

信息系统领域的主题变化研究

——基于对 MISQ 的文献分析

任明¹，张瑾²，马宝君³

(¹中国人民大学信息资源管理学院，北京，100872；²中国人民大学商学院，北京，100872；

³北京邮电大学经济管理学院，北京，100876；)

摘要 信息系统研究领域在过去的几十年中不断发展，不断衍生出新的概念和主题。分析这些发展变化，有助于理解学科的核心问题，把握学科发展动向，为本领域的研究人员提供参考和借鉴。本文基于对 MIS Quarterly 最近30年来的文献的分析，总结了信息系统领域研究主题的变化，抽取了近五年来的新兴主题；进一步地，把新兴主题归为四类，通过分别提取对应文章的摘要中的高频词，总结了各类研究的关注热点。

关键词 信息系统；主题；MIS Quarterly；内容分析

1. 背景

信息技术和信息系统应用作为全球经济和社会发展的重要推动力量而得到越来越广泛的关注。在这个新兴且日新月异的研究领域中，研究问题和研究方向不断涌现和变化着，研究的层面也各有不同，例如从个人、团队、组织、市场的层面研究信息技术和信息系统与之的相互作用[8]。这些多元化的研究兴趣使得人们对复杂的现象不断获得更加丰富的认识，从而推动科学研究的发展[5][7]。回顾过去的几十年，人们对于信息系统的理解和关注点的确在不断推陈出新。

审视信息系统领域的发展和趋势也是一个有意义的研究问题。它旨在梳理一个研究领域的知识体系和知识脉络，剖析研究领域和学科的核心问题与发展趋势，同时也可以为研究人员提供参考和借鉴。在信息系统研究的发展过程中，也有学者对一定时期的演变进行分析。例如，Alavi and Carlson 对 8 个 MIS 相关期刊 20 年（1968-1988）的文章进行了抽取和分析[1]，Banker and Kauffman 根据 Management Science 创刊 50 年以来的文章抽取了 5 大类研究内容[3]，Sidorova 等使用 3 个期刊 22 年间（1985-2006）的论文总结了信息系统领域的知识核心[8]。此外，还有一些研究关注更为具体的某个子

领域，例如 Lee 等分析了电子商务研究 10 年（1996-2005）间的多元性[5]，Zhang and Scialdone 则选取 ICIS2009 的文章分析了当时有关 IT 构件 (artifact) 的研究关注点[9]。尽管文献范围有所差异，年代和跨度也各不相同，但分析维度是比较一致的。研究人员比较关注的文献特征包括研究的主题 (topic)、所属的层面 (analysis of unit)、及文献所采用的研究方法 (methodology)，其它还有采用的理论 (theory)、作者所属机构 (affiliation)、相关的学科 (related discipline) 等。上述研究大多以文献全文为研究对象，也有些使用了关键词、摘要等。在分析方法上，大都是搭建分类框架或者采用前人的分类框架，再对目标文献的全文、关键字等进行人工解读和编码，构建文献的特征向量，用于后续的分析和对比。Sidorova 等在分析过程中采用潜在语义分析 (LSA) 模型[4]，把文献的摘要和词映射到同一个低维的主题空间，进而发现研究主题[8]。

本文把目光聚焦在近 30 年来的信息系统研究，通过对其间研究文章的分析，总结信息系统研究主题的发展变化及新兴的主题。具体的，本文的文献分析聚焦在研究主题的维度上，希望回答下列问题：近年来信息系统领域的研究问题随着时间发生了怎样的更替？出现了哪些新的主题？这些主题又有着怎样的关注点？

本文的研究是基于对 MIS Quarterly 期刊发表文献的分析。在信息系统领域, MIS Quarterly 一直是公认的顶级“纯 MIS”期刊之一[1][8], 发表在其上的成果在很大程度上反映了信息系统领域研究的主流工作和最高水平。鉴于此, 本文以 MIS Quarterly 为例, 分析了发表在其中的 1983 年 1 月至 2013 年 6 月内的文章信息, 对上述问题进行了回答。在分析方法上, 本文并不完全依赖于人工编码, 而是采用内容分析的方法[2], 并在个别情况下进行必要的人工校正。

2. 研究方法

本文选取了文章的主题词和摘要进行分析。主题词是在标引和检索中用以表达文章主题的词或词组, 是规范化的、有组织、已定义的名词术语。与关键词相比, 主题词是经过规范化处理的术语, 它对文献中出现的同义词、近义词、多义词和同一概念不同书写形式等进行严格的控制和规范, 保证主题词的含义明确, 因此是本文分析的参照依据。摘要以相对简洁的语言概括了一份研究的相关信息, 如背景、内容、方法和贡献。与关键字、主题词等相比, 摘要能够传达更加丰富的信息; 与全文相比, 摘要又简明扼要, 处理的复杂度大大降低。

