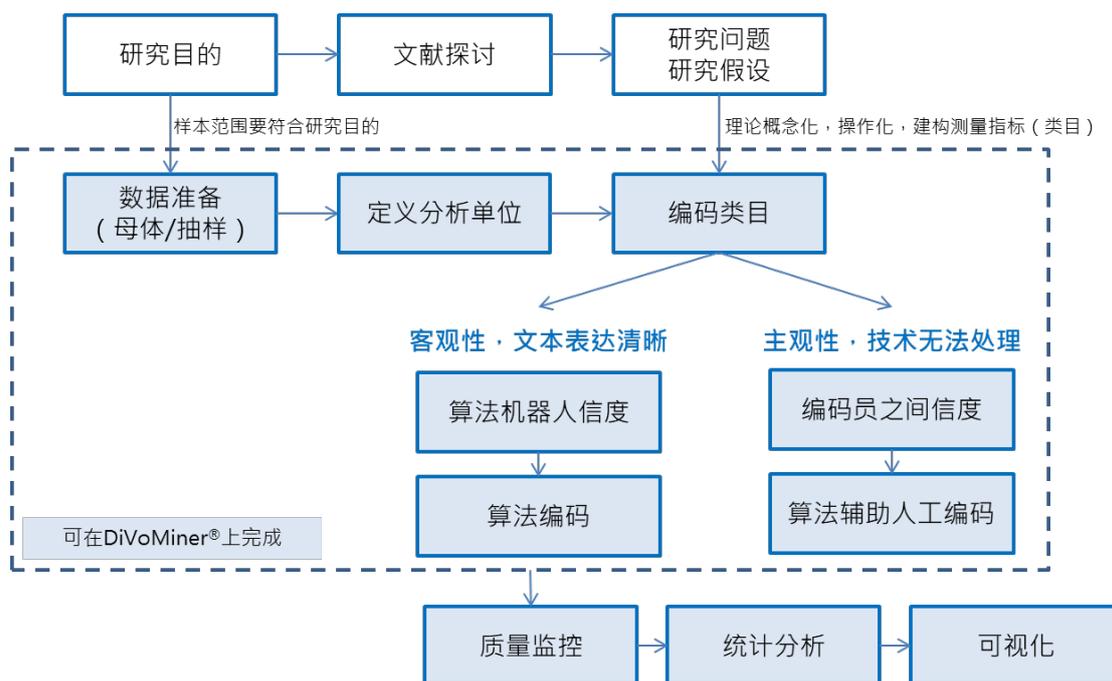
新冠肺炎疫情改变了生活，也给研究界带来了深远的影响，产生许多以疫情为主题的论文成果，尤其是疫情期间，产生了大量的文本资料，如新闻报道，社交媒体上的网络讨论、政策文本等等，有不少研究文本为研究对象，完成论文。随生活回归常态，也借此机会，盘点疫情研究。本研究集中关注以内容分析法为主要研究方法的疫情研究成果，尤其关注在线内容分析法的运用，使用文献分析，理解传统和在线内容分析法在方法论上的相关理论及操作要点异同。

二、传统及在线内容分析法

内容分析法作为一种研究方法，经过长时间的发展和运用，已在社会科学中占有一席之地。通过一整套具有科学性和系统性的程序，测量文本中某些变量的数值 (Berger, 1991)，并运用统计分析方法分析这些数值涉及的关系，从而得出文本中的意义 (Riffe, Lacy, & Fico, 2005)。使用关键词“内容分析法”在中国知网 CNKI 上检索，2020 年内（到 9 月 30 日）已发表 1,845 篇相关论文。内容分析法广泛运用在传播学及其他领域，主要用来描述内容以及检验理论 (Lacy et al, 2019)。

使用内容分析法为主要研究方法的疫情研究中，部分研究沿用传统的内容分析法，以人工处理为主，如芦何秋等研究者，以疫情期间出现的网络谣言为分析对象，从重大公共卫生事件突发期谣言的内容建构、谣言传播和辟谣应对等角度，理解谣言的传播机制和辟谣对策 (2020)。该研究采用完成的内容分析法流程，完整汇报了研究理论和分析框架、数据取样范围、内容编码过程（编码员、信度测试结果）和数据结论等；也有部分研究在传统的内容分析法基础之上，采用更新的方法论，使用计算机技术和大数据算法编码完成研究，例如程潇潇团队以抗疫背景下中医媒介形象之变化为研究主题，采用大数据技术辅助在线内容分析法完成研究 (2020)。本研究关注在传统内容分析法的理论基础之上，结合计算机技术和大数据算法，从社会科学的角度，梳理大数据技术辅助在线内容分析法的理论要点和操作流程。

表 1：大数据技术辅助在线内容分析法的主要操作流程



与所有研究的起点相同，在研究开始，需要明确研究目的，在文献探讨的基础之上，提出研究问题或研究假设。进入到数据内容分析的过程，首先需要准备数据，样本通常需要根据研究目的做设定，选取符合研究需求的数据样本。在这个环节，可以考虑使用母体样本，或是对数据进行抽样处理。在处理和数据分析之前，需要定义分析单位，即是考虑最小的独立分析的单元是如何。编码类目是内容分析的大框架，主要来自于理论的概念化和操作化，根据理论的具体化结果，建构测量指标。根据编码类目的属性不同，决定接下来的分析方法，由于技术的改良，对于大部分具有客观性并且文本表达相对清晰的编码类目，已经可以由大数据算法完成。但是依然有部分主观性较强，尤其是在观念上有多层次表达的编码类目上，当前的技术水平还暂时无法准确分析这部分的维度，因此需要更多的人工介入，由算法辅助完成人工编码。在整个编码分析的过程中，都需要对数据质量进行监控，如果是人工编码，需要考虑数据质量和编码员绩效，在算法编码中，需要考虑算法的准确性。最后统计分析并且产生可视化的数据结果，以供研究解读。接下来具体说明每个环节。

（一）研究目的

与各类型研究一样，研究之初，需要确定研究目的，这是后续做好数据分析的基础。使用内容分析法的研究，从概念化和数据分析的关系角度出发，不同的研究目的会导向不同的研究设计。研究目的可以是描述性研究，以描述内容变量结果为主，这种需求下的统计分析过程也可能相对简化，因为不涉及理论的验证；还有研究目的也可以是通过挖掘内容变量来

回答研究问题或研究假设 (Lacy, et al, 2019)。

疫情相关的论文中，新闻传播学领域研究目的较多为研究媒体主体（传统新闻、社交媒体）呈现出的传播特点，是否存在某种报道框架，是否存在舆论引导效果，以及传播内容中表达出的媒介形象等。以政策研究为主体的研究论文多关注政府部门管理、医疗系统等主体在应对疫情、及后续复工复产需求下的政策运用情况，研究目的决定了分析的方向和复杂性。

（二）理论到数据的落地

在以理论验证或回答研究问题的研究中，会涉及到由理论概念化，概念测量等要求。研究者从一个研究话题相关的观点出发到测量的过程，是一个概念化、操作化和测量的过程 (Babbie, 2007)。概念本身难以直接测量，这是因为在概念到文本数据之间存在对应理解落差，因此在文本分析的研究中，通常会有一个将概念具体化的步骤，即是将概念落实到维度与测量指标。从理论过渡到数据的解读，落实到操作层面，则是以编码类目为桥梁。

1. 拟定研究问题或提出研究假设

提出研究问题，或者研究假设，是在研究目的之下明确研究的方向，并且以此作为数据研究的指向。

2. 指标及类目

提出研究问题后，很多时候难以直接用于文本分析，即是在测量层面上有难度。因此，需要给涉及的每个概念设定维度和指标，其中指标层级就是概念的具体化结果，可用于制作内容分析的编码类目，提供给编码员使用的编码簿。程潇潇团队在回顾相关文献后，并考虑到数据采集、清理和分析等方面可能遇到的覆盖度 (coverage)、测量 (measurement) 和解释 (interpretation) 难题，进一步将研究问题拆分成为不同的维度，并且需要为每个维度设立了测量指标。