本研究所分析的文章信息是在 EBSCO 数据库中采集的, 包括文章的题目、主题词、摘要、卷号。在 MIS Quarterly 中选择 1983 年 1 月至今共计 30 年半的文章, 共返回 1395 条结果。通过删除缺少主题词和摘要信息、摘要极其简短、特殊类型的文章信息 (比如 Editor's Comments、Editor's Notes、Executive Overview、Introduction of Special Issue), 本研究筛选出 984 篇文章的信息用于后续分析。这些文章的类型绝大多数是 article, 也有个别的 product review 和 case study。

为了更好地分析主题的变化, 全部数据被切分成了 3 个部分, 分别对应三个阶段: 阶段 1 (1983.01- 1992.12)、阶段 2 (1993.01- 2002.12)、和阶段 3 (2003.01- 2013.06)。在分析最新出现的主题时, 又把最近的十年半切分成阶段 3a (2003.01- 2007.12) 和阶段 3b (2008.01- 2013.06), 在最近五年半 (即

阶段 3b) 的分析时又着重观察了 2013 年的情况。

下面说明分析依据的指标。一个主题词的频度越高, 说明这一主题的相关研究越被关注, 所以在指定阶段 i 的频度 F_i 是下面分析的主要指标之一。通过观察主题词的频度在各个阶段的分布可以总结出该主题的发展趋势。因此, 本研究设定的另一个指标是在指定阶段的比重 R , R_j 表示在阶段 j 的频度占总频度的比重。同理, 这些指标也适用于从摘要中抽取出的词。

每篇文章提供了个数不等的主题词, 主题词的处理主要是对每篇文章的主题词集合进行了主题词的分隔和提取, 用于后续分析。摘要的分析采用了内容分析的方法。首先使用 R 语言中的 text mining 包, 对摘要部分进行了预处理, 包括去除标点、去除停用词以及词干化, 然后再进行词频统计, 用于统计分析。

3. 研究发现

对所选文章的主题词进行初步分析, 共涉及 2618 个主题词, 这些主题词平均出现了 4.7 次, 也就是说平均出现在 4.7 篇文章中。在所有的主题词中, 频度最高的是 INFORMATION resources management, INFORMATION technology, MANAGEMENT information systems, 均高于 300, 这与它们是高层 (high-level) 的、通用的主题词有关。总频度在 100 以上的还有: RESEARCH, INFORMATION resources, STRATEGIC planning, DECISION making, INFORMATION storage & retrieval systems, KNOWLEDGE management。

3.1. 主题词的不同发展趋势

总频度不小于 10 次的主题词共计 247 个。分析这些主题词的 (R_1, R_2, R_3), 可归纳出四类发展趋势: 衰退型、屋顶型、平稳型、上升型, 从图 1 的 a, b, c, d 中分别可以看出这几类的发展趋势。

衰退型的特征是, 在第一个 10 年中曾经备受关注, 但后来相关研究迅速减少, 难以与往日的数量相比。表 1 给出了个代表性的衰退型主题词 ($F = F_1 + F_2 + F_3 \geq 30, R_3 < 15\%$)。

可以看出,由于技术已经成熟和应用热点的转移,电子数据处理、数据库管理、决策支持系统、专家系统等已经较少被提及,系统的分析、设计和开发也不再被大家所关注。

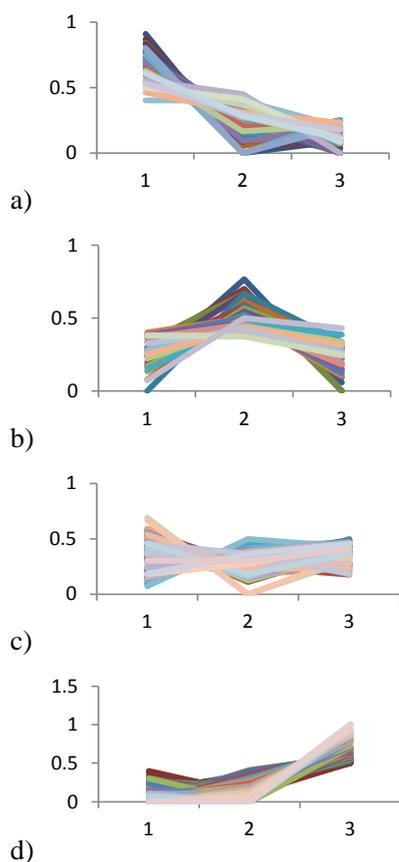


图 1.主题的不同发展趋势

表 1 衰退型的主题词

Computer systems
Computer Systems Design Services*
Cost effectiveness
Decision support systems
Expert systems (Computer science)
Electronic data processing
- End-user computing
- Database management
Industrial management
System analysis
- Systems design
Systems development