（三）数据准备

内容分析法中的研究对象，可以是任意一种可被传播的信息，包括“单词、意义、图像、符号、思想、主题等”，包括书面的、视觉上的或是口头表达的资料均可用于内容分析 (Neuman, 1997)，例如书籍、章节、采访内容、讨论内容、报纸标题和文章、历史资料、演讲、谈话、广告、戏剧、非正式交谈或者任何交流性的语言 (Palmquist, 2013)。在社会科学领域，可以用来做内容分析研究的资料包含新闻报道、社交媒体内容、文学作品、历史档案、访谈、学术文献、政策文本、发言稿、图片和视频等。到了大数据时代，网络数据大而全，同样需要考虑数据齐全、具有代表性，保证数据质量等问题，否则数据样本失焦，难以满足研究要求。

1. 根据研究主题确定数据范围

研究者在确定研究主题（研究假设）后，决定需要使用什么范畴的数据，可以回答研究问题。

2. 架构概念化逻辑，检索获取数据样本

在研究实践中，检索条件的设计优劣，会影响到数据样本结果的数据量和准确度。一种很容易发生的失误是，检索逻辑（关键词条件）范围过大导致数据目标模糊，样本代表性不佳，混入无关和垃圾信息过多，给研究造成困扰。另一种情况是，检索条件方向偏离，导致查询结果与目标数据范围落差过大，进一步分析时会得不到有效数据结果。为避免上述问题，可使用概念化逻辑检索的思路，利用多元检索关键词搭配，配置一定的逻辑语言，建立一套检索概念（有点像建立检索概念数据库），利用检索词在概念上的相关性，检索获得同属一类概念的结果。

3. 清理数据

由语言表达的复杂性和网络信息的杂乱无章，即便使用概念化逻辑条件检索文本数据，依然无法保证数据恰如其分刚好是研究范畴内的数据。所以获取初步数据样本之后，需要清理数据。在实践中，清理数据的方法也有多种做法。

a) 使用关键词排除

例如，在某论文中，想要收集关于医疗人员形象的内容，设置检索数据条件后，检索到的数据结果中，存在某些医疗机构公关文，但是这明显不是研究范畴内的数据。对于这种情况，设置招聘概念类检索条件，把提及“xx 医疗”等内容的文本排除出去，实现数据清理。

b) 使用条件排除

可以对文本长度（字数）、提及关键词次数（相关性）等因素进行考虑，设置排除条件。有论文为了提高数据相关性，提升目标数据覆盖度，于是做了初步预处理，移除信息量过少的文本样本，删除文本少于 200 字的样本。随后研究人员人工检查样本，发现虽然很多文本确实提及了中西医相关词汇，但是报道主题并非与医学直接相关，而且部分文本中仅出现一两次关键词。因此团队再次进行数据清理，剔除单篇文本中提及检索关键词词频少于 5 次的文本。

c) 人工清理

即便采用结合排除词和条件设定方法，进行了属于预处理，然而依然可能有无关数据混入在内，有必要的话，最终也可以补充人工操作。具体方法是研究人员手动删除无关数据。

（三）定义分析单位

分析单位是研究过程中被分析的主体，主要是“什么”或“谁”被分析。就像化学和物理学一样，其中所有物质都可以基于原子或分子单位（忽略更深层次的亚原子粒子物理学），任何其他科学或其他科学理想的领域应该是以元素的不可约的基本组合为基础（Ernest, 2016）。内容分析将分析单位定义为一段明确的内容，特点是将其归入一已设定的类目中（Holsti, 1969）。亦即是说，当编码员针对一段传播内容编码进行一个类目编码时，这一段传播内容就是一个记录单位。记录单位是搜集抽样单位中的资料，提供分析的基础。

（四）信度测试

根据编码类目的属性不同，决定了后续执行内容编码的方式有两大类，大数据技术的算法编码及算法辅助人工编码，侧重点各自是算法和人工。由于执行编码的主题不同，在信度测试方面也有不同的要求。人工编码要求进行编码员之间的信度，在信度达到接受的一致性后，再执行正式编码；而算法编码则要求检验人工和算法之间的准确性。