* 该词已经不出现在最新的主题词表中

屋顶型也体现了主题的衰退,不同的是,热点时期出现在中间十年,而不是第一个十年中。在表 2 中,可以看到屋顶型的主题词的代表 ($F \geq 10$, $R_2 > 60\%$), 重构、计算机培训、创新能力、服务质量这些曾经在十年前的热门主题,已随着社会和市场环境的变化

淡出我们的视线。

表 2 屋顶型的主题词

REENGINEERING (Management)
Vocational Rehabilitation Services
Computer Training
OCCUPATIONAL training
CREATIVE ability*
QUALITY of service

* 该词已经不出现在最新的主题词表中,以~ in technology, ~ in business 替代

平稳型的主题词在三个十年中的波动不如其它类型明显,基本上保持平稳的态势。这一类除了高层的、较为通用的主题词(如信息资源管理、信息技术、信息科学),还涉及组织变革、竞争优势、最终用户、信息服务等主题,它们在近些年得到较为持续的关注。具体请参见表 3 ($F \geq 35$, $\text{Max}(R_1, R_2, R_3) - \text{Min}(R_1, R_2, R_3) \leq 30\%$)。在图 1 的 c) 中最下方两条线索对应的主题词是 Custom Computer Programming Services 和 COMPUTER programming。虽然它们在中间十年从未被提及,但在最近十年又结合了新的背景(如开源软件、极限编程等)出现在信息系统领域中。

表 3 平稳型的主题词

INFORMATION resources management
INFORMATION resources
INFORMATION technology
STRATEGIC planning
DECISION making
INFORMATION science
BUSINESS planning
ORGANIZATIONAL structure
ORGANIZATIONAL change
COMPETITIVE advantage
PROBLEM solving
COMPUTER software development
END users (Information technology)
ECONOMIC aspects
INFORMATION services
TEAMS in the workplace

上升型的主题词的特征表现为近十年的相关研究数量处于高位,代表着获得研究人员关注的新问题。代表性的上升型主题词有:社会信息学、网络出版和传播以及搜索引擎、结构方程建模、外包、电子购物等。具体请参见表 4 ($F > 20$, $R_3 > 70\%$)。这些新

问题、以及新背景下的原问题(如顾客行为)快速吸引了研究人员的注意力。

表 4 上升型的主题词

SOCIAL informatics
Internet Publishing and Broadcasting and Web Search Portals
CONTRACTING out
STRUCTURAL equation modeling
INTERNET
ELECTRONIC commerce
Electronic Shopping
Business to Business Electronic Markets
SCHOLARLY publishing
CONSUMER behavior

3.2. 新兴的主题

在上升型的主题词中,有一些是几乎近五年才登上舞台的主题词。由于出现时间较短,可能无法进入 3.1 小节的分析(要求 $F \geq 10$)。既然它们代表着被关注的最新问题,这一小节就集中在这些新兴的主题词上。

表 5 给出了 21 个主题词 ($F > 5$ 、 $R_{3b} > 0.90$)。绝大多数的 $R_{3b} = 1$, 仅有 3 个在 0.90 左右,说明它们是新颖的。通过观察这些主题词的共现情况,结合其研究点,将它们分为四类,在表 3 中用短实线隔开。

表 5 新的主题词 (21)

技术	WEB 2.0 WIKIS (Computer science) SOCIAL networks SOCIAL media VIRTUAL communities VIRTUAL reality AVATARS (Virtual reality) SECOND Life (Game)
组织	COMPETITION (Economics) VALUE capture VALUE chains OFFSHORE outsourcing CONTRACTING out -- Economic aspects ORGANIZATIONAL performance
个人	WEB design COMPLIANCE NONCOMPLIANCE AMAZON.COM Inc.
方法	DESIGN science STATISTICAL reliability RESEARCH methodology evaluation

第一类代表着近年来出现的新技术、新应用和新模式,如在 Web 2.0 等技术的支持

下,社交媒体、虚拟社区等新型应用模式成为了关注热点。虚拟现实、第二人生等技术也成为研究人员的兴趣。第二类体现为从组织的层面入手的主题,例如竞争、价值链、离岸外包、合同外包等,这些代表着研究人员关注的企业内的新问题。第三是个人有关的研究,包括 web 设计相关以及信息安全策略的配合和不配合行为等。最后是研究方法有关的 3 个主题词,反映了研究人员对方法越来越重视,并不断吸收和采纳 Design science 等方法。