1. 编码员之间信度

传统内容分析法要求在正式编码之前，需要进行编码员之间信度测试。信度是指在研究过程中，测量数据所独立于测量工具的程度。即不同的研究者对同一现象进行重复测试后，所产生结果的一致程度（Babbie, 2007）。也就是说，如果测量过程要进行两次或多次，所得出的结论应该是类似的，这确保了内容分析的有效性或可靠性。信度的测试可以在科学抽样的基础上进一步防范研究者对数据产生的影响，编码员之间信度可保证对数据较为一致的处理，使研究结果更为客观。

可以利用必要的工具来计算信度指标，如手动（利用公式计算）或计算机程序。我们以内容分析法中进行编码员之间信度评估为例，具体的操作步骤如下：

a) 依据编码簿制作编码指引，编码指引需明确统一。帮助编码员熟悉议题，理解编码类目，保证所有编码员均理解类目所指代的含义。

b) 进行编码测试。选取少量样本作为测试编码之用的样本，这部分样本的测试编码过程中，各编码员需独立编码，不能相互讨论或指导。若使用了机器编码，则直接执行机器编码的信度测试。

c) 进行编码校正。测试编码结果如果未能达到理想信度，需重新测试编码。重新测试之前，对编码员再次进行培训和指导，尤其针对编码结果差异较大之类目，需再次给予说明。若有机器编码的部分，则需重新检查并修正编码簿各选项的关键词，尽可能完善类目的选项，再重新执行机器编码信度测试。

d) 正式编码。当所有编码员达至理想信度，即可开始正式编码。

图 1：编码员之间信度测试的操作流程



2. 算法机器人的信度要求

算法编码是指采用算法或计算机挖掘数据内部规则，来得出数据编码结果的方式 (Zamith & Lewis, 2015)。在大数据技术辅助内容分析法中，应用了算法编码，信度测试就不仅仅是编码员和编码员之间进行计算，需要对算法机器人的信度进行评估。算法编码可以由研究人员设定编码规则，让计算机代为执行编码过程，优点是极为高效，可以在很短时间内完成数据分析，信度百分之百。但是编码效度则是依赖于研究人员对编码规则的定义水平 (Lacy et al, 2019)。在操作实践中，有研究者试图采用人机比对的方法，检验算法编码是否可以达到可接受水平。程潇潇的研究团队随机抽取 1% 的样本作为检验样本（107 篇），执行人工编码，对照大数据算法自动编码结果，计算信度，所有类目准确度均在 0.87 以上，显示大数据自动编码结果良好，可采用该结果作为解读依据。这一步骤是为确保大数据算法自动编码的准确程度。

我们在其他的研究中也发现有类似做法，例如，张文瑜教授（2020）在一项健康传播的研究中，同样随机抽取了 1% 的数据作为比对样本，由 4 位经过训练的研究助理完成人工编码的部分，且编码员之间信度达到 Cohen's Kappa = 0.78 ($p < 0.001$)，95% CI (0.604, 0.948)。最终人机对比一致性达到 80%，证明算法编码结果可接受 (Chang, A. Schulz, P. J. & Cheong, A. 2020)。

（五）内容编码

编码是将内容分析的文本，按照类目进行归类的过程。人工编码一般按照编码说明检阅文本，填答编码簿，即完成了一次人工编码，所有文本的编码结果最终汇总到一起进行统计分析，即完成了内容分析的量化过程。传统内容分析法采用人工内容编码的方式，相对效率较低。编码员可能存在个体认知差异，因此在正式编码之前，需要进行编码员之间的信度测试。随计算机和大数据技术发展，内容分析法逐渐引入了技术和算法支持，优化内容分析操作流程，并且给研究方法本身带来了更多的可能性。目前阶段，内容编码可以区分为人工编码和算法编码，适用于不同的应用场景。根据编码类目的属性，客观性较强、可以使用明确文字表达的类目建议使用算法编码，效率高，准确性高，例如文本主题、提及人物等。主观性较强、无法界定准确的语义表达的类目建议采用人工编码，例如多层次的意见态度、有指向性的情感分析等情况。当然技术在人工编码的场景下也可以提供一些协助，比如给出某些关键词提示等。

1. 大数据算法编码

对于文本大数据研究来说，借助计算机算法辅助进行算法编码，在人工少量介入的情况下，可提升编码效率同时确保编码质量。但是算法编码由算法机器人执行，指引机器对文本进行关键词的自动化标注和自动化填答选项，以完成自动编码，准确性依然依赖于人工介入的部分，尤其是人工对算法的设定规则，决定了数据的质量。最终由通过信度测试的编码员对大数据算法编码结果进行随机检查和校正，进一步提升了编码准确度，完成数据结果的收集工作。

例如在编码类目“中西医类别”这一题中，选项区分了中医、西医、中西医结合三个选项，接下来由研究人员对每个选项的编码规则设置条件，例如，文本中提及关键词“中医 OR 汉医”等，则自动判定为“中医”选项。

这是相对较为简单的算法编码，在算法的编辑上，可以考虑更丰富的场景和需求，或采用机器学习的方法执行算法编码，那么需要有人工编码的可靠训练集作为学习基础，暂不展开说明。

算法辅助人工编码以编码员为主导。通过编码员之间的信度测试后，正式进行内容编码。计算机技术和大数据算法可以提供某种程度的辅助，例如，类目选项中，同样可以设置某些关键词，文本中会自动化标注关键词，预设选项，协助人工编码。

（六）统计分析，可视化解读结果

选择何种统计方法未见和研究主题有直接关系。最常用的数据描述方法，单变量频次的

呈现和多变量交叉分析、时间序列表达，在各类研究主题中均有运用。在可视化表达方面，图表的应用同样丰富，图表应用目的以比较、占比、相关和趋势四大类为主，也可以借助LDA模型进行主题分析、社交网络或语义网络分析等。

表 2：统计分析方法说明

比较	比较各数值之间的差异						
占比	部分占总体的百分比						
相关	显示各数值的关系						
趋势	各数值随维度变化的情况						

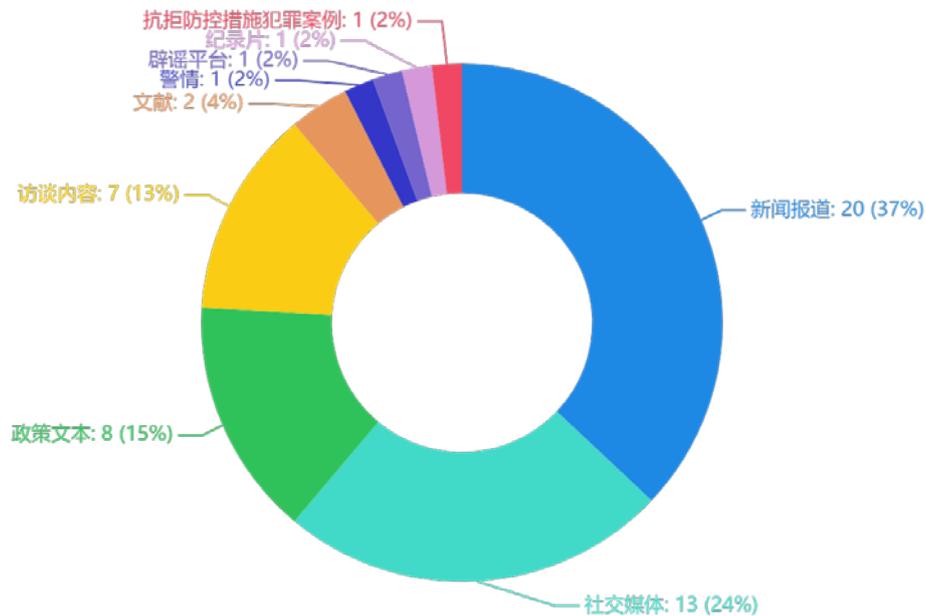
三、基于内容分析法的疫情研究文献分析结果

在中国知网上使用“疫情+内容分析法”为检索关键词，至2020年9月20日共发表49篇学术期刊论文，集中在社会科学领域，涉及新闻传播、公共卫生政策、教育、社会心态等多方面主题。说明内容分析法在具体学科及研究领域的运用较为广泛，使用场景多元。