据此对近 5 年半的 276 篇文章进行分类,如果一篇文章提供的主题词包括了在某一类中的主题词,就把它标上该类别的标签。对于涉及 AMAZON.COM Inc.的 5 篇文章,鉴于 AMAZON 的数据被广泛的用于技术、企业等多方面的研究,对其标签进行了人工修正。

按照上述说明,选出了 99 篇文章,占近 5 年半文章总数的 36%。文章在各类别、各年的分布请见图 2 和图 3,可见较为均匀的覆盖了各类别、各年份,且保证了最近年份的比重。一篇文章可能在多个类别中出现。结果显示,多类别标签的文章共有 12 篇,占文章总数的 12%。

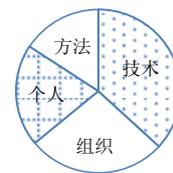


图 2. 文章在各类中的分布

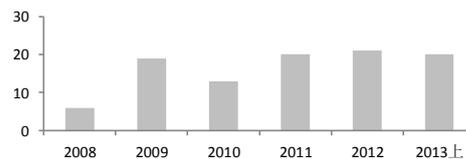


图 3. 各年选出的文章数

3.3. 新兴主题的词频分析

通常,同一类主题的研究在用语上有一定的共同之处,这些用语从一定程度上反映了相关研究的背景、着眼点、方法等信息。本小节对上述四类文章的摘要进行词频分析。

具体地，针对每一类的文章统计词频，包括每年的频度和总频度，以及每年包含该词的文章数量。然后，根据 2013 年的频度 F_{2013} 和总频度 F 分别对词语进行排序，综合考虑二者的排名，选出前 10 个高频词。调整权重，观察排名，结果显示比较稳定，最终设定二者的权重分别为 0.8 和 0.2。需要说明的是，在高频词的选取中回避了在各类中均高频出现但不与主题直接相关的词语（如 effect, study），以及只出现在极个别文章中的高频词。

表 6 列出了新技术相关研究中的前十高频词。可以发现，社交和 wiki 相关的技术和模式最受关注，社群——尤其是借助社交平台形成的社群是热点之一，在这些背景下的内容和知识也得到重视。同时，相关研究不仅关注用户，还关注与商业的关联。

表 6 技术相关研究中的高频词

social	community	user	
business	knowledge	wiki	
website	strategy	content	online

表 7 显示了组织/企业相关研究中的前十高频词。相关研究从公司、行业甚至商业（模式）的角度展开，它们关注组织/企业在数字环境下的相关战略、价值相关的活动、IT 相关的能力及流程等。

表 7 组织相关研究中的高频词

strategy	business	firm	digital
value	industry	capability	
technology		process	design

表 8 显示了个人相关研究中的前十高频词。相关研究关注的是个人的行为、动机及受其它因素（如环境）的影响。在社交平台下的研究比较多。研究较多的除了采纳相关问题，还有安全政策的配合/不配合问题。相关研究还较多的关注了跨背景下的对比，例如跨产品、跨网站、跨组织、跨文化的对比研究。

表 8 个人相关研究中的高频词

influence	security (policy)	individual	
social	behavior	process	
employee	motivation	adopt	across

表 9 展示了方法相关研究中的前五高频词。相关研究中讨论较多的有模型、数据、方法和理论等。除了对方法的严谨性进行核实，还引入了 design science 的方法。这也体现出在大数据时代信息系统学科对常规方法与 design science 等方法配合运用的需求。

表 9 方法相关研究中的高频词

design(science)	model
data	method theory

3.4. 研究启示

基于上述分析的初步结果，我们总结出如下几点启示。

首先，从新兴主题可以清晰地看到不同的关注点和研究层面（技术、组织、个人、方法），这种划分与前人的工作基本一致，说明在层面上相对稳定，但是具体的关注点发生了迁移。例如在技术相关研究中，[8]提到，2000 之前最受关注的 DSS 在 2002-2006 年间被（信息系统开发中的）合作、风险管理、培训等取代，如今的关注点已转移到互联网相关技术，聚焦在社交和 wiki、虚拟化技术上。互联网上的海量的内容（content）是备受瞩目的信息资源。再如，在个人相关研究中，在 2002-2006 年间被关注的技术接受[8]虽然仍有相关研究，但是热度已经稍减；更加吸引人的是社会化网络环境中的个人行为与动机，以及在具体的情境下的行为及对比研究。