（一）文本来源类型

在文本数据的来源和类型方面，以新闻报道和社交媒体内容最多，合计占六成（61%，33篇），政策文本和访谈内容（含调查问卷开放题）分别占15%（8篇）和13%（7篇），也是较多的类型。绝大多数分析文本以文字为主，也有部分视频等数据形式，例如，有4篇论文以短视频（来自抖音等网站）为研究对象，体现出新媒体研究的特点。

图 2：文本来源类型



以分析新闻报道为主的研究论文多是讨论媒体在疫情中信息传播策略和媒介形象等方面主题，如《新闻叙事写作策略研究——以《人物》在疫情期间的报道为例》（姜文洁，2020）、《官方与非官方媒体疫情报道框架的对比研究——以新华社和三联生活周刊为例》（沈雨柔，2020）等，以媒体报道策略和报道框架为研究核心；社交媒体研究较多关注微信、微博等平台内容在新冠疫情这类公共突发事件中的传播特征和作用，如《社交平台中主流媒体的公共突发事件报道创新研究——以《人民日报》官方微博对新冠肺炎报道为例》（邵慧 & 李怡菲，2020）、《新冠疫情中政务微信的信息发布——以“珠海发布”为例》（彭兰，2020）等，研究传统新闻媒体、政府信息公开在新媒体上的效果和作用。

政策文本方面，对疫情的研究涉及卫生、经济等领域，例如疫情期间公共卫生政策、医疗机构应对疫情政策、复工复产等经济政策等，如《我国新型冠状病毒肺炎疫情防控政策文本计量分析》（张文丽等，2020）、《基于政策工具的我国新冠肺炎疫情背景下企业复工复产政策文本分析》（赵雪芹, 李天娥 & 莫长镭，2020）。访谈内容/问卷开放题类型，对普通民众及医疗人员为访谈/调研对象较多，涉及社会心态、心理学等范畴较多。

（二）数据样本数量

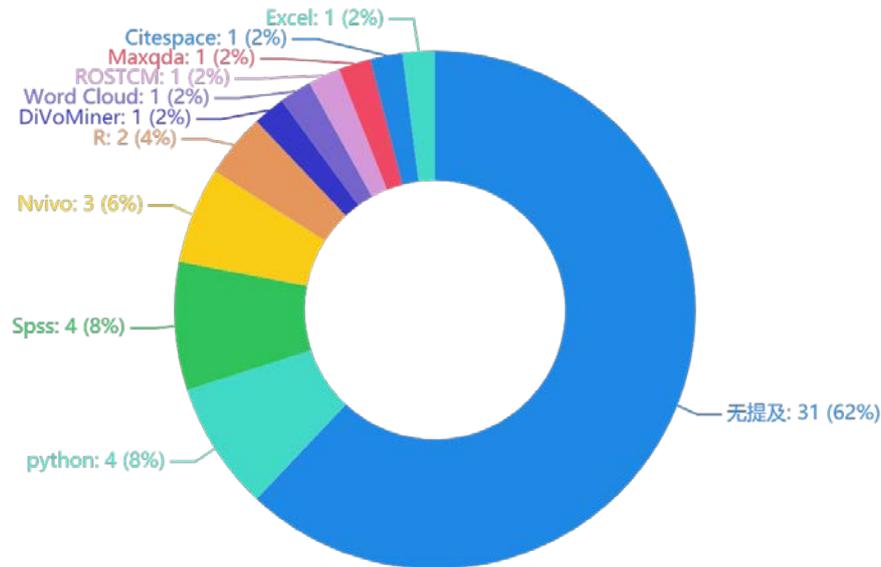
本次研究检视的疫情研究文献数据中样本量多寡不一，大部分的样本数量在一千条以内，有 38 篇，占有所有样本的 78%，样本数在一千条或以上有 7 篇，占 14%。说明相关疫情的内容分析法研究，以小样本为主，用到较大数据体量的研究相对较少。分析文本数量最多的研究

是程潇潇团队，最终采用样本量 9,981 条。

（三）文本处理工具

在使用的数据分析工具方面，62%论文未汇报具体数据分析工具。部分论文使用 python、Spss、Nvivo、R 等工具完成数据处理部分。

图 3：文本处理工具



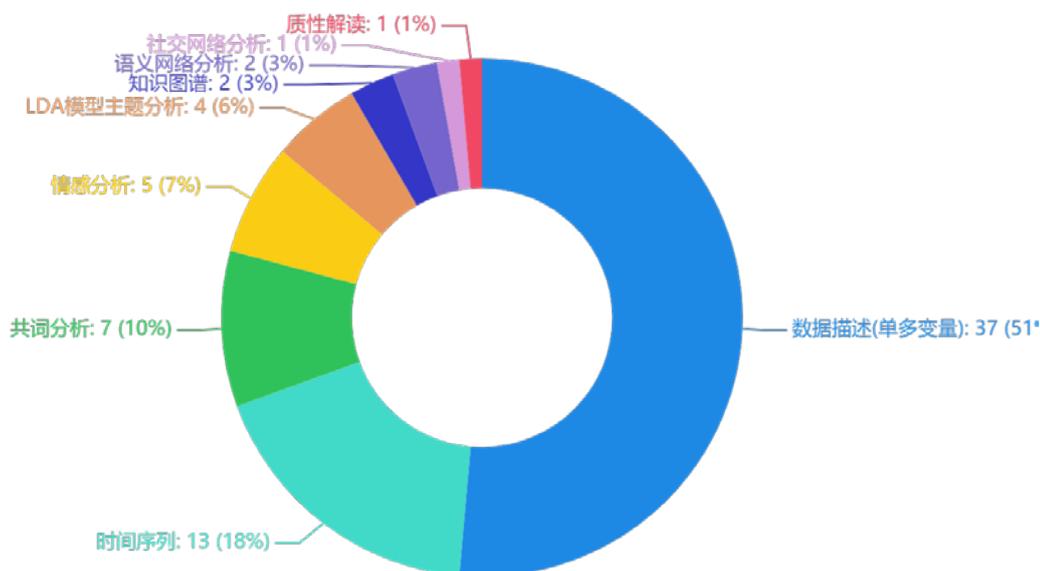
（四）信度

在对疫情研究的文献分析中，仅有 3 篇论文汇报了信度测试结果，占比相当低（1%），反应了当前研究普遍对信度的重视度不足。在 3 篇论文中，2 篇采用人工编码，汇报编码员之间信度测试结果，另 1 篇结合人工编码及算法编码，同时汇报有编码员之间信度测试和人工与算法之间的准确性结果。

（五）数据统计分析方法

在宣称使用内容分析法的论文中，采用的数据分析方法、数据挖掘方法多样，对数据描述（包括单变量频次及多变量交叉）占 51%（7 篇），接下来是情感分析和 LDA 模型主题分析，分别占 7%（5 篇）和 6%（4 篇）。

图 4：数据统计分析方法



从研究方法层面理解上述疫情研究文献情况，可见内容分析法作为经典研究方法，内容分析法在计算机技术和大数据算法的发展背景下，在方法论理念和操作执行层面也有创新，体现在对数据类型的兼容性更高，借助大数据算法执行数据编码和统计分析等。有研究者认为，当前计算机技术已经在内容分析法各项流程中广为运用，例如数据库的建构，基于关键词获取大数据集、内容编码和分析等环节。但是本质上，内容分析法依然是以“人”为中心的研究方法，即便是自动化分析，也需要由研究者设计分析规则，把关数据质量(Lacy et al, 2019)。即便是体现在人工智能、机器学习编码方法上，也需要由人工编码的结果作为训练集，离不开“人”的卷入。

四、结语

本研究借助疫情研究的论文成果，梳理以内容分析法为核心的研究在理论和操作上的做法，观察内容分析法给方法论和流程上带来的技术优化，但是当前的论文文献分析结果也反映出，大部分的研究采用小样本作为数据研究，使用大数据研究的项目较少。在内容分析法的操作执行上欠缺规范，例如大部分论文缺乏对信度的汇报，对数据样本的采集方式说明不足等。部分研究会采用计算机或大数据技术辅助（38%论文有汇报所采用的数据处理工具），其中的部分仅是使用某种软件来管理数据，采用技术来辅助数据挖掘或分析的研究相对较少，且形式较为单一，应用程度较浅。