第二，互联网还给我们带来了丰富的、可以方便获得的数据，像 AMAZON.com 的数据就多次出现在相关研究中，还有在互联网上的大数据、用户生成内容等。这些一方面为研究人员提供了方便，另一方面也提出了新的问题和挑战。虽然相关研究还没有在 MISQ 的文章里集中出现，但从管理学的角度应用大数据技术以支持商业分析和决策，以数据驱动为主导的金融、市场、战略、营销和运作管理研究和实践指导，在未来有着广阔的发展空间[10]。

4. 结语

审视信息系统领域的发展和趋势，能够

剖析研究领域和学科的核心问题和发展趋势，同时为研究人员提供参考和借鉴。本文基于对 MIS Quarterly 期刊上最近 30 年来发表的文献的分析，总结了信息系统领域研究主题的四种变化趋势：衰退型、屋顶型、平稳型、上升型。进一步的，本文抽取近五年新兴的主题，并归纳为四类，即技术类、组织类、个人类、方法类。通过分别提取对应文章的摘要中的高频词，我们总结了四类研究的关注热点。

本文也存在一些局限。因为是从期刊文献中找趋势，是总结性的而非预测性的，适合对之前的研究热点和变动趋势做总结，而难以对未来研究热点进行准确预测。同时，由于期刊发表周期的存在，所总结的主题发展趋势不可避免的存在时滞。对未来研究热点的预测需要结合技术、行业、市场和社会的大环境识别新的需求热点。本文所总结的新兴主题是那些新出现的、并且已经显现一定的规模的主题。当然，这些主题的一部分还将在一定时间内延续、甚至继续崛起。

目前文章给出了主题分析的初步结果，还有许多问题值得下一步研究，包括在更多的维度上对文献进行对比分析、使用语义分析方法对研究主题进行探索等。

参考文献

- [1] Alavi, M. and Carlson, P. A Review of MIS Research and Disciplinary Development, *Journal of Management Information Systems*, 1992, Vol 8, No 4, pp. 45-62.
- [2] Berelson, Bernard. (1952) Content analysis

in communication research. New York, NY, US: Free Press.

- [3] Banker, Rajiv D., Kauffman, Robert J. The Evolution of Research on Information Systems: A Fiftieth-Year Survey of the Literature in Management Science, *Management Science*, Vol 50, No 3, 2004, pp. 281-298.
- [4] Deerwester, S., Dumais, S., Furnas, G., Landauer, T., and Harshman, R. 1990. Indexing by Latent Semantic Analysis, *Journal of the American Society for Information Science* (41:6), pp. 391-407.
- [5] Lee, Sang M., Hwang, Taewon and Kim, Jaekyung. An Analysis of Diversity in Electronic Commerce Research, *International Journal of Electronic Commerce*, 2007, Vol 12, No 1, pp.31-67.
- [6] Rainer, Jr., R.K., and Miller, M.D. Examining References Across Journal Rankings, *Communications of the ACM*, 2005, Vol 48, No 2, pp.91-94.
- [7] Robey, D. Research commentary: Diversity in information systems research: Threat, promise, and responsibility. *Information Systems Research*, 1996, Vol 7, No 4, pp.400-408.
- [8] Sidorova, A., Evangelopoulos, N., Valacich, J.S. and Ramakrishnan, T. Uncovering the intellectual core of the information systems discipline, *MIS Quarterly*, 2008, Vol 32, No 3, pp.467-482.
- [9] Zhang, P., and Scialdone, M.J. State of IT Artifacts: An Analysis of ICIS 2009 Research Papers, *Proceedings of the PACIS2010*, July 2010, Taipei, Taiwan.
- [10] 冯芷艳, 郭迅华, 曾大军, 陈煜波, 陈国青, 大数据背景下商务管理研究若干前沿课题, *管理科学学报*, 第 16 卷第 1 期, 2013, pp.1-9.

On the Change of the Subject Terms in Information System Studies

-- Based on a literature analysis of MIS Quarterly

Ming Ren¹, Jin Zhang², Baojun Ma³

¹School of Information Resource Management, Renmin University of China, Beijing, 100872;

²School of Business, Renmin University of China, Beijing, 100872; ³School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing, 100876)

Abstract The information system studies continue to develop and evolve, resulting in changing subjects. An analysis of these subjects will help to understand the intellectual core of the discipline as well as its evolvement, and thus provide guidance for researchers in this field. A literature analysis of MIS Quarterly is conducted, and the emerging subject terms and the fading ones are extracted. Then the emerging subject terms are divided into four categories, and the focus of each category is discussed based on the frequent words in the abstracts of the literatures.

Keywords Information System; Subject term; MIS Quarterly; Content analysis