对于大数据技术辅助在线内容分析法而言,以内容分析法为核心方法论,在文本数据的处理和挖掘上做出改良,例如,采用科学的方法检索数据和清理数据,建构具有代表性和符合研究目标的样本库,可以利用网络挖掘和大数据技术快速挖掘初步的数据结果,观察数据概貌。接着从研究理论出发,建构测量指标,即是编码类目的设置依据。根据类目的不同,决定采用人工编码或算法编码,通过不同的质量控制体系(编码员信度,人机结果对比等),最终获得统计结果,解读数据意义。

参考文献

- 程潇潇,金兼斌,张荣显 & 赵莹. (2020). 抗疫背景下中医媒介形象之变化. 西安交通大学学报(社会科学版)(04), 61-70. doi:10.15896/j.xjtuskxb.202004007.
- 姜文洁. (2020). 新闻叙事写作策略研究——以《人物》在疫情期间的报道为例. 声屏世界(07), 25-26. doi:.
- 彭兰. (2020). 新冠疫情中政务微信的信息发布——以“珠海发布”为例. 新闻前哨(04), 17-19. doi:.
- 邵慧 & 李怡菲. (2020). 社交平台中主流媒体的公共突发事件报道创新研究——以《人民日报》官方微博对新冠肺炎报道为例. 现代视听(02), 16-19. doi:.
- 沈雨柔. (2020). 官方与非官方媒体疫情报道框架的对比研究——以新华社和三联生活周刊为例. 新闻论坛(04), 83-86. doi:10.19425/j.cnki.cn15-1019/g2.2020.04.025.
- 赵雪芹,李天娥 & 莫长镛. (2020). 基于政策工具的我国新冠肺炎疫情背景下企业复工复产政策文本分析. 情报理论与实践(08), 21-28. doi:10.16353/j.cnki.1000-7490.2020.08.004.
- 张文丽,芮天奇,徐娟,张玲玲,张治国,黄文婧,沈丽宁. 我国新型冠状病毒肺炎疫情防控政策文本计量分析[J]. 医学与社会, 2020, 33(08):54-60.
- Babbie, E. (2007). 社会研究方法= *The Practice of Social Research: 第11版*: 英文. 清华大学出版社.
- Berger, A. (1991). *Media research techniques*. Newbury Park, CA: Sage, 25.
- Chang, A. Schulz, P. J. & Cheong, A. (2020). Online Newspaper Framing of Non-communicable Diseases: Comparison of Mainland China, Taiwan, Hong Kong and Macao. *International Journal of Environmental Research and Public Health*. (SSCI/ Scopus, Impact factor=2.849, 5-year impact factor = 3.127)

- Ernest, P. (2016). The unit of analysis in mathematics education: bridging the political-technical divide?. *Educational Studies In Mathematics*, 92(1), 37-58. doi:10.1007/s10649-016-9689-4
- Holsti, O. R. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley Publishing Company.
- Lacy, Stephen, et al. (2019). *Analyzing Media Messages : Using Quantitative Content Analysis in Research*, Taylor & Francis Group, ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/umac/detail.action?docID=5732528>.
Created from umac on 2020-09-23 23:46:19.
- Neuman, W. (1997). *Social research methods: qualitative and quantitative approaches*. Needham, Heights, MA: Allyn & Bacon, 272-273.
- Palmquist, M. (2013). Content analysis. Retrieved from <https://www.ischool.utexas.edu/courses>
- Riffe, D., Lacy, S., & Fico, F.G. (2005). *Analyzing Media Messages: Using Quantitative Content Analysis in Research (2nd ed.)*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers, 20.中文译本参见:[美]赖斯, 菲克等著, 嵇美云译《内容分析法: 媒介信息量化研究技巧》, 清华大学出版社 2010 年版。 .
- Zamith, R., & Lewis, S. C. (2015), Content Analysis and the Algorithmic Coder: What Computational Social Science Means for Traditional Modes of Media Analysis. *Annals of the American Academy of Political & Social Science*, 659(1), 307-318.